Rescue and regulation of proboscipedia: a homeotic gene of the Antennapedia Complex

FILIPPO M. RANDAZZO, DAVID L. CRIBBS* and THOMAS C. KAUFMAN

Howard Hughes Medical Institute, Institute for Cellular and Molecular Biology, and Programs in Genetics and Cellular, Molecular, and Developmental Biology, Department of Biology, Indiana University, Bloomington, Indiana, U.S.A.

*Present address: Centre de Recherche de Biochimie et Genetique Cellulaires, 118, route de Narbonne 31062 Toulouse Cedex, France

Summary

The extraordinary positional conservation of the homeotic genes within the Antennapedia and the Bithorax Complexes (ANT-C and BX-C) in Drosophila melanogaster and the murine Hox and human HOX clusters of genes can be interpreted as a reflection of functional necessity. The homeotic gene proboscipedia (pb) resides within the ANT-C, and its sequence is related to that of Hox-1.5. We show that two independent pb minigene P-element insertion lines completely rescue the labial palp-to-first leg homeotic transformation caused by pb null mutations; thus, a homeotic gene of the ANT-C can properly carry out its homeotic function outside of the complex. Despite the complete rescue of the null, the minigene expresses pb protein in only a subset of pb's normal domains of expression. Therefore, the biological significance of the excluded expression pattern elements remains unclear except to note they appear unnecessary for specifying normal labial identity. Additionally, by using reporter gene constructs inserted into the Drosophila genome and by comparing pb-associated genomic sequences from two divergent species, we have located cis-acting regulatory elements required for pb expression in embryos and larvae.

Key words: proboscipedia, homeotic gene, Antennapedia Complex, rescue, regulation, sequence conservation.

Introduction

Determining how the three-dimensional architecture of an organism results from information encoded in the linear DNA molecule is a primary goal of developmental geneticists. The process of segmentation and homeosis in Drosophila melanogaster defines a paradigm that addresses aspects of this question. The segmented body-plan of Drosophila derives from a complex hierarchy of coordinately expressed genes. Enumerating the primary metameric pattern of the embryonic germ band involves a repertoire of genes classified into four groups according to their function: (a) the coordinate genes, (b) the gap genes, (c) the pair rule genes, and (d) the segment polarity genes (Nüsslein-Volhard and Wieschaus, 1980; Nüsslein-Volhard et al. 1987). The products of the coordinate genes are laid down by the mother and set up the axial polarity of the organism. The gap genes divide the embryo into sets of contiguous segments, whereas the pair rule genes divide the embryo into alternating segments or parasegments. Finally, the polarity of pattern elements within each segment are set up by the segment polarity genes. The actions of each class occur as a cascade of dependent events and taken together specify the proper number of segments in the head, thorax and abdomen. The homeotic genes then interpret this positional information and act as ontogenetic switches to specify differing segmental identities (García-Bellido, 1977; Lewis, 1978).

Eight of these homeotic genes in Drosophila cluster into two separate complexes, the Antennapedia Complex (ANT-C) and the Bithorax Complex (BX-C) (Kaufman et al. 1980, 1989; Lewis, 1978). The BX-C contains three homeotic genes that specify thoracic and abdominal segment identity: Ultrabithorax (Ubx), abdominal A (abdA) and Abdominal B (AbdB) (Lewis, 1978; Sánchez-Herrero et al. 1985; Peifer et al. 1987). The ANT-C includes five homeotic genes which specify head and thoracic segment identity: Antennapedia (Antp), Sex combs reduced (Scr), Deformed (Dfd), proboscipedia (pb) and labial (lab) (Denell et al. 1981; Struhl, 1981; Abbott and Kaufman, 1986; Mahaffey and Kaufman, 1987; Merrill et al. 1987; Regulski et al. 1987; Pultz et al. 1988; Diederich et al. 1989). A number of non-homeotic genes also reside within the ANT-C, including the pair rule gene fushi tarazu (ftz), the maternal effect gene bicoid (bcd) which encodes a morphogen important in the specification of anterior-posterior polarity, and the zygotic gene zerknüllt (zen) which is required for the specification of dorsal structures (Wakimoto et al. 1984; Driever and Nüsslein-
Volhard, 1988a, 1988b; Rushlow et al. 1987). The ANT-C also includes two transcription units with no genetically discernible function – the IgA homology-containing gene amalgam (ama) and the zen-pattern-like gene \( z^2 \) – and a cluster of eight transcripts which show sequence similarity to cuticle genes (Seeger et al. 1988; Rushlow et al. 1987; Pultz et al. 1988; Fechtel et al. 1988; Pultz, 1988).

Mutations in the homeotic loci result in the replacement of one or several segmentally derived structures by those associated with another metamere. Null proboscipedia mutations, for example, cause a distriproboscis-to-leg transformation while hypomorphic \( pb \) alleles lead to a distriproboscis-to-antennal transformation (Bridges and Dobzhansky, 1933; Kaufman, 1978).

No other homeotic gene displays a dosage-dependent transformation to one of two differing tissues. Additionally, loss of function \( pb \) alleles result in a size reduction of the maxillary palps, which has been interpreted as a transformation toward antennal identity (Kaufman, 1978). The \( pb \) gene product is not required for adult viability and its absence does not result in any visible changes in the embryo, an unusual characteristic for a homeotic gene. As with all other homeotic genes, \( pb \) presumably determines segment identity by controlling the expression of a number of ‘downstream’ or ‘realisor’ genes which either directly or indirectly control the differentiation of tissues (García-Bellido, 1977).

Homeotic proteins are localized in the cell nucleus, consistent with their inferred gene-regulatory functions. These proteins possess a highly conserved 60 amino acid sequence called the homeodomain (McGinnis et al. 1984; Scott and Weiner, 1984), which has DNA binding properties (Desplan et al. 1985; Müller et al. 1988). Additionally, tissue culture experiments demonstrate that Ultrabithorax and Antennapedia proteins, both of which are homeobox-containing, can activate and/or inactivate transcription of target genes in vivo in a manner generally compatible with previous genetic analysis (Krasnow et al. 1989; Winslow et al. 1989). Thus, homeotic genes encode DNA binding proteins that are presumably capable of activating and/or inactivating transcription of other target genes.

Molecular homologues to the homeotic genes of the ANT-C and the BX-C are found in a variety of other species. Results from a chromosome walk in \( D. pseudoobscura \) (Randazzo, Seeger, Huss and Kaufman, in prep.) indicate that homologues to the ANT-C genes lie in the same linear order on the chromosome as in \( D. melanogaster \). Furthermore, genetic analysis of the red flour beetle Tribolium castaneum demonstrates a closely ordered and similarly aligned set of ANT-C-like and BX-C-like genes referred to as the HOM-C (Beeman et al. 1989). Considerably more remarkable, the murine Hox and human HOX gene complexes contain clusters of genes with homeobox sequences strikingly similar to those of the homeotic genes of the BX-C and the ANT-C (Duboule and Döllé, 1989; Graham et al. 1989; Kappen et al. 1989; Boncinelli et al. 1988). The genes of the Hox complexes are arranged in the same linear order and relative transcription direction on the chromosome as the BX-C and ANT-C homeotic genes. Moreover, the spatial expression domains in the CNS reflect the linear array of the genes in mammals as in insects. This extraordinary conservation of linkage, linear array, and spatial expression pattern suggests the possibility that there is some necessary functional coupling of the genes of the ANT-C, BX-C, HOM-C and Hox complexes.

Because of the substantial size and complexity of the homeotic genes, rescue of homeotic function has in practice been difficult. However, mutations in some of the smaller non-homeotic genes in the ANT-C – \( bcd, \) \( zen \) and \( ftz \) – have been rescued by \( P \)-element mediated germline transformation (Berleth et al. 1988; Rushlow et al. 1987; Hiromi et al. 1985). We set out to examine the expression and function of a proboscipedia minigene with several goals in mind: (1) to prove that the identified \( pb \) DNA corresponds to the genetically defined \( pb \) gene; (2) to determine whether a homeotic gene that normally is located in the ANT-C must reside there to function properly; and (3) to determine the relationship between the expression of \( pb \) protein in specific tissues to its function in those tissues. In this paper, we present our progress toward these goals by demonstrating that the \( pb \) homeotic phenotype may be rescued by a \( pb \) gene located outside the ANT-C. Additionally, through the use of \( pb/i acZ \) reporter fusion genes transformed into Drosophila and DNA sequence comparison between \( D. melanogaster \) and \( D. pseudoobscura \), we have located cis-acting regulatory elements necessary for \( pb \) expression.

Materials and methods

\( pb \) minigene construction

The minigene was constructed by inserting a restriction fragment containing \( pb \) upstream, promoter and RNA leader sequences into an existing hsp70 promoter/\( pb \) gene fusion construct (Cribbs, Randazzo and Kaufman, in prep.). The hsp70 promoter/\( pb \) gene fusion construct consists of the mini-gene (an EcoRI fragment) which is a derivative of Bm \( \Delta \)-W (Pirrotta et al. 1985) inserted into the Xhol site of pUCHneo vector (Steller and Pirrotta, 1985) from which the neo sequences have been deleted. Genomic \( pb \) sequences were inserted into this vector such that the hsp70 RNA leader was fused to the downstream half of the \( pb \) exon-1 RNA leader. In order to make the minigene, a \( pb \) genomic restriction fragment containing 7.3 kb of upstream sequence and the upstream half of the exon-1 extending to an NcoI restriction site was inserted into the heat shock/\( pb \) vector at a unique Xbal site created by insertion of a linker at the hsp70 RNA leader/\( pb \) exon-1 RNA leader junction. These manipulations lead to a net insertion of 27 nt (nucleotides) between CCATG and G of the NcoI site in the first exon of \( pb \). The inserted sequence reads as follows: GTCTAGAGCGG-CCATG and G of the Ncol site in the first exon of \( pb \).

Additionally, as a result of the manipulations described above, the \( pb \) minigene contains 191 nt of hsp70 5' flanking sequences and the first 92 nt of the hsp70 RNA leader. These sequences are 7.3 kb and 5.3 kb upstream of the \( pb \) and \( z^2 \) transcription start sites, respectively. Flies of the genotype \( P[w^\#, pb]_1/P[w^\#, pb]_1; pb^*/pb^b \) are completely rescued in
terms of the homeotic transformation when grown at 18°C. A stock of P[w\(^+\), pb] homozygous flies was grown at 29°C to see if mild heat shock has any dominant pb-associated effects on the animals. When hsp70 promoter/pb gene fusion lines are grown at 29°C, thickening of aristae, transformation of antennae to maxillary palps and increased pupal lethality phenotypes are observed (Cribbs, Randazzo and Kaufman, in prep.). These phenotypes are inferred to be due to ectopic expression of pb protein. None of these changes are observed in P[w\(^+\), pb] homozygous flies when cultured at 29°C. Therefore, it appears that the hsp70 sequences in the construct do not detectably affect the pb portion of the minigene.

pb/lacZ reporter gene fusion construction

The pbG2 construct was made by first removing the ninaE sequences from pDM66A (kindly provided by D. Misra) as a KpnI/BamHI fragment, blunt-ending with T4 DNA polymerase, and then blunt end ligating the vector to recreate the BamHI site. A 7.8 kb fragment, beginning at the BamHI site 7.3 kb upstream of the pb transcription initiation site and ending at an NcoI site within the first exon of pb which was modified into a BamHI site using linkers, was cloned into the BamHI site of the modified pDM66A vector. This plasmid was then cut with NotI to separate the pHSS7 sequences from the pb/lacZ cassette that was inserted into the NotI site of a CoSpeR NotI transformation vector (kindly provided by J. Tamkun) which was derived from the original CaSpeR vector (Pirrotta, 1988).

The pbG1 construct was made by inserting pb sequences beginning at the BamHI site that lies downstream of the ATG translation start site in the second pb exon to the BamHI site 7.3 kb upstream of the pb transcription start site into the pCaSpeR-\(^{\beta}\)gal vector (Thummel et al. 1988). Next, an oligonucleotide fragment with BamHI sticky ends was inserted into the downstream BamHI site. This manipulation resulted in the in-frame fusion of pb coding sequences to lacZ coding sequences. The pCaSpeR-\(^{\beta}\)gal vector contains P-element ends and the mini-white gene as the visible selectable marker. The pbG1/intron construct was made by inserting a 1.6 EcoRI fragment containing the pb intron-2 conserved elements from D. melanogaster into an EcoRI restriction site in the 2\(^{\beta}\) transcrip of the pbG1 construct.

Culture of Drosophila

Flies were cultured at 25°C on standard Drosophila media supplemented with baker’s yeast. In(3R)MAP17 is a V-type position effect pb mutation that is suppressed at 29°C (Pultz, 1988). Therefore, homozygous In(3R)MAP17 flies were cultured at 29°C.

P-element transformation

Standard P-element mediated germline transformation techniques were utilized for obtaining independent insertion lines of the pb minigene and both the pbG1 and pbG2 constructs (Robertson et al. 1988). All reported transformants derived independently from separate injections, except that pbG2 line B' was obtained by jumping pbG2 line A' onto another chromosome using the P[\(^{\Delta}2\)Δ-3(99B)] element (Robertson et al. 1988).

Rescue of pb

Flies of the genotype w; P[w\(^+\), pb]/+; Ki pb\(^{p/2}\) TM3 Sb were inbred and the four expected classes of progeny were scored. Homozygous pb\(^p\) flies were scored by both the presence of Ki/Ki phenotype and the absence of Sb. Animals carrying the minigene had yellow eyes and their siblings, lacking the minigene, had white eyes. Presence of the Sb phenotype is indicative of the pb\(^{b}\) heterozygous class of flies. pb\(^{b}\) null stocks carrying either the P[w\(^+\), pb] insertion or the P[w\(^+\), pb] insertional phenotype have been propagated for numerous generations. Hundreds of flies have been observed and all have displayed a completely rescued homeotic phenotype. Flies were stored in 70% ethanol before preparation for SEM. Specimen preparation and SEM methods have been described by others (Merrill et al. 1987).

Antibody preparation

A \(\beta\)-galactosidase/pb fusion construct was made using a pWR590-1 \(\beta\)-galactosidase fusion cassette (Guo et al. 1984), which incorporated the 5\(a\) pb cDNA (Cribbs, Pultz, Johnson and Kaufman, in prep.). The fusion protein was expressed in JM101 E. coli cells, extracted by standard Laemmli techniques (Mahlaffey and Kaufman, 1987), isolated on a standard 8% SDS-PAGE gel, and then injected into rabbits. A pATH11/pb 5a cDNA fusion construct was used to make a TrpE/pb fusion protein (Dieckmann and Tzagoloff, 1985). The fusion protein was coupled to Affi-Gel 10 (Bio-Rad) in order to make an affinity chromatography column. Finally, the terminal bleed was purified over the TrpE/pb column twice and then stored at 4°C.

Antibody staining

Collection, fixation, and staining of embryos were performed using standard techniques (Mahlaffey and Kaufman, 1987). The secondary antibody used for the staining of wild-type and P[w\(^+\), pb]; pb\(^{b}\)/pb\(^{b}\) embryos was EAL grade affinity purified polyclonal goat anti-rabbit antibodies conjugated to horse-radish peroxidase (HRP) used at a 1:150 dilution (Bio-Rad). Embryos from the pbG1, pbG2, and pbG1/intron stocks plus a P[r\(^{\Delta}2\)]/[\(\alpha\)/lacA] positive control (Hiromi et al. 1985) and a w\(^{+}\)/w negative control stock which lacks a \(\beta\)-galactosidase containing construct were collected, fixed, and then immunostained using rabbit polyclonal antibodies directed against the \(\beta\)-galactosidase protein and goat anti-rabbit HRP conjugated secondary antibodies. The positive control exhibited the expected \(\alpha\)-like pattern of expression plus hindgut staining. The hindgut staining pattern is a characteristic of this \(\beta\)-galactosidase antisera.

Third instar larvae were stained according to the techniques of Angela Pattatucci (Pattatucci and Kaufman, 1991). FAB' goat anti-rabbit HRP conjugate (1:50 dilution) was the secondary antibody used for the third instar larval staining (Protos Immuno Research). Stained embryos and imaginal disks were photographed using a Zeiss Axioshot with Nomarski optics except for Figs 5D-G which were photographed under normal light optics and Figs 4A-D which were photographed under Nomarski optics using a Zeiss Microscope III.

DNA sequencing

The restriction fragments were subcloned into either mp18, mp19, mWB238, or mWB239 vectors (Barnes et al. 1983). Single stranded template was prepared using standard techniques as described in the Sequenase\textsuperscript{®} Version 2.0 kit (U.S. Biochemical Corp.). The sequencing reactions were performed using such a kit.

Results

A pb minigene can completely rescue the pb homeotic phenotype

Because the length of DNA contained within pb's
genetic boundaries approximates to the technical limits of P-element germline transformation, a minigene rather than a full sized transgene was utilized (Fig. 1). The choice of the pb sequences included in the minigene was based on previous molecular and genetic analysis (Pultz et al. 1988; Cribbs, Pultz, Johnson and Kaufman, in prep.). Searches for cDNAs representative of pb mRNA in standard libraries to date have not been successful (Cribbs, Pultz, Johnson and Kaufman, in prep.). Therefore, three restriction fragments of genomic DNA encompassing sequences inferred to be relevant for pb function were used to construct the minigene.

The proximal and distal genetic boundaries of pb lie between the In(3R)MAP17 breakpoint and the In(3R)lab*76 breakpoint, respectively (Pultz et al. 1988). Both of these mutations fully complement pb deficiencies and, therefore, define the maximum size of the pb gene. The proximal breakpoint of In(3R)MAP17 is approximately 7 kb upstream of the pb transcription start site and the minigene includes all of the 5' sequences defined by In(3R)MAP17. The genetically defined 3' boundary of pb is In(3R)lab*76, which breaks in the labial transcription unit approximately 34 kb downstream of the most 3' detectable pb sequences (Pultz et al. 1988). The construction excludes all of the sequences downstream of the first polyadenylation site of pb (Cribbs, Pultz, Johnson and Kaufman, in prep.). The transgene excludes most of the large intron composing the central 25 kb of the gene and consequently an alternatively spliced 15 bp microexon, positioned in the middle of the 25 kb intron, which was found in most pb cDNAs generated by PCR amplification (Cribbs, Pultz, Johnson and Kaufman, in prep.). The final minigene thus includes all those sequences corresponding to pb mRNA, as defined by northern blot, S1 nuclease protection, and DNA sequence analysis with the single exception of the 15 bp microexon. Closely associated exon flanking sequences are also included (Pultz et al. 1988; Cribbs, Pultz, Johnson and Kaufman, in prep.). The minigene also contains a 27 bp oligonucleotide insertion in the middle of the pb 5' RNA leader and some hsp70 promoter sequences that lie 7.3 kb upstream of the pb transcription start site (see Materials and methods for details).

Introduction of this construct into the Drosophila genome via P-element mediated germline transformation resulted in the recovery of three independent insertion lines: P[w+, pb]1, a homozygous viable insertion located on the second chromosome; P[w+, pb]2, a recessive semi-lethal also located on the second chromosome (we interpret the semi-lethality as being due to a disruption of the chromosomal locus of insertion); and P[w+, pb]3, a male semi-lethal located on the X chromosome.

The following experiments demonstrated that the pb minigene lines 1 and 2 are functionally competent. Stocks were constructed bearing one of the three pb

![Fig. 1](image_url)

The diagram represents the molecular map of the pb region in D. melanogaster. Darkened boxes depict protein coding sequences, whereas white boxes delineate noncoding mRNA sequences. Asterisk(s) designate homeobox sequences. The darkened stippled bar below the molecular map represents the inversion In(3R)MAP17 which defines the proximal boundary of the pb gene; the adjacent cross hatches illustrate the range of uncertainty with respect to the breakpoint. The two black bars directly above the molecular map overlie the pb genomic sequences contained within the minigene. The linear diagram above outlines the entire minigene construct. It consists of two genomic fragments containing the 5' and 3' clusters of pb exons. The construct excludes most of the large intron and the alternatively spliced microexon. The transgene therefore includes the z2 pseudogene. The pb minigene also contains the mini-white gene as a visible marker, P-element ends, and pUC8 sequences (see Methods). As a consequence of its derivation from an hsp70 promoter/pb transgene, the minigene construct contains part of the hsp70 promoter at a location 7.3 kb upstream of the pb transcription start site and a 27 nt adapter sequence inserted in the middle of the pb RNA leader (see Methods).
minigene lines in a pb^5 homozygous (pb-null) background (see Materials and methods for details). Fig. 2A shows a fly with wild-type mouthparts. This structure has been described previously in detail (Hodgkin and Bryant, 1978). The mouthparts include the anterior labellum (labial palps) which contains six pseudotracheal rows on each side. The dorsal half of each pseudotrachea is smooth while the ventral half is convoluted. The posterior end of the labellum contains a characteristic set of bristles. Fig. 2B illustrates a
typical \( pb^5 \) homozygous animal possessing prothoracic legs in place of the labial palps. Flies carrying one \( P[w^+, pb]_1 \)-bearing chromosome in a \( pb^2 \)-homozygous background display completely normal mouthparts (Fig. 2C, compare to 2A). The six pseudotracheal rows on each labella form perfectly, the bristle patterns develop normally, and no indication of tissue transformation is observed. Likewise, the \( P[w^+, pb]_2 \) insertion line completely rescues the \( pb^5 \) homeotic phenotype. Typical \( pb^2/pb^5 \) adults also display a reduced maxillary palp phenotype. Presence of one copy of either the \( P[w^+, pb]_1 \)-bearing chromosome or the \( P[w^+, pb]_2 \) chromosome in the \( pb^6 \)-null background rescues the overall maxillary palp morphology, although in some animals the maxillary palps remain bent at the base.

The third insertion line, \( P[w^+, pb]_3 \), shows only partial rescue of the \( pb^5 \) homeotic phenotype. Unlike the ectopic leg phenotype seen in \( pb \) null animals, hypomorphic alleles of \( pb \) result in a transformation of distiproboscis to antennae (compare Figs 2B and 3A). Fig. 3B exemplifies the effect of one \( P[w^+, pb]_3 \) chromosome in a \( pb^6 \)-null background. The observed phenotype clearly resembles that produced by \( pb \) hypomorphic alleles. In addition to the above effects, flies carrying the \( P[w^+, pb]_3 \) insertion chromosome exhibit thickened aristae, a male associated wings-held-straight-out phenotype, and male-semi-lethality. The presence of thickened aristae also is observed also in a number of \( hsp70 \) promoter/\( pb \) gene fusion lines and has been interpreted in those cases to be a result of misexpression of \( pb \) protein in the antennal anlagen (Cribbs, Randazzo and Kaufman, in prep.). Thus, the \( P[w^+, pb]_3 \) insertion may be misexpressing the \( pb \) transgene protein in the antennal anlagen, resulting in the thickened aristae phenotype. Subsequent transposition of the \( P[w^+, pb]_3 \) transgene onto another chromosome resulted in simultaneous reversion of the three above mentioned phenotypes. This transposed minigene, however, still only partially rescues the homeotic phenotype. Therefore, we suspect that the \( P[w^+, pb]_3 \) transgene may have been damaged during the original insertion event, resulting in reduced \( pb \) function and the consequent partial rescue of the mutant phenotype. However, the most likely cause of the apparent misregulation is a position effect in the original line which can be ameliorated by transposition to a new site.

The \( pb \) transgene protein distribution pattern is a subset of the normal \( pb \) pattern in embryos as well as in third instar larvae.

In order to determine whether the \( pb \) transgene is properly regulated, its derived embryonic and larval protein localization pattern was examined. The \( pb \) transgene protein exhibited a wild-type pattern of expression in the labial, maxillary and mandibular lobes of embryos and in the labial disks of third instar larvae. On the other hand, we found that aspects of \( pb^5 \)'s neural and mesodermal pattern in embryos and \( pb^5 \)'s neural pattern in larvae were either significantly downregulated or absent.

Embryos from a \( pb^2/TM3 \) stock were probed with the \( pb \) antisera to determine whether \( pb^2/pb^5 \) embryos expressed detectable amounts of \( pb \) protein. Twenty-five percent of the embryos were expected to possess the \( pb^2/pb^5 \) genotype and approximately 25% (9 of 42) lacked any \( pb \) protein staining while all of the remaining embryos (33 of 42) exhibited the wild-type \( pb \) staining pattern (mid-stage 10–stage 16 embryos); thus, it appears that the \( pb^2 \) allele is a protein null, at least with respect to the antisera used in this analysis.

Embryos obtained from a \( P[w^+, pb]_1/\ P[w^+, pb]_1; pb^5/pb^5 \) and a \( P[w^+, pb]_2/ +; pb^5/TM6B \) stock were probed with the \( pb \) polyclonal antibodies in order to determine the \( pb \) transgenic protein distribution pattern. Because the \( pb^5 \) allele does not encode a protein product detectable with this \( pb \) antisera, the observed pattern exclusively reflects \( pb \) protein accumulation resulting from expression of the transgene. The wild-type embryonic pattern of \( pb \) expression has been

![Fig. 3. (A) Scanning electron micrograph (SEM) preparation of the head of a \( Ki pb^5/Ki pb^5 \) adult animal focusing in on the mouthparts. \( pb^5 \) is a null allele while \( pb^6 \) is a hypomorphic allele. The transformation of mouthparts is to aristae (a). (B) SEM showing the mouthparts of a typical partially rescued adult animal. The genotype of this female is \( P[w^+, pb]_3/+; Ki pb^5/Ki pb^5 \).](image-url)
described previously (Pultz et al. 1988). In wild-type embryos, pb protein is not detectable until the beginning of stage 10 (approx. 4 h 20 min) (embryonic stages established by Campos-Ortega and Hartenstein, 1985). During stage 10 (4 h 20 min–5 h 20 min), the protein appears abruptly in a set of mesodermal cells that reside anterior to the maxillary lobes and span transversely across the embryo as a triangular belt (Fig. 4A). The protein derived from the transgene, on the other hand, seems to accumulate at significantly lower levels in those mesodermal cells. Furthermore, the P[w\textsuperscript{+}, pb]\textsuperscript{1} and P[w\textsuperscript{+}, pb]\textsuperscript{2} transgene-derived proteins subsequently become undetectable, while in wild-type animals the level of protein accumulation is maintained (compare Fig. 4B with 4A).

The start of stage 10 (4 h 20 min) is marked by the beginning of the stomodeal invagination which occurs ventrally near the anterior end of the embryo. The stomodeum gives rise to the larval mouth. During stage 10 in wild-type embryos just after the onset of the mesodermal pb expression, the pb protein begins to gradually accumulate in the labial and maxillary lobes and a small set of ectodermal cells in the mandibular lobe all of which will eventually contribute to parts of the larval head structure. Protein from the pb transgene expressed by the P[w\textsuperscript{+}, pb]\textsuperscript{1} line, however, becomes detectable in the labial and maxillary lobes and 5 or 6 ectodermal cells of each mandibular lobe slightly earlier in stage 10 when the stomodeal invagination is shorter (data not shown).

Wild type embryos accumulate pb protein in a small subset of ventral nerve chord cells in a metamerically arranged fashion. Normal pb expression in the ventral nerve chord first becomes detectable after head involution starts at the beginning of stage 14 (10 h 20 min). During stage 15 (11 h 20 min–13 h) the labial lobes migrate anteroventrally and fuse on the ventral midline, prior to their involution. At this time the number of ventral nerve chord cells with detectable pb expression, positioned close to the ventral midline, has increased. After the labial lobes involute and the cells of the maxillary lobes move near the anterior pole of the stage 16 embryo (13–16 h), additional ventral nerve chord cells accumulate detectable amounts of pb protein more laterally from the median. In contrast to the wild-type pattern, pb protein expression in homozygous P[w\textsuperscript{+}, pb]\textsuperscript{1}; pb\textsuperscript{b} embryos was not detected in the ventral nerve chord at any time during embryonic development (compare Fig. 4C with 4D). Elsewhere, in the central nervous system, a small number of cells in the supraesophageal and subesophageal ganglia express detectable amounts of pb protein in both rescued and pb\textsuperscript{b} embryos. The pattern of protein accumulation, while in the remaining areas pb protein seems to be localized in closely paired sets of cells. In two rescued lines, P[w\textsuperscript{+}, pb]\textsuperscript{1}; P[w\textsuperscript{+}, pb]\textsuperscript{1}; pb\textsuperscript{b}/pb\textsuperscript{b} (Fig. 5F) and P[w\textsuperscript{+}, pb]\textsuperscript{2}; pb\textsuperscript{b}/pb\textsuperscript{b} third instar larvae fail to detectably accumulate pb protein in the ventral ganglion except for 5–8 cells in the most anterior region. These results are consistent with the absence of ventral nerve chord staining in embryos. On the other hand, pb protein does accumulate in what appears to be the same cells of the supraesophageal and subesophageal ganglia where pb is normally expressed. The levels of protein accumulation, however, are greatly reduced (compare Figs 5D and 5F).

In order to determine whether the transgene protein expression was under proper negative regulatory controls, we examined third instar larvae of the following genotypes: (1) P[w\textsuperscript{+}, pb]\textsuperscript{1}/ P[w\textsuperscript{+}, pb]\textsuperscript{1}; pb\textsuperscript{b}/pb\textsuperscript{b}, (2) P[w\textsuperscript{+}, pb]\textsuperscript{2}/ P[w\textsuperscript{+}, pb]\textsuperscript{2}; pb\textsuperscript{b}/pb\textsuperscript{b}, and (3) P[w\textsuperscript{+}, pb]\textsuperscript{3}/ P[w\textsuperscript{+}, pb]\textsuperscript{3}. Ectopic expression of trans-
gene derived pb protein was not detectable in these lines (data not shown). We observed the pb protein accumulation pattern in third instar larvae homozygous for the In(3R)MAP17 mutation to determine if sequences upstream of the proximal breakpoint are necessary for any portions of the pb pattern. We could not detect any missing aspects from the wild-type pattern (data not shown).

Results from pb/β-galactosidase fusion transgene lines demonstrate that sequences necessary for the pb pattern of expression are located downstream of the pb translation start site

In order to determine the location of the cis-acting regulatory elements responsible for the pb expression pattern, pb/lacZ reporter gene constructs were made and introduced into the Drosophila genome. The pbGl construct consists of 7.3 kb of sequence upstream of the pb transcription initiation site, exon-1, intron-1, and two thirds of exon-2 including the ATG translation initiation site (Fig. 6). The exon-2 coding sequences are fused in-frame with the lacZ coding sequences to make a pb/β-galactosidase fusion protein. The pbG2 construct also contains 7.3 kb of sequence upstream of the pb transcription initiation site (Fig. 6). In this case, however, RNA leader sequences in the first exon of pb are fused to the alcohol dehydrogenase (ADH) RNA leader. The ADH sequences also include an ATG translation start site and ADH coding sequences which are fused in-frame with lacZ sequences. Both the pbGl and the pbG2 constructs include precisely the same pb 5′ upstream sequences as the minigene. Six independent pbGl insertion lines and two pbG2 lines were obtained via P-element mediated germline transformation. Embryos and third instar larvae carrying the different reporter gene insertions were immunostained using polyclonal antibodies directed against β-galactosidase.

The results from the six pbGl transformant lines demonstrate that sequences upstream of the pb translation start site are not sufficient for the pb pattern of expression in embryos and third instar larvae. In both embryos and larvae, all six of the pbGl transgene lines and both pbG2 lines show different patterns of expression, or no expression at all, and none of the patterns are similar to that of pb (data not shown). Thus, rather than displaying major components of or the full pb pattern of expression, both the pbGl and pbG2 reporter genes seem to be acting as 'enhancer traps', i.e. the pb promoter sequences in the constructs are presumably being driven or repressed by the cis-acting regulatory elements of other genes located nearby the sites of insertion in the genome.

Because the minigene rescues the homeotic phenotype and also expresses the normal pb protein pattern in the embryonic and larval head, sequences necessary for this pattern must be present in the minigene but absent from these two reporter constructs (see Fig. 6). Thus, important pb regulatory elements that are necessary for this pattern must be located downstream of the pb pattern.

Fig. 6. The internal negative control is photoreceptor cells (prc) in a metameric pattern. (D) Similarly staged homozygous P[w+, pb+]; pb+ embryo showing complete absence of lacZ expression.

Fig. 5. Dissected portions of third instar larvae stained with polyclonal rabbit antiserum specific for the pb protein and a secondary antibody conjugated to horseradish peroxidase (HRP). Top is anterior and bottom is posterior. 5A–C. Imaginal disks from third instar larvae stained in the same reaction tube. 20x magnification. 5A. The internal positive control (w/w) is wild-type for pb expression in the labial disks (lab). The small arrowhead near the bottom right points to tracheal epithelial cell staining. (B) Internal negative control – pb null (pb+/pb−). The labial disks are transformed into first thoracic leg disks (T1) and there is no pb staining, only nonspecific background staining. This nonspecific staining occurs in all imaginal disks and is independent of the specific antibody used (it is not specific to just pb antisera). The intensity of the nonspecific background staining is dependent on the elapsed time of the HRP color reaction. This particular batch (A–C) was exposed to the HRP reaction for a longer period of time, hence the visible background staining. Batches of pb antisera-treated larvae that were exposed to the HRP reaction for shorter amounts of time showed proportionally less background. (C) Labial disks from a rescued larva (homozygous P[w+, pb+]; pb+; Ex/Ex) stained strongly for pb. The disks are not flattened on the slide as much as with the positive control (A). The labial disk on the right was torn during mounting. The small arrowhead near the bottom right shows staining in tracheal epithelial cells somewhat out of the plane of focus. The boxed-in area is shown magnified 4x in H to highlight the punctate nuclear staining. (D–F) Central nervous system from third instar larvae stained in the same reaction tube. The brain hemispheres (br) are above and the ventral ganglion (vg) is below. (D) The internal positive control (w/w) is wild-type for pb expression. The lines point to pb expression in cells of the supraesophageal (spg) and subesophageal ganglia (sbg). Cells staining in the sbg of the left brain lobe is out of the focal plane. (E) Internal negative control – pb null (pb+/pb−). (F) Homozygous P[w+, pb+]; pb+. The upper arrowhead points to weak spg staining and the lower right arrowhead points to weak sbg staining. The arrowhead on the lower left points to the area where seven anterior vg cells show pb accumulation. (G) 40x magnification of wild-type pb protein accumulation in the CNS.

sequences included in the pbGl construct but not in either the large intronic region absent from the minigene or downstream of the first pb poly(A) site.
Fig. 8. (A) pbGl/intron reporter gene line stained with antibodies directed against β-galactosidase and secondary antibodies conjugated to horseradish peroxidase (HRP). 20x magnification. Embryo is oriented anterior to the left and posterior to the right. β-galactosidase accumulates in the labial lobes (lab), the maxillary lobes (max) and a subset of ectodermal cells of the mandibular lobes (see arrowhead). Negative controls (w/w), i.e. animals without a pbGl/intron reporter gene, had the same nonspecific background staining levels as the areas outside lab and max (data not shown). (B,C) Dissected third instar larvae showing labial disks (lab) stained with antibodies directed against β-galactosidase and secondary antibodies conjugated to HRP. Anterior is above and posterior is below. 20x magnification. (B) Internal negative control (w/w). (C) pbGl/intron reporter gene line stained in the same reaction tube as B. The HRP/chromogen reaction was not allowed to continue as long as with the embryo shown in A. Thus, nonspecific background staining levels are lower with the negative control and the staining intensity of C is concomitantly lower.
DNA sequence comparison between D. melanogaster and D. pseudoobscura indicates the location of putative pb cis-acting regulatory elements

In order to help pinpoint the location of the important cis-acting regulatory elements of pb, DNA sequences of two divergent species were compared. Because D. pseudoobscura is approx. 46 million years diverged from D. melanogaster (Beverley and Wilson, 1984), sequence conservation between these two species should reflect functional relevance.

Lambda EMBL4 genomic clones with high sequence homology to specific pb probes from D. melanogaster were obtained in a chromosome walk through the Antennapedia Complex in D. pseudoobscura (Ran-dazzo, Seeger, Huss and Kaufman, in prep.). Regions with high sequence similarity in D. pseudoobscura were localized further by Southern blot analysis, subcloned, and then sequenced. The D. pseudoobscura DNA sequence was then compared with the existing pb sequence from D. melanogaster. Outside of mRNA-coding sequences, blocks of strongly conserved sequences were observed both in the second pb intron, slightly downstream of the second exon, and about 300 bp upstream of the pb mRNA 5' terminus in D. melanogaster (Fig. 6).

Fig. 7A shows DNA sequence in the exon-2/intron-2 boundary region of pb in both species. The end of the second exon of pb in D. pseudoobscura was inferred, based on sequence identity to the second pb exon in D. melanogaster, which was defined by S1 nuclease protection analysis and PCR-generated cDNAs (Cribbs, Pultz, Johnson and Kaufman, in prep.). Third base degeneracy which preserves identity at the amino acid level is evident in 4 of the last 12 codons of exon-2. Numerous blocks of sequence identity reside downstream of the second exon of both species. Five elements ranging in size from 18–28 bp and three elements between 38 and 59 bp in length exhibit 100% sequence conservation between the two species. Shorter blocks of 100% sequence conservation are...
Fig. 7. (A) DNA sequence comparison of the exon-2/intron-2 regions of proboscipedia in *D. melanogaster* and *D. pseudoobscura*. The *D. melanogaster* sequence is above and the *D. pseudoobscura* sequence is aligned directly below. The exonic sequences are boxed-in as are all of the intronic sequences with complete sequence identity. The first box at the top of the figure consists of the last 12 codons of exon-2. The vertical dotted lines separate codons. The asterisks below designate third base degeneracy in the codons. Amino acids are designated by their single letter abbreviations above the codon sequences. The rest of the figure consists of intronic sequences. For alignment purposes, the dots represent sequences absent in the respective species. Directly repeated sequences are overlined by arrows. The stippled area highlights a long but interrupted direct repeat named r1, while r2 is a directly repeated sequence that is contained within r1 and is also found in a third location between the two r1 repeats. An 11 bp directly repeated sequence is referred to as r3, while two additional repeats are designated as r4 and r5. (B) Upstream DNA sequence comparison between *D. melanogaster* and *D. pseudoobscura*. Boxed-in areas show complete DNA sequence identity. This conserved element lies approximately 300 bp upstream of the *pb* 5' mRNA terminus in *D. melanogaster*. 
clustered around these elements. Between the two species, sequences separating these conserved elements differ in both nucleotide sequence and in length. We also have cloned the putative pb homologue from *D. virilis*, a species estimated to be ~62 million years diverged from *D. melanogaster* (Beverley and Wilson, 1984). We have sequenced a portion of the inferred intron-2/exon-2 region in *D. virilis* and preliminary results show a substantial amount of sequence identity with the conserved intron-2 region in *D. melanogaster* and *D. pseudoobscura* including 100% conservation of the 'r1' motif (see below).

A number of directly repeated sequences reside within these conserved stretches. Included within two of the conserved blocks is a directly repeated 32 bp motif, ANCTGTTATNNGGTCTAGTTGNNAAATTT, referred to as 'r1' (stippled and indicated by arrows in Fig. 7A). A seven nucleotide motif consisting of the sequence TACATTG, and referred to as 'r2', is located within the r1 repeat and also is found at a third location between the two r1 repeats (indicated by arrows in Fig. 7A). Additionally, an r2 sequence is found in the opposite orientation approximately 150 bp downstream of the exon-1 splice donor site in *D. melanogaster* (data not shown). The corresponding region, however, has not yet been sequenced in *D. pseudoobscura*. An 11 bp directly repeated sequence, TTTCCTGTCC, is found in two locations and is designated as 'r3' (indicated by arrows in Fig. 7A). Two sets of conserved direct repeats, 'r4' and 'r5', overlap one another and are located downstream of the second r3 element (indicated by arrows in Fig. 7A). The sequences of r4 and r5 are similar; the r4 sequence is GTTGC while the r5 sequence is TTGCATT.

These conserved intron-2 sequences are contained within the minigene but are excluded from the pbG1 and pbG2 reporter gene constructs (Fig. 6). Because the minigene expression pattern resembles normal pb expression, whereas the β-galactosidase reporter genes fail to reveal a pb-like expression pattern in embryos and third instar larvae, these blocks of conserved sequences are good candidates for cis-acting regulatory sequences necessary for the correct pattern of pb expression.

Located 305 bp upstream of the pb transcription start site and 696 bp downstream of the z2 poly(A) site in *D. melanogaster* is a highly AT rich 60 bp stretch sharing 55 bp of sequence identity with *D. pseudoobscura* (Fig. 7B). The distances from the second pb exon to the conserved upstream sequence are similar in both species, i.e. 2.6 kb in *D. melanogaster* versus 1.85 kb in *D. pseudoobscura*, though we do not yet know the precise size or location of exon-1 in *D. pseudoobscura*. The DNA sequences surrounding the conserved block diverge, presumably due to lack of functional constraints. This upstream conserved element is included in the minigene and in both the pbG1 and pbG2 reporter gene constructs (Fig. 6). The protein expression patterns of the minigene and of both these reporter genes indicate that while these sequences may be necessary they are not sufficient to direct any aspect of the pb pattern of expression. It is highly unlikely that the upstream conserved element is involved in z2 rather than pb regulation; even though the 11.5 kb distance between zen and pb is approximately conserved between *D. melanogaster* and *D. pseudoobscura*, preliminary evidence indicates that z2 coding sequences are not present in *D. pseudoobscura* between zen and pb (Randazzo, Seeger, Huss, and Kaufman, in prep.).

The region containing the conserved intronic elements is necessary for pb expression in the embryonic and larval head

We next set out to determine whether the conserved element region in the second intron was necessary for pb expression in the embryonic and larval head. A 1.6 kb EcoRI restriction fragment containing the conserved intronic elements was inserted into the pbG1 lacZ reporter gene construct 1 kb upstream of the pb transcription start site (Fig. 6). The fragment was inserted into z2 in the same relative orientation as its original position in pb. Two independent lines were obtained by P-element transformation in *Drosophila*. β-galactosidase protein accumulation was monitored in both embryos and third instar larvae. Both reporter gene lines showed a pattern of β-galactosidase accumulation in the embryonic head identical to that of pb protein accumulation observed in rescued embryos (compare Figs 8A with 4B). The labial lobes, the maxillary lobes, and 5 or 6 ectodermal cells of the mandibular lobes show β-galactosidase accumulation. In older embryos of one of the independent pbG1/ intron lines, ectopic expression is observed in cells of the central and peripheral nervous systems. As with wild-type and rescued third instar larvae, both of the independent pbG1/intron lines exhibit significant β-galactosidase expression in the labial disks, as compared with the internal negative control (compare Figs 8C with 8B). No pb-like pattern of expression is observed in the CNS. Additionally, larvae from both lines exhibit ectopic expression in the distal most portion of the antennal and T1, T2, and T3 leg disks. Taken together, the results from the embryonic and larval analyses of these transgenic lines demonstrate that the restriction fragment containing the intron-2 conserved elements is necessary for pb expression in the head. At present we do not know why these reporter constructs also show ectopic expression of β-galactosidase, a feature not observed in the minigene lines. However, we suspect that the nervous system expression observed in the one line is best explained by position effects and that the ectopic disc expression in both lines reflects the absence of a negative regulatory element in the reporter constructs. These possibilities are currently under investigation.

**Discussion**

**Rescue of a homeotic gene**

Rescue of the *proboscipedia* gene with two independent insertion lines directly demonstrates that the DNA
from 7.3 kb upstream to just downstream of the proposed pb transcription unit is responsible for pb homeotic function. Two transcription units reside within this interval, z2 and pb. The pattern of expression for z2 closely resembles that of zen (Rushlow et al. 1987), but genetic analysis has failed to ascribe any function to z2 (Pultz et al. 1988). Taken together, previous genetic and molecular evidence (Pultz et al. 1988) and the rescue results presented here demonstrate that pb function can indeed be attributed to the transcription unit previously identified as pb.

Immuno-staining of rescued embryos and third instar larvae with antibodies directed against pb protein shows that the transgenic pb pattern of expression largely mimics the normal endogenous pattern of ectodermal expression in the head region of both embryos and third instar larvae. Additionally, ectopic expression is not observed. These results seem to exclude the possibility that the observed rescue of the homeotic phenotype is mediated by unregulated, global expression of the transgenic pb protein.

To date, rescue of other homeotic genes of the ANT-C and the BX-C with minigenes that utilize a nonheterologous promoter and regulatory elements has not been reported. However, it should be noted that labial, another homeotic gene in the ANT-C, has been partially rescued using a minigene, i.e. rescue of the embryonic but not adult phenotype (Chouinard and Kaufman, in prep.). Since by homeotic standards pb and labial have relatively small cis-acting regulatory regions (Pultz et al. 1988; Diederich et al. 1989), it would appear that difficulties in performing rescue of homeotic genes stem in part (if not primarily) from the problems associated with the great size and complexity of these genes and not necessarily from residential requirements in a specific chromosomal domain.

What is the functional significance of the ANT-C complex structure?

The homeobox-containing genes of the ANT-C and the BX-C presumably arose through duplication and divergence events prior to the ancestral split between the deuterostomes and protostomes during the late Precambrian period. But why have these genes remained both clustered and in the same linear order for over 500 million years of evolution? The ANT-C and the BX-C in Drosophila melanogaster presumably arose from an event that split a larger single ancestral complex. Additionally, genetic manipulations using deficiencies and translocations have demonstrated that Ubx, abd-A and Abd-B can be separated without compromising function under laboratory conditions (Struhl, 1984; Tiong et al. 1987). Moreover, Antp and Ubx reporter gene constructs inserted into chromosomal locations outside of the homeotic gene complexes have been shown to closely mimic the patterns of expression of their respective genes (Boulet and Scott, 1988; Irvine et al. 1991). The two pb minigene transgenic lines that completely rescue the homeotic phenotype are second chromosome insertions, whereas the ANT-C resides on the third chromosome. Thus, our results demonstrate, at least in terms of homeotic regulation and function (we define homeotic function in terms of the visible cuticular transformations), that the pb gene works out of its normal chromosomal context. Clearly, it is possible to separate at least some of the homeotic genes from the presumed configuration of the ancestral gene cluster and from the present ANT-C and BX-C structure in D. melanogaster and still have a viable fly. Whether or not these flies can survive the selective pressures imposed in a nonlaboratory environment is unclear at this point.

At least two possibilities exist for the evolutionary selection of these homeotic gene complexes. One possible explanation is that genes within the cluster may possess shared or interspersed cis-acting regulatory elements. On the other hand, inversions in the ANT-C that break within one gene fail to affect the other genes in the complex and therefore suggest that the homeotic genes of the ANT-C may not share regulatory elements (Abbott and Kaufman, 1986; Pultz et al. 1988; Diederich et al. 1989).

A second possible explanation for the selection throughout evolution against the dispersion of these homeotic genes is the detrimental consequences of position-effects on these genes, since inversion events within the complex could influence the transcription of the nearby homeotic genes by bringing in new sequences that (1) contain enhancer-like elements and/or (2) change the local chromatin structure. While ectopic expression of some genes in Drosophila fails to discernibly affect the organism (Bonner et al. 1984; Basler and Hafen, 1989), ectopic expression of the homeotic genes Antp, Ubx and pb have been found to cause highly deleterious effects on the organism (Gibson and Gehring, 1988; Mann and Hogness, 1990; Gonzalez-Reyes et al. 1990; Cribbs, Randazzo and Kaufman, in prep.). Thus, the homeotic genes of the ANT-C may reside within the complex as 'prisoners' whose escape, in many cases, may cause detrimental effects on the organism as a whole. The minigene line that partially rescues may be a case in point, for its viability is severely compromised. In addition, G0 minigene transformants may have failed to survive (given the low efficiency of recovery compared with the reporter gene constructs), although this type of negative evidence is difficult to quantitate and remains distressingly anecdotal in character.

Putative elements missing from the minigene result in the loss of the full pb expression pattern

We were able to show, in animals null for endogenous pb, that levels of transgene-derived pb protein accumulation in the supraesophageal and subesophageal ganglion seem normal in embryos and are significantly reduced in larvae. It is possible that expression is initiated in the embryo but is not maintained throughout larval development. Additionally, transgene-derived pb protein expression, in pb- animals, is initiated but not maintained in the mesodermal cells of the embryo where pb normally accumulates. Therefore,
tissue-specific maintenance elements may be missing from the minigene. In \( pb^- \) animals, detectable ventral nerve chord expression of \( pb \) in embryos and detectable ventral ganglion expression of \( pb \) in third instar larva are absent in animals carrying the minigene. Consistent with all of these results, the \( pb \) reporter gene constructs, which contain only a subset of the total \( pb \) DNA sequences used to construct the minigene, fail to demonstrate the ability to initiate or maintain a \( pb \)-like pattern of expression in the mesoderm or nervous system. Missing cis-acting regulatory elements, the missing microexon, or changes in genome context could explain the defects in nonhomeotic expression. The presence of the 15 bp alternatively spliced microexon, an exon missing from the minigene, may be necessary for \( pb \) protein stability in the CNS and mesoderm or for the accumulation of the \( pb \) transcript in those cells. Analysis of the Ubx gene products indicates that alternate splice products are differentially localized in the epidermis and mesoderm versus the CNS (Kornfeld et al. 1989). At this point we cannot formally rule out the possibility that these results are due to the minigene being outside of the context of the ANT-C. However, because the minigene is significantly truncated as compared with the endogenous \( pb \) gene and because the remaining aspects of transgene derived \( pb \) expression are normal, we believe the defects in minigene expression are due to missing cis-acting control elements or the missing microexon and not to changes in chromosome context. These putative missing cis-acting control elements would be located either in the major intron or downstream of the first \( pb \) poly(A) site but not upstream of the proximal \( In(3R)MAP17 \) breakpoint, because embryos (Pultz, 1988) and larvae homozygous for this inversion seem to show neither reduced levels nor missing aspects of the \( pb \) protein accumulation pattern. Furthermore, because a clear role for \( pb \) expression in the mesoderm and nervous system does not exist, the possibility remains that \( pb \) expression in these tissues is controlled fortuitously by long range elements.

A correlation exists between defective neural expression and changed behavior. Rescued homozygous \( P[w^+, pb]1; pb^2 \) and \( P[w^+, pb]2; pb^3 \) adults do not readily attempt escape on removal of the plug from the culture vessel but rather tend to stay on or near the bottle. These animals also do not fly properly. Homozygous \( P[w^+, pb]1 \) and wild-type control animals, however, promptly escape when given an opportunity and commonly migrate to the top of the bottle. Unfortunately, drawing comparisons with \( pb \)-null adult controls seems inappropriate because the mouth-to-leg transformation prevents the animal from eating and thus affects behavior via malnutrition. We are in the process of testing the possibility of a cause and effect relationship between the neural expression and the behavioral phenotype. Additionally the homozygous \( P[w^+, pb]1; pb^2 \) and \( P[w^+, pb]2; pb^3 \) stocks fail to propagate well, the latter much more so than the former. The significant differences between these two lines, however, indicates this is a position dependent phenomenon and not a function of the minigene per se – \( P[w^+, pb]2 \) is a semilethal insertion.

**Regulation of \( pb \) expression**

Through the use of a \( pb \) minigene, \( pb/lacZ \) reporter gene fusions, and DNA sequence comparison between two divergent species we have been able to locate sequences necessary for much of \( pb \)'s complex pattern of expression. This includes \( pb \) protein expression in the labial, maxillary and mandibular lobes, subesophageal ganglia, and supraesophageal ganglia of the embryo and in the labial disks, subesophageal ganglia, and supraesophageal ganglia of the larvae. Regulatory elements necessary for this expression lie just downstream of the second \( pb \) exon. This region contains a number of long blocks of DNA sequence that are conserved between \( D. melanogaster \) and \( D. pseudoobscura \), species approx. 46 million years diverged. Present within these conserved sequences are a number of directly repeated elements. Some eucaryotic enhancers include directly repeated sequences. The SV40 virus, for example, contains two directly repeated 72 bp enhancer elements (Moreau et al. 1981). The intron-2 direct repeats therefore may represent enhancer elements responsible for \( pb \) localization in the embryonic and larval head region. There also is a possibility that rather than acting alone, these conserved elements require the coordinate action of other sequences to bring about proper \( pb \) distribution in the head. One candidate for such a sequence is the conserved element located upstream of the \( pb \) transcription start site. This sequence is included in the minigene and also in the \( pb/lacZ \) reporter gene fusion constructs. Both the upstream conserved element and many of the intron-2 conserved blocks of sequence, especially the direct repeats, are AT rich. Sequences that bind homeoproteins have been shown to be AT rich (for review see Biggin and Tjian, 1989). Thus, these conserved repeats may interact with homeobox-containing trans-regulators.

**Possible significance of the putative internal regulatory elements**

It has been hypothesized that the large introns of the homeotic genes are necessary for the proper timing of translation initiation, i.e. longer primary transcripts require more time for their transcription and thus a greater lag time results between transcription initiation and translation (Kornfeld et al. 1989). Absence of the large \( pb \) intron, however, does not affect \( pb \)'s homeotic function. We hypothesize that the presence of large introns may function as a means for distancin internal enhancer elements from the promoters of other nearby homeotic genes, thus preventing cis interactions between genes. A number of smaller Drosophila genes possess regulatory elements within their introns (Basler et al. 1989b; Gasch et al. 1989). Most of the homeotic genes of the ANT-C and the BX-C possess uncharacteristically large transcription units as compared with other Drosophila genes. For example, \( pb \), Scr, Ubx, and Antp contain transcription units of 33 kb, 65 kb, 77 kb, and...
100 kb respectively (see Kaufman et al. 1989 and Kornfeld et al. 1989). The Antp gene has been shown to contain regulatory elements within its introns (Boulet and Scott, 1988). Unpublished data on labial indicate that regulatory elements reside within its largest intron (Chouinard and Kaufman, in prep.). Here we have presented evidence for cis-acting regulatory elements contained within the large pb intron. Long range cis interactions between genes within the ANT-C occur when a deletion mutation juxtaposes fushi tarazu (ftz) promoter sequences 20 kb upstream of the z2 promoter. In this mutant, the ftz sequences drive the expression of z2 in a ftz-like striped pattern (Rushlow and Levine, 1988). This hypothesis is consistent with the one forwarded above as to why these genes have remained clustered and in the same order over hundreds of millions of years of evolution.

We thank Rudi Turner for the scanning electron microscopy work, Dwayne Johnson for sequencing, Marie Mazzulla for the preparation of pb antibodies, Angela Pattacucci for her expert advice on larval staining, John Tamkun, Carl Thummel and Drs. Izmro for the vectors, Mark Seeger and Kathy Huss for cloning in D. pseudoobscura, and Jessica Kissinger for cloning in D. virilis. We are grateful to Charles Langley for the D. pseudoobscura and Doug Cavener for the D. virilis genomic libraries, Scott Chouinard and Elizabeth Raff for critical reading and Chris Sayre for proof reading of the manuscript. We also would like to thank Dee Verostko for her administrative assistance. This work was supported by a National Institutes of Health Predoctoral Fellowship (GM7577) to F.M.R. and an NIH grant (GM24299) to T.C.K. T.C.K. is an investigator of the Howard Hughes Medical Institute.

References


