Cell, Vol. 47, 1033–1040, December 26, 1986, Copyright © 1986 by Cell Press

Conservation of a Large Protein Domain in the Segmentation Gene *paired* and in Functionally Related Genes of Drosophila

Daniel Bopp, Maya Burri, Stefan Baumgartner, Gabriella Frigerio, and Markus Noll
Department of Cell Biology
Biocenter of the University
Klingelbergstrasse 70
CH-4056 Basel, Switzerland

Summary

Extending our search for homologous domains of the Drosophila *paired* gene, two closely linked genes at the *gooseberry* locus have been isolated. Both genes are expressed with a single segment periodicity but with different spatial and temporal expression patterns. While the transcripts of one gene appear earlier and are equally distributed between ectoderm and mesoderm, those of the second gene accumulate preferentially in neuroblasts. The similarity of these expression patterns to the 14-band pattern of the *paired* gene suggests a functional relationship. Such a functional link may be reflected in the two structurally homologous domains shared with the *paired* gene: a new type of homeo box extended by 18 amino acids at the 5′ end, and a new domain, the paired box, consisting of a sequence of 128–135 amino acids. Thus, together with the PRD repeat, the *paired* gene contains at least three different domains, each defining a gene set thought to be important for development.

Introduction

The generation of the metameric pattern in a Drosophila *melanogaster* embryo has been demonstrated to be governed by genes that fall into three categories, the gap genes, the pair-rule genes, and the segment polarity genes (Nüsslein-Volhard and Wieschaus, 1980; Nüsslein-Volhard et al., 1982). It is thought that the embryo is organized first by the activity of the gap genes into relatively large spatial domains comprising several adjacent segment primordia, which are subsequently subdivided by the pair-rule genes into repeating units of a double-segment periodicity, and finally into single segments by the segment polarity genes (Nüsslein-Volhard et al., 1982). Five pair-rule genes (Scott et al., 1983; Holmgren, 1984; Kuroiwa et al., 1984; Fjose et al., 1985; Ish-Horowicz et al., 1985; Poole et al., 1985; Kilchherr et al., 1986; Harding et al., 1986) and one gap gene (Preiss et al., 1985) have been isolated and their interaction with other segmentation genes has been examined by comparing their spatial expression pattern in wild type with that in various segmentation mutant embryos (Carroll and Scott, 1986; Howard and Ingham, 1986; Harding et al., 1986). These studies indicate that these genes interact in a regulatory network (Carroll and Scott, 1986; Harding et al., 1986) rather than in a strict hierarchical fashion.

We have recently isolated the pair-rule gene *paired* (*prd*) which exhibits a dynamic pattern of expression, switching from a 7 band pattern of double-segment periodicity to a 14 band pattern with a single-segment repeat during formation of cellular blastoderm (Kilchherr et al., 1986). Thus *prd* is first expressed in a pattern similar to that of other pair-rule genes like *fushi tarazu* (*ftz*) (Hafen et al., 1984), *hairy* (*h*) (Ingham et al., 1985), and *even-skipped* (*eve*) (Harding et al., 1986) and later expressed in a manner expected for segment-polarity genes. Therefore *prd* might be involved in the control of the transition from the spatial organization with a double-segment periodicity to one with a single.

Based on evolutionary considerations, we have proposed that one approach to investigate the link between genes integrated in a functional network might consist of finding not only one but all different domains of a gene and their structural homologues. In addition to the generally accepted principle that the structural homology reflects a functional relationship, we assume that related domains are found preferentially in the same integrated gene system, as, for example, in the network of genes regulating early development. As a consequence, only a relatively small number of different domains might be used to make up the genes of such an integrated system. Following this rationale, we have searched the *prd* gene for domains shared with other genes and found the PRD repeat and two new types of homeo boxes (Frigerio et al., 1986). While many genes containing a homeo domain (McGinnis et al., 1984a, 1984b; Scott and Weiner, 1984; Kuroiwa et al., 1985; Laughon et al., 1985; Regulski et al., 1985) are known from genetic studies to be functionally related (Lewis, 1978; Lewis, 1980a, 1980b; Kaufman et al., 1980; Denell et al., 1981; Wakimoto and Kaufman, 1981; Struhl, 1982; Wakimoto et al., 1984), only few genes containing the PRD repeat (Frigerio et al., 1986) are presently known to play a role in early developmental processes, namely *bicoid* (*bcd*) (Frohnhöfer and Nüsslein-Volhard, 1986) and *prd* (Nüsslein-Volhard and Wieschaus, 1980). For the validity of our approach, it is crucial to demonstrate a functional link between genes sharing structurally homologous domains different from the homeo box. Continuing this line of investigation, we show here that the *prd* gene contains a third domain about twice the size of the homeo domain, the paired box, which is homologous to two genes encoded at the *gooseberry* (*gsb*) locus. In addition, these two genes share with *prd* the homeo domain that is extended by 18 amino acids at its 5′ end. If these two genes indeed represent the *gsb* locus, their relationship to the *prd* gene is likely to be of functional significance since *gsb* belongs to the class of segment polarity genes.

Results

Isolation of Genomic Clones Homologous to the BS Probe of *paired*

In our attempt to find structural homologues to different domains of the *prd* gene, we have been screening the ge-

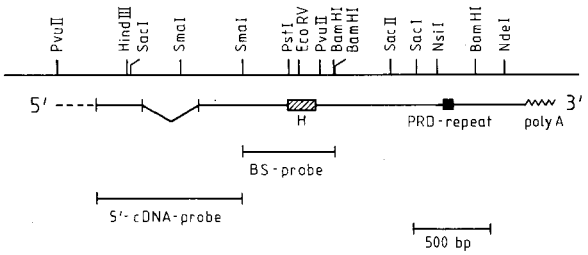

Figure 1. Map of the *paired* Gene and cDNA Probes Used to Screen for Homologous Domains of *paired*

At the top, a restriction map of the sequenced *paired* gene (Frigerio et al., 1986) is shown. Below, the direction and extent of its transcript and the location of the intron are indicated. Two cDNA probes used to screen for homologous domains of *prd* are shown: the 0.6 kb BS probe between SmaI and BamHI containing a homeo box H and the abutting 0.5 kb 5′ cDNA probe comprising the 5′ end of the cDNA c7340.4 (Frigerio et al., 1986). Not shown is the 0.6 kb cDNA probe covering the carboxy-terminal domain and containing the PRD repeat with which genes of the PRD gene set have been isolated (Frigerio et al., 1986).

nome for sequences homologous to a 0.6 kb cDNA probe coding for the carboxy-terminal of the *prd* protein. A set of 11 genes was isolated that are expressed during early development and have the PRD repeat in common (Frigerio et al., 1986). When comparing the sequence of one of these genes, PRD gene 4, which is likely to be *bcd* (Frohnhöfer and Nüsslein-Volhard, 1986), with that of the *prd* gene (Frigerio et al., 1986), another homologous domain was found which turned out to be a homeo domain. Both the *prd* and the *bcd* homeo domain deviate considerably from each other and from all previously known homeo boxes such that they might be considered to represent two new types of homeo domains (Frigerio et al., 1986). In our screen for additional domains of *prd* that are shared with other genes, we used two cDNA probes, one located in the middle of the *prd* gene (BS probe in Figure 1) and the other continuing upstream and including the 5′ end (5′ cDNA probe in Figure 1). After hybridization at low stringency to whole genome Southern blots, both probes revealed weak bands in addition to the strong signal originating from the *prd* gene (not shown). In a first set of experiments, the BS probe was used to screen a genomic library at reduced stringency of hybridization. Two strongly positive clones, BSH4 and BSH9, which were negative after hybridization with sequences flanking the *prd* gene on either side, were isolated and analyzed further with respect to their cytological location on polytene chromosomes.

Surprisingly, both clones hybridized to the same chromosomal bands 60E9-F2 in the telomeric region of the right arm of the second chromosome (Figure 2). As two segmentation genes have been mapped to this region, the segment polarity gene *gooseberry* and the gap gene *Krüppel* (*Kr*) (Nüsslein-Volhard et al., 1984), it is possible that the two clones contain *gsb* or *Kr* or both.

Spatial Distribution of BSH4 and BSH9 Transcripts

To test the supposition that the genomic clones BSH4 and BSH9 coded for *gsb* or *Kr*, subclones derived from these clones and hybridizing with the BS probe of *prd* were used as hybridization probes to study the spatial distribution of complementary transcripts in embryonic tissue sections.

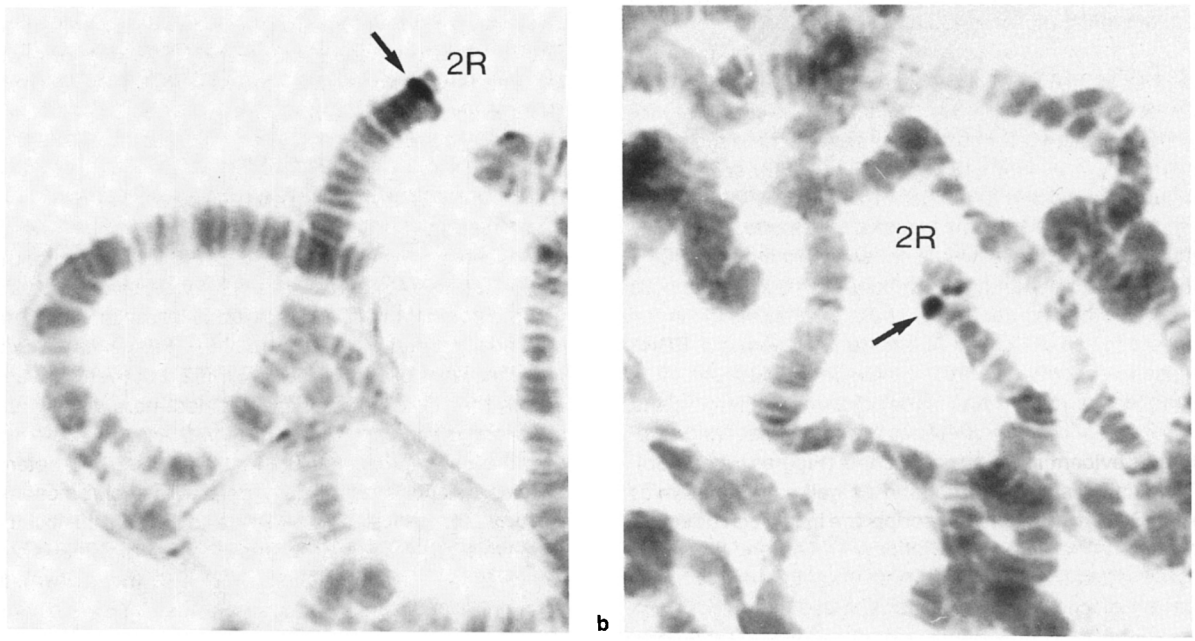

Figure 2. Localization of BSH4 and BSH9 on Polytene Chromosomes Carrying *gooseberry* Deletions

In situ hybridizations of biotinylated BSH4 (a) or BSH9 (b) DNA to salivary gland chromosomes of *Df(2R)IIX62/CyO* (a) or *Df(2R)SB1/CyO* (b) larvae were carried out according to Langer-Safer et al. (1982). Note that in (a) the signal extends only over the *CyO* chromosome (arrow) and is absent in the slightly shorter *Df(2R)IIX62* chromosome while in (b) the signal on the *Df(2R)SB1* chromosome is much weaker than on the *CyO* chromosome (arrow). The telomere of the right arm of the second chromosome (2R) is indicated.

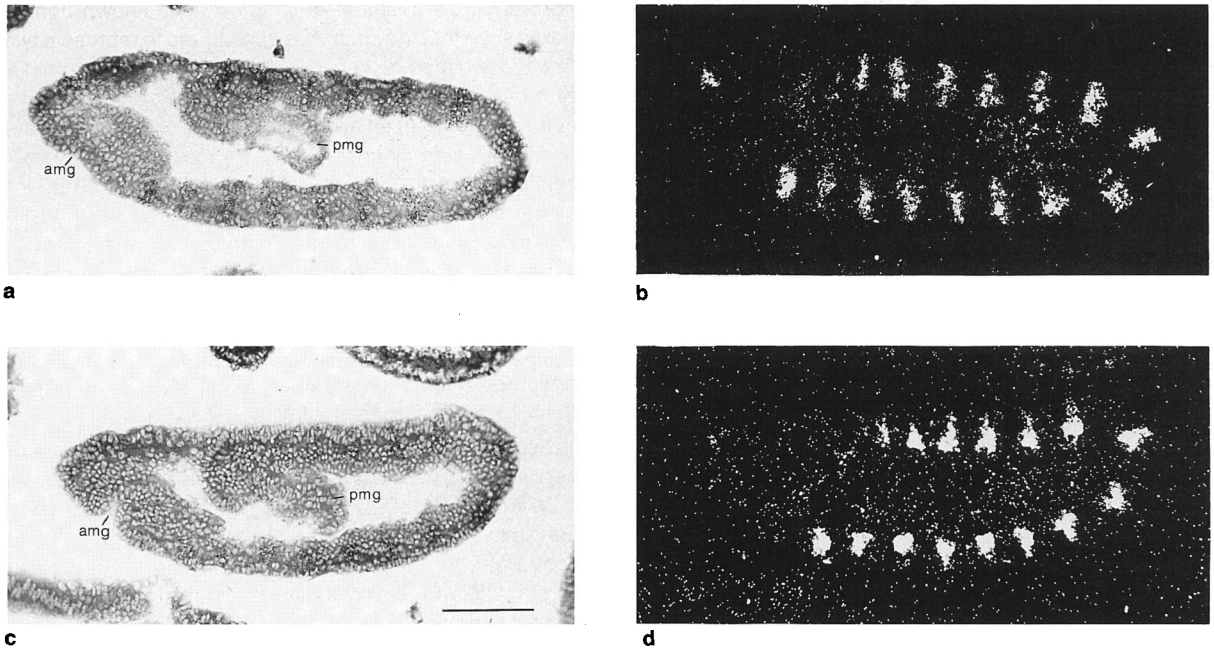

Figure 3. Localization of BSH4 and BSH9 Transcripts in Tissue Sections of Embryos at the Extended Germ Band Stage

The panels show photomicrographs of the same embryos taken under phase-contrast (left) or dark-field illumination (right). Parasagittal sections through embryos at the extended germ band stage (about 5 hr development at 25°C) were hybridized in situ with ^{3}H-labeled genomic probes according to Hafen et al. (1983), with a 1.3 kb BamHI–EcoRI fragment of BSH9 containing the homeo box (the BamHI site is located between the paired box P and the homeo box H in Figure 4) in (a, b), or with the proximal 2.4 kb terminal EcoRI fragment of BSH4 (Figure 4) in (c, d). Autoradiographic exposure was for 28 days. Tissue sections are oriented with their dorsal side up and their anterior end to the left. amg: anterior midgut invagination. pmg: posterior midgut invagination. The horizontal bar in (c) indicates a length of 0.1 mm.

As evident from Figure 3, both probes detect similar patterns of transcripts during germ band extension. Transcripts complementary to the BSH9 subclone appear at the beginning of gastrulation in a pattern of 15 regularly spaced bands of a single-segment repeat (not shown) still present at the fully extended germ band stage (Figures 3a and 3b). In addition, cells on the dorsal side at the anterior end of the embryo, probably corresponding to the labral segment primordium (Jürgens et al., 1986), are labeled (Figures 3a and 3b). Transcripts hybridizing with the BSH4 subclone exhibit the same spatial distribution as BSH9 transcripts but are delayed with respect to those of BSH9 and appear only during germ band elongation (the cells labeled at the anterior dorsal end are barely visible in the section shown in Figures 3c and 3d). Furthermore, BSH4 transcripts appear to accumulate in neuroblasts at a higher concentration than in the ectoderm (Figures 3c and 3d) whereas no such difference between meso- and ectoderm is evident for BSH9 transcripts (Figures 3a and 3b). The detailed spatial distribution as well as the dynamics of the BSH4 and BSH9 transcripts are in fact considerably more complex. As the description would exceed the space of this paper, it will be the subject of a later study. For the present discussion, it suffices to stress that both BSH4 and BSH9 transcripts exhibit a single-segment periodicity. Such a pattern is consistent with that of a segment polarity phenotype in which homologous regions of each segment are deleted and replaced by mirror image duplications of the remaining part of the corresponding segments (Nüsslein-Volhard and Wieschaus, 1980; Nüsslein-Volhard et al., 1982). Therefore, it is likely that BSH4 and/or BSH9 contain sequences coding for *gsb*. On the other hand, the observed patterns of transcripts (Figure 3) as well as the EcoRI restriction maps of BSH4 and BSH9 (Figure 4) differ from those expected for the *Kr* gene (Knipple et al., 1985; Preiss et al., 1985).

BSH4 and BSH9 Are Closely Linked within a *gsb* Deletion

The question of whether BSH4 or BSH9 or both harbor the *gsb* gene was examined by in situ hybridization to salivary gland chromosomes of heterozygous larvae in which one chromosome carried a deletion of the *gsb* gene. Two *gsb* deletions were used. One, *Df(2R)IIX62*, uncovers *gsb* and *zipper*; the other, *Df(2R)SB1*, is located more distally and uncovers *gsb* and *Kr* (Nüsslein-Volhard et al., 1984). In situ hybridization of BSH4 or BSH9 to *Df(2R)IIX62/CyO* heterozygotes detects a signal only on the *CyO* chromosome (Figure 2a). Hence, both BSH4 and BSH9 are located within this *gsb* deficiency. On the other hand, BSH4 is outside of the other *gsb* deletion, *Df(2R)SB1* (not shown). In contrast, most of BSH9 is located within *Df(2R)SB1* as is evident from the weak signal on this chromosome after in situ hybridization of BSH9 to *Df(2R)SB1/CyO* polytene chromosomes (Figure 2b). Since the deletion *Df(2R)SB1* is distal to *Df(2R)IIX62* (Nüsslein-Volhard et al., 1984), BSH9 is located closer to the telomere than BSH4.

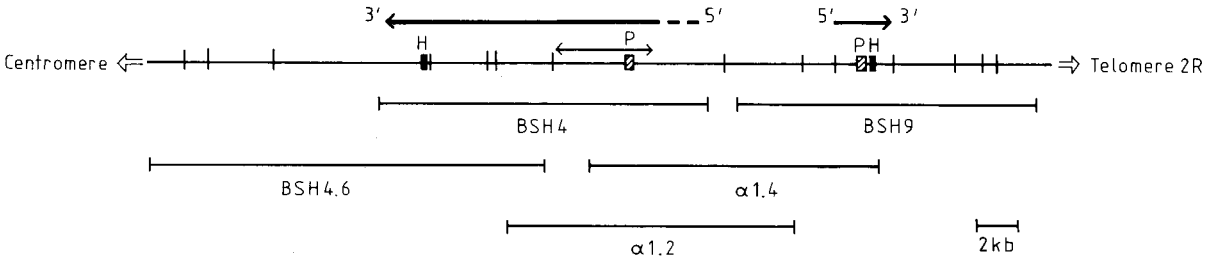

Figure 4. Map of the Chromosomal Region at 60F1 Comprising Two Genes with Structural Homologies to the *paired* Gene

An EcoRI restriction map of the region comprising the two clones BSH4 and BSH9 and its orientation with respect to the telomeric region of the right arm of the second chromosome are shown. BSH4 and BSH9 were isolated from a genomic library (Frigerio et al., 1986) in EMBL4 (Frischauf et al., 1983) by hybridization with the BS probe (Figure 1) at low stringency. The two clones were joined by a "chromosomal walk" (Bender et al., 1983) in which the clones BSH4.6, α1.2, and α1.4 were isolated. The clone BSH9 crosses the proximal boundary of *Df(2R)SB1* (Figure 2b) which has been mapped to the chromosomal band 60F1 (Nüsslein-Volhard et al., 1984). At the top, the orientation and extent of two transcripts corresponding to the two genes carrying the extended *prd* homeo box H and the paired box P are shown. Above the paired box of the BSH4 gene, a double-headed arrow indicates that a small portion of the 3′ end of the paired box is located close to the left arrowhead and is separated by an intron from the main part of the paired box which has been mapped to a 2.7 kb HindIII fragment with its distal end at the right arrowhead. The paired box and the extended *prd* homeo box of BSH4 are further separated by a large intron of about 6 kb.

The in situ hybridization to chromosomes further suggests that BSH4 and BSH9 are closely linked. This was verified in a "chromosomal walk" in which the two clones were shown to be separated by only 1.4 kb (Figure 4). Sequence analysis of isolated cDNA clones corresponding to the BSH4 and BSH9 regions showed that the two regions code for separate transcripts on opposite DNA strands whose 5′ ends are separated by less than 10 kb and possibly as little as 6 kb, as indicated in Figure 4. The spatial distribution of BSH4 and BSH9 transcripts and the chromosomal location and close proximity of their transcriptional units strongly suggest that either BSH9 or BSH9 and BSH4 are responsible for the *gsb* phenotype. The possible involvement of two genes receives some support from the observation that only deficiencies and no point mutations have been isolated for *gsb* (Nüsslein-Volhard et al., 1984). On the other hand, a defect in BSH9 might suffice to produce the cuticular *gsb* phenotype while BSH4 mutants might primarily affect neurectodermal derivatives (Figures 3c and 3d) and hence escape detection in a screen for ectodermal phenotypes.

Extended Homeo Domain of *paired*, BSH4, and BSH9

Because the BS probe turned out to contain a homeo domain (Frigerio et al., 1986), we wanted to know whether the homologous structures detected were domains related more closely to the homeo domain of the *prd* gene. When cDNA clones of BSH4 and BSH9 containing the entire open reading frame were isolated, sequenced, and their sequences compared with the *prd* gene within the region homologous to the BS probe of *prd*, one homologous domain, the *prd* homeo domain, was found (Figure 5). Interestingly, not only does this homeo domain consist of the characteristic sequence of 60 amino acids, but the homology between *prd*, BSH4, and BSH9 extends by 18 amino acids at the 5′ end of the normal homeo box (positions −18 to −1 in Figure 5b). The protein homology in this extended portion is 83% between BSH4 and *prd*, 67% between BSH9 and *prd*, and 61% between BSH4 and BSH9. In addition, the homology within the normal homeo domain of 60 amino acids is rather high, both between *prd* and BSH4 (87%) or BSH9 (85%), as well as between BSH4 and BSH9 (92%). As shown previously (Frigerio et al., 1986), the *prd* homeo domain deviates considerably from all previously known homeo domains and hence might belong to a separate class of homeo boxes together with the homeo domains of BSH4 and BSH9. It is probably significant that wherever *Antp-Ubx*-like homeo boxes differ from the new type of homeo domains in positions that are conserved in the *Antp-Ubx* type, the amino acids in *prd*, BSH4, and BSH9 are again mostly identical (boxed in Figure 5b). This observation further supports the notion that amino acids at these positions are characteristic for a particular class of homeo domains (Frigerio et al., 1986). Of particular interest is an accumulation of amino acids characteristic for the *prd* homeo domain in the second α-helix of the helix-turn-helix model for homeo domains (positions 43, 44, 46, and 50 in Figure 5b) because this region is thought to interact specifically with DNA sequences (Laughon and Scott, 1984).

The Paired Box

As mentioned above, whole genome Southern analysis with a third cDNA probe covering the 5′ region of the *prd* gene (5′ cDNA probe in Figure 1) also revealed weak bands, indicating that domains homologous to this part of the *prd* gene existed elsewhere in the genome. According to our concept, such domains are characteristic features of genes functioning as members of an integrated system. Therefore, in a first screen, we examined whether the homologous domains detected by the 5′ cDNA probe occurred in any gene known to be related to *prd* by a structurally homologous domain, i. e., in the genes of the PRD gene set (Frigerio et al., 1986) or in BSH4 and BSH9. Although no hybridization of the 5′ cDNA probe to any of the PRD genes was detectable, both BSH4 and BSH9 showed a clear signal after hybridization at reduced stringency (not shown).

Comparison of the sequence of the *prd* gene region used as probe (Frigerio et al., 1986) with the sequenced

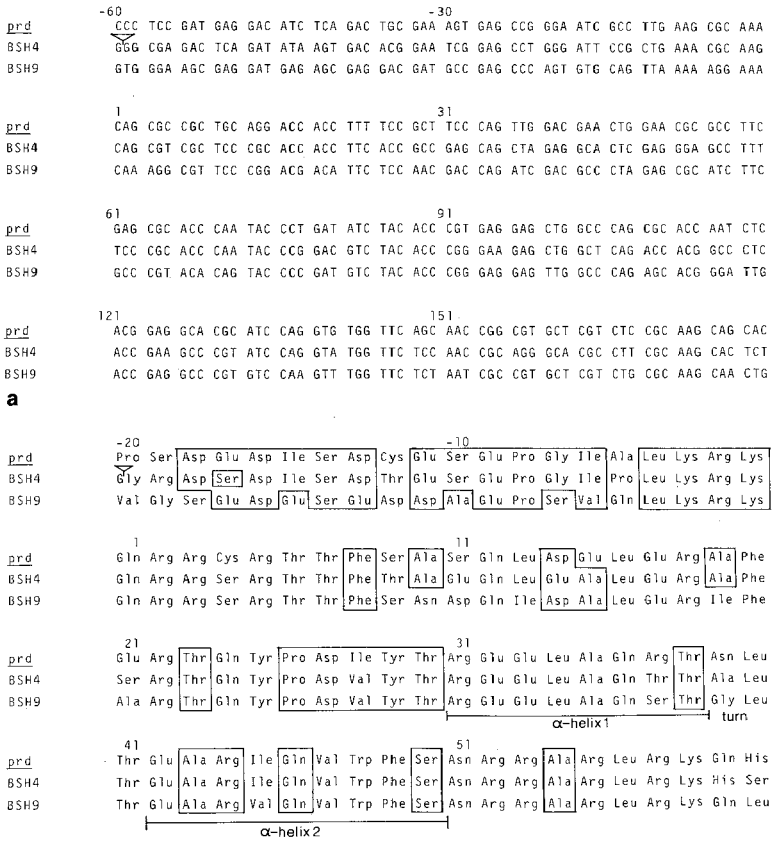

Figure 5. Extended Homeo Box Sequence of *prd*, BSH4, and BSH9

The DNA (a) and corresponding amino acid sequences (b) from the homeo box regions of the *prd* gene (Frigerio et al., 1986), BSH4, and BSH9 are shown. Codons of the common open reading frames, which have been determined from nearly full-length cDNAs comprising the entire open reading frames, are aligned. The homology with previously characterized homeo boxes extends from positions 1 to 180 of the DNA sequence (Frigerio et al., 1986). In this portion, the DNA homology is 72% between BSH4 and *prd* or BSH9 and 71% between *prd* and BSH9. In the extended portion at the 5′ end of the homeo box (positions −1 to −54), the DNA homology is 57% between *prd* and BSH4, 46% between *prd* and BSH9, and 50% between BSH4 and BSH9. In (b), amino acids between positions 1 and 60 that deviate from amino acids conserved in the *Antp-Ubx*-like homeo boxes and which hence are characteristic for the *prd* type of homeo box are boxed. In addition, amino acids conserved with respect to the *prd* gene are boxed in the extended portion of the homeo domain (positions −18 to −1). The region thought to interact with DNA in the helix-turn-helix model (Laughon and Scott, 1984) is indicated in (b). The position of an intron at the 5′ boundary of the extended homeo box of BSH4 is indicated by a triangle.

cDNAs of BSH4 and BSH9 revealed a rather large homologous domain which we call "paired domain" or "paired box" (Figure 6). As indicated in Figure 6b (boxed amino acids), only few nonconservative amino acid changes occur within this domain of 128–135 amino acids. All three proteins are homologous over a length of 128 amino acids while the homology between BSH4 and BSH9 extends by 7 amino acids toward the amino-terminal (Figure 6b). The homology at the protein level is 88% between *prd* and BSH4, 86% between *prd* and BSH9, and 87% between BSH4 and BSH9. It is extremely high in the first 75 amino acids of the paired box (97% between *prd* and BSH4). In contrast, the DNA sequences exhibit considerable divergence (69% homology between *prd* and BSH4, 70% between *prd* and BSH9, and 69% between BSH4 and BSH9). That selection occurs at the protein rather than the DNA level becomes evident if one considers the DNA divergence at those wobble positions where the amino acids have been conserved. Thus one finds 59 nucleotide differences at 94 wobble positions between *prd* and BSH4, which corresponds to complete randomization.

The first half of the paired domain is fairly basic (10 arg or lys and 2 glu of 64 amino acids) and a large portion of the second half is highly charged (9 arg or lys and 8 asp or glu of the 40 amino acids between positions 72 and 111; Figure 6b). The basic part of the domain might interact with nucleic acids or acidic proteins whereas the highly charged portion could be a binding site for itself or other protein domains.

We have sequenced most of the genomic DNA of BSH9 and a considerable portion of BSH4. Comparison with the corresponding cDNA sequences shows that the paired box and the homeo box are separated by an intron in BSH4 yet closely associated in BSH9 (Figure 4). In addition, there is at least another intron close to the end of the paired domain in BSH4 whereas in BSH9 only a single intron at the 5′ end of the paired box has been found (Figure 6). In the *prd* gene, the paired box and extended *prd* homeo box are not separated by an intron (Frigerio et al., 1986). It might be significant that both in BSH4 and BSH9 nearly the same number of amino acids (19 and 20, respectively) separate the paired domain from the extended *prd* homeo domain. In contrast, in the *prd* gene this link consists of 40 amino acids (compare Figure 6b with the *prd* sequence in Frigerio et al., 1986).

Discussion

In our search for domains of the *prd* gene that occur as structurally homologous domains in other genes, we have found, in addition to the previously characterized PRD repeat (Frigerio et al., 1986), two domains that are shared with other genes: the extended *prd* homeo box and the paired box. Both domains are found in two genes, BSH4 and BSH9, that are likely to represent the *gsb* locus. Assuming that *gsb* has been isolated in these clones, two points need to be emphasized in this context. One, the two domains are present in both *prd* and *gsb*. This is in line

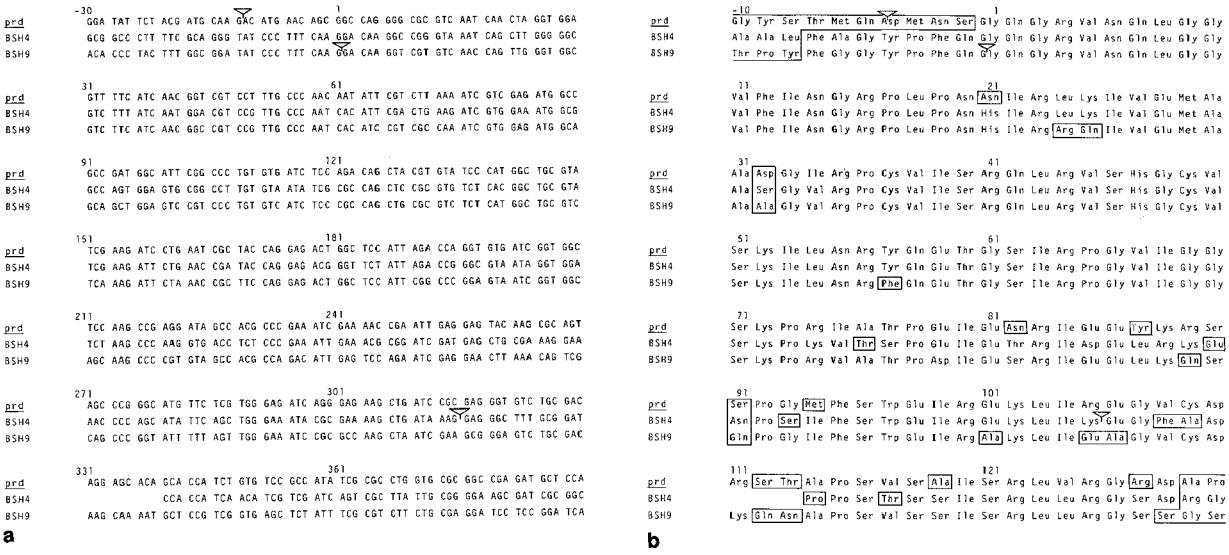

Figure 6. Paired Box Sequence of *prd*, BSH4, and BSH9

The DNA (a) and corresponding amino acid sequences (b) from the paired box regions of the *prd* gene, BSH4, and BSH9 have been aligned as described in the legend to Figure 5. Amino acids that are neither conserved nor represent conservative changes in the three genes are boxed in (b). Optimal homology of the BSH4 protein with that of BSH9 and *prd* is obtained only if one assumes a deletion of three amino acids at position 111–113 in BSH4. The location of introns is indicated by triangles. Additional introns may occur in the 5′ portion of the *paired* box of BSH4 since this region of the genomic DNA has not yet been sequenced.

with our previous observation that the PRD repeat is frequently found to be linked to a homeo domain (Frigerio et al., 1986). Both results are consistent with a relatively small number of different domains that are used preferentially in the same system of integrated genes, such as that determining early developmental events. A small number of such domains mediating specific developmental functions would be sufficient to generate a large number of developmental genes if each gene consists of a combination of several domains (Frigerio et al., 1986). Two, since *gsb*, a segment polarity gene, and *prd*, a pair-rule gene, are both required for proper segmentation, the structurally homologous domains shared between these genes might be responsible for the functional link that is likely to exist between them. The combination of the extended *prd* homeo domain with the paired box in these genes raises the question of whether the postulated functional link is mediated by one or both of the two different domains. That the function of the paired box is not always associated with that of the *prd* homeo box is suggested by preliminary experiments, in which two additional genomic loci have been isolated that hybridize at reduced stringency with the paired box but not with the extended *prd* homeo box. Hence, the specific combination of both domains in *prd* and *gsb* might be important for their functional linkage. It further follows that the paired box genes constitute a new gene set and are not merely a subset of the homeo box gene set.

It will be interesting to find all paired box and (extended) *prd* homeo box genes to determine the size of these gene sets. In addition, it will be necessary to determine the frequency of two different domains occurring in the same gene. This should provide a rough estimate on the number of different domains used in an integrated system of a certain number of genes. Another question is whether there are other segmentation genes containing a paired box which is neither found in the sequence of *ftz* (Laughon and Scott, 1984) nor in the sequence of *en* (Poole et al., 1985).

It is striking that the spectrum of different homologous domains thought to be members of a functional gene network regulating development exhibits such an enormous diversity with respect to length and composition as exemplified by the paired domain and the PRD repeat (Frigerio et al., 1986). The functions of the paired box and of the PRD repeat are unknown. It has been proposed that the carboxy-terminal half of the homeo domain specifically recognizes DNA sequences (Laughon and Scott, 1984). On the basis of in vitro binding studies and model building, O'Farrell has recently postulated that the amino-terminal portion of the homeo domain is a site for protein dimerization (personal communication). It is interesting in this context that the *prd* homeo domain is extended at its amino-terminal end by 18 amino acids. The striking feature common to this extended portion of all three *prd* homeo domains is a cluster of 5–7 acidic amino acids within its first ten amino acids followed, after a gap of five amino acids, by a cluster of six basic amino acids within eight amino acids (Figure 5b). This extension of the domain might be required to specify which proteins may participate in the formation of homo- and heterodimers.

Although we consider it likely that the homeo domain, PRD repeat, and paired domain mediate specific developmental functions, a direct demonstration of what these functions are is still lacking. If we assume a developmental function for the paired box, it might be interesting to

look for its conservation in other organisms. A limitation of the paired box to segmentation genes would further suggest its restriction to segmented animals.

Experimental Procedures

Screening of a Genomic Library with *prd* cDNAs at Low Stringency

A genomic library of Drosophila was prepared in EMBL4 according to Frischauf et al. (1983) and screened with cDNA probes of the *prd* gene at reduced stringency (McGinnis et al., 1984a) as described previously. The cDNA probes were prepared as DNA fragments from subclones of cDNAs isolated from a λgt10-cDNA library of poly(A)$^+$ RNA from 0–4 hr old embryos according to standard procedures (Maniatis et al., 1982). Poly(A)$^+$ RNA was prepared according to the method of Chirgwin et al. (1979).

DNA Sequencing

All DNA sequences were analyzed on both strands, reading each sequence at least twice on independent clones. The DNAs were sequenced in the M13 vector mWB3296 (Barnes et al., 1983; Frigerio et al., 1986) by the dideoxynucleotide sequencing procedure of Sanger et al. (1977).

In situ Hybridization to Polytene Chromosomes

In situ hybridizations to polytene chromosomes were carried out with biotinylated probes according to the method of Langer-Safer et al. (1982) as described (Frei et al., 1985).

In situ Hybridization to Embryonic Tissue Sections

Tissue sections through young embryos were prepared and hybridized in situ with ^{3}H-labeled probes according to Hafen et al. (1983).

Acknowledgments

We are indebted to Christiane Nüsslein-Volhard and Herbert Jäcale for providing the *gsb* mutant stocks and to Marek Mlodzik for a gift of bio-dUTP. We thank Erich Frei and Hans Noll for many stimulating discussions and comments on the manuscript. This work has been supported by the Swiss National Science Foundation grant No. 3.600-0.84 and by the Kanton Basel.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received October 1, 1986.

References

Barnes, W. M., Bevan, M., and Son, P. H. (1983). Kilo-sequencing: creation of an ordered nest of asymmetric deletions across a large target sequence carried on phage M13. Meth. Enzymol. *101*, 98–122.

Bender, W., Spierer, P., and Hogness, D. S. (1983). Chromosomal walking and jumping to isolate DNA from the *Ace* and *rosy* loci and the bithorax complex in *Drosophila melanogaster*. J. Mol. Biol. *168*, 17–33.

Carroll, S. B., and Scott, M. P. (1986). Zygotically active genes that affect the spatial expression of the *fushi tarazu* segmentation gene during early *Drosophila* embryogenesis. Cell *45*, 113–126.

Chirgwin, J. M., Przbyla, A. E., MacDonald, R. J., and Rutter W. J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry *18*, 5294–5299.

Denell, R. E., Hummels, K. R., Wakimoto, B. T., and Kaufman, T. C. (1981). Developmental studies of lethality associated with the Antennapedia gene complex in *Drosophila melanogaster*. Dev. Biol. *81*, 43–50.

Fjose, A., McGinnis, W. J., and Gehring. W. J. (1985). Isolation of a homeo box-containing gene from the *engrailed* region of *Drosophila* and the spatial distribution of its transcripts. Nature *313*, 284–289.

Frei, E., Baumgartner, S., Edström, J.-E., and Noll, M. (1985). Cloning of the *extra sex combs* gene of *Drosophila* and its identification by P-element-mediated gene transfer. EMBO J. *4*, 979–987.

Frigerio, G., Burri, M., Bopp, D., Baumgartner, S., and Noll, M. (1986). Structure of the segmentation gene *paired* and the *Drosophila* PRD gene set as part of a gene network. Cell *47*, 735–746.

Frischauf, A.-M., Lehrach, H., Poustka, A., and Murray, N. (1983). Lambda replacement vector carrying polylinker sequences. J. Mol. Biol. *170*, 827–842.

Frohnhöfer, H. G., and Nüsslein-Volhard, C. (1986). Organization of anterior pattern in the *Drosophila* embryo by the maternal gene *bicoid*. Nature, *324*, 120–125.

Hafen, E., Levine, M., Garber, R. L., and Gehring, W. J. (1983). An improved *in situ* hybridization method for the detection of cellular RNAs in *Drosophila* tissue sections and its application for localizing transcripts of the homeotic Antennapedia gene complex. EMBO J. *2*, 617–623.

Hafen, E., Kuroiwa, A., and Gehring, W. J. (1984). Spatial distribution of transcripts from the segmentation gene *fushi tarazu* during Drosophila embryonic development. Cell *37*, 833–841.

Harding, K., Rushlow, C., Doyle, H. J., Hoey, T., and Levine, M. (1986). Cross-regulatory interactions among pair-rule genes in *Drosophila*. Science *233*, 953–959.

Holmgren, R. (1984). Cloning sequences from the *hairy* gene of *Drosophila*. EMBO J. *3*, 569–573.

Howard, K., and Ingham, P. (1986). Regulatory interactions between the segmentation genes *fushi tarazu, hairy,* and *engrailed* in the Drosophila blastoderm. Cell *44*, 949–957.

Ingham, P. W., Howard, K. R., and Ish-Horowicz, D. (1985). Transcription pattern of the *Drosophila* segmentation gene *hairy*. Nature *318*, 439–445.

Ish-Horowicz, D., Howard, K. R., Pinchin, S. M., and Ingham, P. W. (1985). Molecular and genetic analysis of the *hairy* locus of *Drosophila*. Cold Spring Harbor Symp. Quant. Biol. *50*, 135–144.

Jürgens, G., Lehmann, R., Schardin, M., and Nüsslein-Volhard, C. (1986). Segmental organization of the head in the embryo of *Drosophila melanogaster*. Roux's Arch. Dev. Biol. *195*, 359–377.

Kaufman, T. C., Lewis, R., and Wakimoto, B. (1980). Cytogenetic analysis of chromosome 3 in *Drosophila melanogaster*: the homoeotic gene complex in polytene chromosome interval 84A-B. Genetics *94*, 115–133.

Kilchherr, F., Baumgartner, S., Bopp, D., Frei, E., and Noll, M. (1986). Isolation of the *paired* gene of *Drosophila* and its spatial expression during early embryogenesis. Nature *321*, 493–499.

Knipple, D. C., Seifert, E., Rosenberg, U. B., Preiss, A., and Jäckle, H. (1985). Spatial and temporal patterns of *Krüppel* gene expression in early *Drosophila* embryos. Nature *317*, 40–44.

Kuroiwa, A., Hafen, E., and Gehring, W. J. (1984). Cloning and transcriptional analysis of the segmentation gene *fushi tarazu* of Drosophila. Cell *37*, 825–831.

Kuroiwa, A., Kloter, U., Baumgartner, P., and Gehring, W. J. (1985). Cloning of the homeotic *Sex combs reduced* gene in *Drosophila* and *in situ* localization of its transcripts. EMBO J. *4*, 3757–3764.

Langer-Safer, P. R., Levine, M., and Ward, D. C. (1982). Immunological method for mapping genes on *Drosophila* polytene chromosomes. Proc. Natl. Acad. Sci. USA *79*, 4381–4385.

Laughon, A., and Scott, M. P. (1984). Sequence of a *Drosophila* segmentation gene: protein structure homology with DNA-binding proteins. Nature *310*, 25–31.

Laughon, A., Carroll, S. B., Storfer, F. A., Riley, P. D., and Scott, M. P. (1985). Common properties of proteins encoded by the Antennapedia complex genes of *Drosophila melanogaster*. Cold Spring Harb. Symp. Quant. Biol. *50*, 253–262.

Lewis, E. B. (1978). A gene complex controlling segmentation in *Drosophila*. Nature *276*, 565–570.

Lewis, R. A., Kaufman, T. C., Denell, R. E., and Tallerico, P. (1980a). Genetic analysis of the Antennapedia gene complex (ANT-C) and adjacent chromosomal regions of *Drosophila melanogaster*. I. Polytene chromosome segments 84B-D. Genetics *95*, 367–381.

Lewis, R. A., Wakimoto, B. T., Denell, R. E., and Kaufman, T. C. (1980b).

Genetic analysis of the Antennapedia gene complex (ANT-C) and adjacent chromosomal regions of *Drosophila melanogaster*. II. Polytene chromosome segments 84A-84B1,2. Genetics *95*, 383–397.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory).

McGinnis, W., Levine, M. S., Hafen, E., Kuroiwa, A., and Gehring, W. J. (1984a). A conserved DNA sequence in homeotic genes of the *Drosophila* Antennapedia and bithorax complexes. Nature *308*, 428–433.

McGinnis, W., Garber, R. L., Wirz, J., Kuroiwa, A., and Gehring, W. J. (1984b). A homologous protein-coding sequence in Drosophila homeotic genes and its conservation in other metazoans. Cell *37*, 403–408.

Nüsslein-Volhard, C., and Wieschaus, E. (1980). Mutations affecting segment number and polarity in *Drosophila*. Nature *287*, 795–801.

Nüsslein-Volhard, C., Wieschaus, E., and Jürgens, G. (1982). Segmentation in *Drosophila*, a genetic analysis. Verh. Dtsch. Zool. Ges., 91–104.

Nüsslein-Volhard, C., Wieschaus, E., and Kluding, H. (1984). Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster*. I. Zygotic loci on the second chromosome. Roux's Arch. Dev. Biol. *193*, 267–282.

Poole, S. J., Kauvar, L. M., Drees, B., and Kornberg, T. (1985). The *engrailed* locus of Drosophila: structural analysis of an embryonic transcript. Cell *40*, 37–43.

Preiss, A., Rosenberg, U. B., Kienlin, A., Seifert, E., and Jäckle, H. (1985). Molecular genetics of *Krüppel*, a gene required for segmentation of the *Drosophila* embryo. Nature *313*, 27–32.

Regulski, M., Harding, K., Kostriken, R., Karch, F., Levine, M., and McGinnis, W. (1985). Homeo box genes of the Antennapedia and bithorax complexes of Drosophila. Cell *43*, 71–80.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA *74*, 5463–5467.

Scott, M. P., and Weiner, A. J. (1984). Structural relationships among genes that control development: sequence homology between the *Antennapedia, Ultrabithorax,* and *fushi tarazu* loci of *Drosophila*. Proc. Natl. Acad. Sci. USA *81*, 4115–4119.

Scott, M. P., Weiner, A. J., Hazelrigg, T. I., Polisky, B. A., Pirrotta, V., Scalenghe, F., and Kaufman, T. C. (1983). The molecular organization of the *Antennapedia* locus of Drosophila. Cell *35*, 763–776.

Struhl, G. (1982). Genes controlling segmental specification in the *Drosophila* thorax. Proc. Natl. Acad. Sci. USA *79*, 7380–7384.

Wakimoto, B. T., and Kaufman, T. C. (1981). Analysis of larval segmentation in lethal genotypes associated with the Antennapedia gene complex in *Drosophila melanogaster*. Dev. Biol. *81*, 51–64.

Wakimoto, B. T., Turner, F. R., and Kaufman, T. C. (1984). Defects in embryogenesis in mutants associated with the Antennapedia gene complex of *Drosophila melanogaster*. Dev. Biol. *102*, 147–172.