Structure of two genes at the gooseberry locus related to the paired gene and their spatial expression during Drosophila embryogenesis

Stefan Baumgartner, Daniel Bopp, Maya Burri, and Markus Noll

Department of Cell Biology, Biocenter of the University of Basel, CH-4056 Basel, Switzerland

The gooseberry (gsb) locus contains two closely linked genes, BSH9 and BSH4, which are structurally related to each other and to the paired (prd) gene. Sequence analysis of genomic DNA and cDNA shows that BSH9 and BSH4 can encode proteins of 427 and 452 amino acids, respectively. The structural homology between these two putative proteins and the prd protein consists essentially of two domains forming most of the amino-terminal halves of the proteins: the prd domain of 128 amino acids and a prd-type homeo domain of 60 amino acids, which is extended by 18 amino acids at its amino-terminal end. The temporal profiles of BSH9 and BSH4 transcripts, as characterized by Northern analysis, show a peak shortly after the peak of prd transcripts. The spatial distributions of BSH9 and BSH4 transcripts have been analyzed by in situ hybridization to whole-mount and sectioned embryos. BSH9 transcripts appear in the posterior ventrolateral part of each primordial segment throughout the embryo, including head and tail segments. Transcripts are initially restricted to the ectoderm, in which they arise as two spatially shifted and temporally delayed waves exhibiting double-segment periodicity and anteroposterior polarity. During germ-band extension, BSH9 is induced in the mesoderm in register with the ectoderm and neurectoderm and in the tail segments A9–A11. In contrast, BSH4 transcripts appear with a single-segment repeat, first, in the neurectoderm during germ-band extension and, later, in single neurons during neuronal differentiation. BSH9, BSH4, and prd are activated in cells that are in register along the anteroposterior axis of the embryo in the posterior parts of primordial segments comprising the posterior compartments of engrailed expression.

[Key Words: Drosophila; gooseberry locus; paired box; homeo box; whole-mount in situ hybridization]

Received July 27, 1987; revised version accepted October 6, 1987.

A genetic analysis of the genes involved in early developmental processes in Drosophila melanogaster has shown that the zygotically acting segmentation genes thought to be responsible for the subdivision of the embryo into metameric units may be divided into three classes according to their embryonic phenotype [Nüsslein-Volhard and Wieschaus 1980; Nüsslein-Volhard et al. 1982]. The existence of these classes implies that the process of segmentation is regulated in three different units of spatial organization. In the first class of segmentation genes, represented by the so-called gap mutants, the spatial units of gene action comprise several neighboring segment primordia, which are deleted in homozygous mutant embryos. The second class of segmentation mutants reveals a repetitive spatial organization with a periodicity of two segments (pair-rule genes), whereas the third class exhibits a spatial unit with a single-segment periodicity (segment–polarity genes). Nüsslein-Volhard and Wieschaus (1980) have proposed that these genes act in a sequential fashion corresponding to a progressive subdivision of the embryo along its anteroposterior axis.

The positional information along the anteroposterior axis that activates the zygotically acting segmentation genes is provided by a small number of maternal genes and deposited into the developing oocyte [Nüsslein-Volhard et al. 1982; Boswell and Mahowald 1985; Degermann et al. 1986; Frohnhöfer and Nüsslein-Volhard 1986; Lehmann and Nüsslein-Volhard 1986; Schüpbach and Wieschaus 1986]. For example, the maternal gene bicoid (bcd) provides positional cues for the development of head and thoracic segments [Frohnhöfer and Nüsslein-Volhard 1986], whereas the seven maternal gap genes of the grandchildless-knirps group determine the development of the abdominal region of the embryo [Lehmann and Nüsslein-Volhard 1986; Schüpbach and Wieschaus 1986]. Support for such a concept of segmentation has been provided by the molecular analysis of the expression patterns of particular segmentation genes by in situ hy-
bridization to wild-type embryos and embryos of the various segmentation mutants [for a recent review, see Scott and O'Farrell 1986]. Such analyses have suggested that segmentation genes are functionally linked in a hierarchical regulatory network (Carroll and Scott 1986; Carroll et al. 1986; Harding and Ingham 1986; Howard and Ingham 1986; Jäckle et al. 1986; Scott and O'Farrell 1986).

Based on evolutionary considerations, we have recently proposed (Bopp et al. 1986; Frigerio et al. 1986) that genes controlling complex functions, for example, development, DNA replication, and metabolic pathways, are integrated into networks by the principle of independent assortment of a limited number of different domains, each defining a homologous set. The diversity of genes belonging to a network would be generated by two variables: first, by alterations in the sequence of a particular homology and, second, by the many diverse ways in which different domains may be combined with each other. For example, genes that determine segmentation in Drosophila embryogenesis are integrated into a network by sharing a relatively small number of characteristic domains. Genes controlling closely linked functions may share specific domains defining a subroutine as, for example, determination of segmentation and segment identity, whereas other domains may be less specific and occur in several networks, ultimately interconnecting all networks to make up the phenotype of an organism. If our concept is correct, it should be possible to isolate the genes belonging to a particular network on the basis of structurally homologous domains by starting from a single isolated gene that is part of the network. This approach would be particularly valuable because no direct method exists at present for higher eukaryotes, which permits the isolation of genes that are linked functionally in a network, as, for example, genes regulated by or in parallel to a particular gene cloned previously.

A scan of the pair-rule gene paired (prd), isolated recently (Kilchherr et al. 1986), for homologous domains shared with other genes of Drosophila has provided some preliminary support for the gene network concept. Three of the genes isolated on the basis of three different structural homologies shared with prd have been analyzed genetically. The genetic data suggest that they are all part of the gene network that is required for the establishment of segmentation. The first gene that was isolated was bcd, which shares a PRD repeat with prd and contains a new type of homeo domain (Frigerio et al. 1986). In agreement with the genetic results of Frohnhöfer and Nüsslein-Volhard (1986), we found that bcd mRNA is localized at the anterior pole of mature oocytes and exhibits a specific gradientlike distribution in the anterior third or half of the embryo during cleavage stage and syncytial blastoderm (Frigerio et al. 1986).

The two other genes, BSH9 and BSH4, are related to each other and have two domains in common with prd: a new type of homeo domain extended at its amino terminus and the "paired" domain consisting of about 130 amino acids. Both genes are closely linked and located at the gooseberry (gsb) locus, a segment–polarity gene (Nüsslein-Volhard and Wieschaus 1980; Nüsslein-Volhard et al. 1984). A preliminary analysis of the spatial distribution of transcripts of these two genes in eukaryotic tissue sections has revealed that both genes are expressed with a single-segment periodicity. However, although transcripts of BSH9 appear in the ectoderm, neur ectoderm, and mesoderm, BSH4 is expressed in the neur ectoderm (Bopp et al. 1986). As the gsb phenotype has been defined by segmentally repeated alterations in the pattern of the ectoderm (Nüsslein-Volhard and Wieschaus 1980; Nüsslein-Volhard et al. 1984), BSH9 alone or in combination with BSH4 could be responsible for the gsb⁺ function.

Here, an analysis of the DNA sequence and transcriptional organization of the two genes, BSH9 and BSH4, at the gsb locus is presented. In addition, the complex changes in spatial distribution of their transcripts during embryogenesis are described in detail.

Results

Transcriptional organization of the gsb locus

A preliminary characterization of the transcriptional organization of the two closely linked genes, BSH9 and BSH4, at the gsb locus and their orientation on the second chromosome has been presented previously (Bopp et al. 1986). In Figure 1a, a more detailed map of their major transcripts is shown, as derived from sequence analysis of genomic and two nearly full-length cDNAs. The BSH9 and BSH4 transcripts consist of two and five exons, respectively, and are transcribed in opposite direction. Their 5' ends of nearly full-length cDNAs are separated by about 10 kb of a nontranscribed sequence that contains the proximal breakpoint of the Df(2R)SB1 deletion (Preiss et al. 1985). This breakpoint has been mapped to within 1.2 kb by restriction analysis of the corresponding region from the Df(2R)SB1 DNA (data not shown).

Sequence of the BSH9 gene

The BSH9 transcription unit was sequenced on both strands of the genomic DNA and a cDNA, except for about half of the 1-kb intron sequence. The length of the cDNA, BSH9c2 (Fig. 1b), is 1564 bp [without poly(A)], which is close to the full length of the 1.8-kb BSH9 transcript [with poly(A)], as judged by Northern analysis (Fig. 5). Therefore, the 5' end of the transcript, which has not yet been determined, is probably not far upstream of the 5' end of BSH9c2.

The following features of the DNA sequence shown in Figure 2 may be of interest. The longest open reading frame starts at the first ATG of the BSH9c2 cDNA and would encode a protein of 427 amino acids. Upstream of this ATG, all three reading frames on the cDNA are closed. The reading frame is interrupted after 18 amino acids by an intron located within the first codon of the paired domain, a homologous region of 128 amino acids,
Relationship and expression of gsb and prd

Figure 1. (a) Transcriptional organization of the major transcripts of BSH4 and BSH9 at the gsb locus. An EcoRI restriction map of the gsb locus at 60F1, close to the telomere of the right arm of the second chromosome (2R), is shown. The major transcripts of the two gsb genes, designated BSH4 and BSH9 according to the recombinant phage DNA on which they have been isolated (Bopp et al. 1986), are displayed below. Exons are represented by filled bars, except for the short open bar at the 3' end of the major BSH4 transcript, which indicates multiple sites of poly(A) addition within this region (cf. Fig. 3). \(P, H\) Exons containing the paired and the prd-homeo domain, respectively. The proximal breakpoint of the deficiency \(Df(2R)SB1\) (Preiss et al. 1985) has been mapped to the region represented by the open area of the arrow shown above the restriction map. (b), (c) Restriction maps of genomic DNA and cDNAs of BSH9 and BSH4. Restriction maps of the genomic regions containing the BSH9 (b) and the BSH4 gene (c) are shown. The extent of the sequenced cDNAs and the corresponding intron regions are indicated below. A small part of the genomic sequence in the 1.3-kb \(BamHI-EcoRI\) fragment covering the 3' portion of BSH9 has not been determined. Its equivalence with the corresponding cDNA sequence is assumed from their comparison by restriction analysis, which revealed no intron.

which it shares with the prd gene (Bopp et al. 1986). As indicated schematically in Figure 4, the paired domain is separated from the second domain homologous to prd, the prd-homeo domain, by 20 amino acids, which are rich in Gly and Ser (60%). This domain may be considered to belong to a separate class of homeo domains and is extended by 18 amino acids at its amino end (Bopp et al. 1986).

The carboxy-terminal half of the putative BSH9 protein contains one large and one small region rich in Ser and Ala (Fig. 4). The first region consists of 39 amino acids [2739–2855 in Fig. 2], 14 Ser and 7 Ala, whereas the second region, consisting of 11 amino acids [3000–3032 in Fig. 2], contains 4 Ser and 4 Ala. The termination codon is followed by an untranslated trailer sequence of 181–185 nucleotides with a canonical poly(A) addition signal, AATAAA, 15–19 nucleotides from the 3' end.

Sequence of the BSH4 gene
The sequence of the BSH4 gene is shown in Figure 3.
Figure 2. (See facing page for legend.)
Relationship and expression of gsb and prd

Figure 2. DNA sequence of the BSH9 gene at the gsb locus and corresponding amino acid sequence of the putative BSH9 protein. The DNA sequence between the first Ball site upstream of the 5' end of the cDNA BSH9c2 and the first EcoRI site downstream of the 3' end of BSH9c2 is shown (cf. Fig. 1a, b). About 450 bp of the first intron have not been sequenced. Hence, the numbering of the nucleotides, indicated in the right margin, is only approximate after this gap of nonsequenced DNA. Above the DNA sequence, the amino acid sequence corresponding to the longest open reading frame is shown. Two deviations of the cDNA from the genomic DNA sequence, one of which results in an altered amino acid at the carboxy-terminal end of the paired domain, are indicated below the corresponding positions. The positions of the 5' and 3' end [site of poly(A) addition] of BSH9c2 and of the cleavage sites of some restriction enzymes are marked by vertical arrows. The canonical poly(A) addition signal is underlined. The sequences corresponding to the paired domain (including its extended portion of homology to BSH4 from nucleotides 956-982, separated from the paired domain by the intron) and the prd type of homeo domain, including its 5'-extended portion [Bopp et al. 1986], are indicated.
Figure 3. (See page 1254 for legend.)
Baumgartner et al.

**Figure 3.** DNA sequence of the BSH4 gene at the gsb locus and amino acid sequence of the putative BSH4 protein corresponding to the longest open reading frame of the major BSH4 transcript. The DNA sequence from the first HindIII site upstream of the 5' end up to about 1 kb downstream of the 3' end of the longest cDNA, BSH4c4, is shown (cf. Fig. 1a, c). A single deviation of the cDNA from the genomic DNA sequence in the leader of BSH4c4 is indicated below. Most of the second intron (3.1 kb) and the third intron (6.0 kb) has not been sequenced. Therefore, the numbering of nucleotides, indicated in the right margin, is only approximate after the second intron. The amino acid sequence corresponding to the longest open reading frame is shown above the DNA sequence. The positions of the 5' and 3' ends of the sequenced cDNAs, including multiple sites of poly(A) addition [the two poly(A) addition signals are underlined], and the cleavage sites of some restriction endonucleases are indicated by vertical arrows. The 5' end of three sequenced cDNAs (BSH4c2, BSH4c5, and BSH4c10) coincided with the EcoRI site at position 5193 of exon 3, which is due to incomplete methylation of the EcoRI sites during preparation of the cDNA library. The sequences corresponding to the paired domain (including the extended portion of homology to BSH9 from nucleotides 1277 to 1303, separated from the paired domain by the first intron), and to the prd type of homeo domain, including its extended portion (Bopp et al. 1986), are indicated.

Except for the large second and third introns of 3.1 kb and 6.0 kb, respectively, which have not been sequenced, most genomic DNA and several cDNA sequences (BSH4c2, BSH4c4, BSH4c5, and BSH4c6, Fig. 1c) have been read on both strands. The longest cDNA, BSH4c4, is close to full length [2166 bp without poly(A)], as compared to a length of 2.5 kb [with poly(A)] of the major BSH4 mRNA observed by Northern blot analysis (see below). However, the exact 5' end of the major BSH4 mRNA has not yet been determined.

This transcript consists of five exons and four introns if no additional introns occur upstream of the 5' end of BSH4c4. The first exon (consisting of at least 460, but more likely of ~600, bases) is followed by a short intron of 327 nucleotides and a small second exon of 314 bases. The third exon, consisting of only 112 nucleotides, is flanked by two large introns, whereas the fourth and fifth exons of 594 and 686 nucleotides, respectively, are separated by a small intron of only 114 bases.

The longest open reading frame starts with the ATG at position 1247 in the first exon (all reading frames of the cDNA upstream of this ATG are closed) and would encode a protein of 452 amino acids (see Fig. 3). As shown in Figure 4, the amino-terminal half of this protein is closely related to that of the BSH9 and prd-gene product. It consists of a paired domain and an extended prd-homeo domain. The homology between the BSH4 and BSH9 paired domains extends by 9 amino acids at the amino end, as compared to the homology they share with the prd protein. In both the BSH4 and BSH9 coding region, the extended portion of the paired domain is separated from the main portion of 128 amino acids, which
Thus, in addition to a major 1.8-kb BSH9 RNA, two minuscors of the major BSH9 transcript, which contains a 1-kb intron (Fig. 1b), or could represent differentially spliced RNAs or RNAs transcribed from different promoters.

The temporal pattern of BSH4 RNAs is even more complicated. In addition to a major transcript of about 2.5 kb, minor RNAs of 2.3, 2.7, and 3.0 kb, as well as a very weak 4.3 kb RNA, are observed (Fig. 5). As shown above, the major BSH4 RNA consists of several transcripts that differ in the lengths of their trailer sequences by about 150 nucleotides. Whereas most of the minor RNAs cannot be explained by partially spliced precursors of the 2.5-kb RNA, this possibility is not excluded for the 3.0-kb RNA.

The spatial patterns of BSH4 during embryogenesis were examined by in situ hybridization to both embryonic tissue sections (Figs. 6 and 7) and whole-mount embryos (Fig. 8). A \(^{3}H\)-labeled 1.3-kb genomic BamHI–EcoRI fragment of BSH9, which contains 1 kb of the 3' terminal portion of BSH9c2, comprising the extended prd-homeo domain but not the paired domain (Bopp et al. 1986; Fig. 2), was used for hybridization. Transcripts of BSH9 are first detected in the developing embryo...
Baumgartner et al.

Figure 5. Northern analysis of developmental profile of prd and gsb transcripts. Poly[A]+ RNA (2 µg each), isolated from follicles, 0 to 2-, 2 to 4-, 4 to 8-, 8 to 12-, 12 to 17-, and 17 to 22-hr-old embryos, from first instar [1–2 days], late second and early third instar [3–4 days], and late third instar larvae [6–7 days] and from adult females was separated in 1.1% agarose gels containing formaldehyde (Lehrach et al. 1977). The RNA was transferred to nitrocellulose filters and hybridized to the following [α-32P]dATP-labeled nick-translated probes. (Top panel) prd cDNA c73.2 (Frigerio et al. 1986); (middle panel) the 1.3-kb genomic BamHI–EcoRI fragment of BSH9 comprising the 3'-terminal portion of BSH9c2; (bottom panel) a 2.4-kb genomic fragment of BSH4 covering the last two exons of BSH4c4. For size calibration, end-labeled fragments of a partial RsaI digest of pBR322 and of a HindIII digest of phage λ DNA were used [not shown].

...odicity corresponding to two-segment primordia. They are shifted with respect to each other by a single-segment primordium, thus generating a combined pattern of 15 bands exhibiting a single-segment periodicity at the onset of gastrulation [Fig. 6e–h]. During the initial phase of germ-band elongation, the width of the bands and the gaps in between are both about two [Fig. 6g,h] to three cells [Fig. 6i,j] wide. However, during further extension of the germ band, the gaps increase to twice the band width [Figs. 6k,l and 7a–d].

...No band is detected in the dorsal region [Fig. 6g,h], as most clearly demonstrated by the pattern observed on whole-mount embryos [Fig. 8a,b]. Only band 0, located anterior to the cephalic furrow, extends nearly to the dorsal midline [Figs. 6i,j, and 8a,b]. As shown in Figure 8, c–h, during the extended germ-band stage, this band retracts ventrally to about the same lateral extent as the more posterior bands. Even-numbered bands arise with a clear ventrodorsal polarity, first being expressed in the ventral portions [Figs. 6e,f and 8a–d] and extending laterally during germ-band extension [Fig. 8e–h].

...During this stage, BSH9 transcripts increase in all 15 bands, forming a regular pattern with a single-segment repeat [Figs. 6g–l and 8a–d]. Interestingly, transcripts are initially absent from the forming ventral furrow, giving rise to the mesodermal germ band [Fig. 6e,f], in which they appear only during germ-band elongation [Fig. 6k,l]. Absence of BSH9 transcripts from the primordial mesoderm at the cellular blastoderm stage is also evident in longitudinal sections along the ventral midline [not shown], as well as in cross sections in which no BSH9 transcripts are detectable in the morphologically distinct cells of the primordial mesoderm but only in the neighboring ectodermal cells [insert in Fig. 6c–f]. By mid-germ-band extension [stage 10 in Campos-Ortega and Hartenstein 1985], BSH9 transcripts are observed in mesoderm, neurectoderm, and ectoderm at equivalent positions along the anteroposterior axis [Fig. 6k,l]. No BSH9 transcripts were detected in the mesoderm of stage-9 embryos [not shown].

...The distribution of BSH9 transcripts after germ-band elongation was examined, as shown in Figure 7. When the germ band reaches its fully extended length, BSH9 transcripts appear in three additional posterior bands [bands 15–17 in Fig. 7a–d], and in three pairs of anterior
regions reflecting the bilateral symmetry of the corresponding anlagen (two lateral pairs L1 and L2; Figs. 7a–d and 8e–h), as well as a pair of anterior dorsal spots (AD in Figs. 7a,b and 8e,f). In addition, transcripts are observed in a small region of the dorsal junction between posterior midgut and proctodeum (below band 15 in Fig. 7a,b). This region, which might be at the origin of the Malpighian tubules, comprises only a few cells; hence, transcripts are only detectable in suitable sections. In fact, these transcripts already appear at early stage 10 [Campos-Ortega and Hartenstein 1985]. At the beginning of germ band retraction, transcripts in these

Figure 6. (See next page for legend.)
cells are no longer observed. The variable relative band widths and intensities and the occasional absence of bands observed in the plane of section (Fig. 7a–d) indicate that not all cells at the same distance along the extended germ band express BSH9.

During germ band shortening, BSH9 transcripts disappear from the bilateral AD regions and from mesodermal cells of bands 0–17 (Fig. 7e,f). In contrast, transcripts increase in the ectoderm of the anterior bands 0 and 1, which appear to fuse (Figs. 7g,h and 8i,j), and of band 17 at the posterior end of the retracting germ band (Fig. 7e–h). In addition, BSH9 transcripts accumulate transiently in the anterior ventral cord (Fig. 7g,h). It may be worthwhile noting that at this stage, ectodermal expression extends farther laterally in A1–A4 than in the thoracic and more posterior abdominal segments where it has been reduced to a narrow mid-ventral portion (Fig. 8i,j).

During head involution, BSH9 transcripts also disappear from the ectoderm, first from the thoracic (Fig. 8i–l) and later from the abdominal ectoderm and the ectoderm of the head segments (Fig. 7i,j). Strongly labeled cells remain only below the pharynx posterior to the region of the forming salivary gland and in the anal pads (Fig. 7i,j). During further condensation of the ventral cord, transcripts disappear from the anal pads (Fig. 7k,l) and, by 18 hours of embryogenesis at 25°C, from the entire embryo (not shown).

Spatial transcription patterns of BSH4 during embryogenesis

The distribution of BSH4 transcripts was examined using as hybridization probe a 3H-labeled 2.4-kb genomic fragment of BSH4 covering the last two exons. BSH4 transcripts are detected considerably later than those of BSH9 during embryogenesis. As shown in Figure 9, BSH4 transcripts are first detected in single neuroblasts of the extending germ band (Fig. 9a,b) and, slightly later, of the procephalic lobe (not shown). During the extended germ band stage, a pattern with a single-segment repeat, very similar to that of BSH9, becomes apparent. It consists of 19 segmental (0–17, AD) and 2 nonsegmental clusters, L1 and L2, of neuronal precursor and in some epithelial cells. During further retraction of the germ band, BSH4 transcripts disappear except in regions 0 and 17, where they appear to accumulate even further (Fig. 9g,h). In addition, transcripts appear in a small number of specific neurons per segment spaced at a segmental repeat (Fig. 9g–l). It remains to be elucidated whether these transcripts persist or arise in newly induced neurons. In each ganglion, these neurons are found preferentially in dorsolateral parts of the ventral cord and in the supraesophageal ganglion. No label has been detected in the subesophageal ganglion. We have found expression in single neurons of the 12 ganglia of parasegments 3–14 (Martinez-Arias and Lawrence 1985). Expression of BSH4 in the ventral nervous system persists, presumably in the same neurons, throughout larval development (not shown).

Localization of BSH9 transcripts within primordial segments

As described above, when the odd-numbered bands of BSH9 arise during germ band elongation, alternating bands and gaps of about equal width corresponding to half-primordial segments are detected (Figs. 6e–j and 8c,d). During germ band elongation, the gaps increase until they reach about twice the width of the bands (Figs. 6k,l, 7a–d, and 8e–h). Thus, by the end of germ band extension, BSH9 is expressed over the width of one-third of a segment. As evident from Figure 8i,j, cells expressing BSH9 appear immediately anterior to the posterior segment boundary in the thoracic and abdominal
Figure 7. Distribution of BSH9 transcripts in tissue sections of embryos after germ band elongation. Photomicrographs of the same embryos taken under phase-contrast (left) or dark-field (right) illumination are shown. Sagittal sections of embryos are oriented as in Fig. 6. Stages (according to Campos-Ortega and Hartenstein 1985) shown correspond to the extended germ band (early stage 11) at ~5.5 hr of embryonic development at 25°C (a–d), to germ band retraction at ~9 hr (late stage 12) (e, f) and 9.5 hr (early stage 13) (g, h), to head involution and dorsal closure at about 11.5 hr (early stage 15) (i, j), and to an 18- to 22-hr-old embryo before hatching (stage 17) (k, l). Hybridization and exposure were as described in the legend to Fig. 6. (k) Bar represents 0.1 mm. Abbreviations: (am and pm) Anterior and posterior midgut; (ap) anal pad; (cl) clypeolabrum; (de) dorsal epidermis; (np) neuropile; (pv) proventriculus; (sbg and spg) subesophageal and supraesophageal ganglion; (vc) ventral cord. For other abbreviations, see legend to Fig. 6.
segments. In addition, the visible segment borders in Figure 8i permit us to assign the bands to specific segments. Thus, bands 4–6 are in posterior T1–T3, and bands 7–14 are in posterior A1–A8, whereas bands 15–17 may be assigned to primordial segments A9–A11 of the tail region [Jürgens 1987]. Consistent with this interpretation, band 17 clearly labels the anal pads (Fig. 7g–i), originating from A10 and A11 [Jürgens 1987]. Additional evidence locating BSH9 transcripts in the posterior compartments of T1–T3 and A1–A8 is provided by the observation that the anterior boundaries of BSH9 expression coincide with or appear slightly posterior to the parasegmental grooves [Martínez-Arias and Lawrence 1985] in that region (Fig. 7a,c). Bands 0–3 are expressed in the primordial head segments, band 3 in the labial, band 2 in the maxillary, band 1 in the mandibular, and band 0 in the premandibular or intercalary head segment [Jürgens et al. 1986]. The AD regions expressing BSH9 (Fig. 7a,b) are located within the labral head segment, whereas L1, which may be part of the anlage of the antennal sense organ [Turner and Mahowald 1979; Jürgens et al. 1986], and L2 appear within the procephalic lobe at the time of neuroblast division during stage 10 (Campos-Ortega and Hartenstein 1985).

Relative location of BSH4, BSH9, prd, and engrailed transcripts

Hybridization to adjacent serial sections to detect BSH4 (Fig. 10a), BSH9 (Fig. 10b), or both BSH4 and BSH9 transcripts [Fig. 10c,d] demonstrates that BSH4 and BSH9 transcripts are in register along the anteroposterior axis. Therefore, BSH4 is also expressed in the posterior parts of the thoracic and abdominal segments T1–A8 [bands 4–14]. Also, the remaining regions expressing BSH4 correspond to the primordial head [bands 0–3 and AD] and tail [bands 15–17] segments, with the exception of L1 and L2, which appear in neurogenic regions of the procephalic lobe.

A similar experiment localizing BSH9 with respect to prd transcripts is illustrated in Figure 10e–h. It demonstrates that the odd-numbered bands 1–13 of BSH9 overlap and probably coincide with the bands of prd at early gastrulation. Band 0 of BSH9 is anterior to band 1 of prd transcripts. As we have previously assigned this band to the mandibular segment primordium [Kilchherr et al. 1986], band 0 of BSH9 would be in the next anterior segment, that is, in the premandibular or intercalary segment. This conclusion is consistent with the assignment of BSH9 bands to primordial segments on the basis of the visible segment boundaries in whole-mount embryos [Fig. 8i].

Because BSH9 is transcribed in the posterior parts of primordial segments, we expect its transcriptional pattern to overlap with that of engrailed (en). The patterns observed after in situ hybridizations to the serial sections shown in Figure 10, i–l, confirm this interpretation, although the bands of the superimposed patterns [Fig. 10k,l] are slightly broader than those of en (Fig. 10i) or BSH9 (Fig. 10j) in the abdominal segments A1–A4. In other serial sections, a broadening of these bands in the superimposed pattern with respect to BSH9 bands is not observed [not shown]. Thus, we attribute this effect to a slight variability of the BSH9 band width, depending on the plane of section as clearly observed in Figure 7a–d. As en expression probably marks posterior compartments of larval segment primordia [Kornberg 1981a,b], it follows that at the extended germ band stage, BSH9 and BSH4 are active in the posterior third of thoracic and abdominal segments. At the beginning of germ band elongation, prd and BSH9 are expressed in register in half-primordial segments, two cells wide, which probably also correspond to posterior segment primordia.

**Discussion**

The gsb locus appears to consist of two closely linked genes, BSH9 and BSH4, that are divergently transcribed in segmentally repeated patterns (Bopp et al. 1986). Because only two deletions and no point mutations of gsb have been isolated [Nüsslein-Volhard et al. 1984], it remains to be demonstrated whether or not both genes are functional and responsible for the segment—polarity phenotype produced by the two deletions [Nüsslein-Volhard and Wieschaus 1980]. Although one of the two gsb deficiencies removes only the BSH9 transcriptional unit, a concomitant cis- or trans-inactivation of BSH4 has not yet been excluded.

Detailed sequence analysis, summarized in Figure 4, shows that both genes are related to each other as well as to the prd gene with which they share two domains in their amino-terminal half—the paired domain [a large domain of 128 amino acids] and the prd type of homeo domain [Bopp et al. 1986]. In contrast, the carboxy-terminal portions of these genes do not share long stretches of sequence similarity.
of homologous sequences. The only intron of BSH9 interrupts the open reading frame at the amino-terminal end of the paired domain. The first intron of BSH4 occurs at precisely the same position, whereas the intron of prd is shifted by four amino acids toward the amino-terminal end [Fig. 4]. These findings suggest that the amino-terminal portions of BSH9 and BSH4 arose by duplication after their first intron appeared. One of them

Figure 8. [See facing page for legend.]
might have been generated by earlier duplication and recombination events of the amino-terminal half of the prd gene.

The structural relationship of BSH9 and BSH4 is reminiscent of that of en and invected (inv) at the en locus (Coleman et al. 1987). These two genes, which are expressed in a segmentally repeated pattern similar to that of BSH9 and BSH4, are also juxtaposed and show extensive structural homology to each other, including a homeo domain. Furthermore, expression of inv is delayed with respect to that of en (Coleman et al. 1987), as BSH4 is activated after BSH9 (Figs. 6 and 9). However, BSH4 and BSH9 differ from en and inv in an important aspect. Whereas en and inv are expressed in the same cells during development (Coleman et al. 1987), BSH4 is initially transcribed in a subset of the cells in which BSH9 is active, namely in the posterior third of segments of the neurectoderm, and later, during neuronal differentiation, its activity becomes restricted to a few specific neurons per segment.

The characteristic spatial distribution of BSH9 transcripts, which exhibits a single-segment repeat during germ band elongation, is generated from two patterns of a double-segment periodicity that are shifted by a single primordial segment. The two patterns arise with an anterior-polar and are temporally delayed with respect to each other. Before the posterior bands of the first double-segment repeat (odd-numbered bands) appear, the anterior bands of the second double-segment repeat (even-numbered bands) emerge. Thus, the final pattern may be described as arising from two spatially shifted and temporally delayed waves running from the anterior to the posterior pole of the embryo. Furthermore, the bands of the second wave appear with a ventrodorsal polarity. Due to the relatively short time interval by which the second pattern is delayed in the midventral region, an intermediate pattern with a double-segment periodicity is apparent only in lateral parts of the embryo where the lag of the second wave is longer (Figs. 6e, f, and 8c, d). Interestingly, BSH9 transcripts are restricted to the ventrolateral regions of the ectoderm, whereas induction of BSH9 in the mesoderm occurs only after the second mesodermal mitosis.

The induction of BSH4 is delayed until the beginning of neurogenesis when its transcripts appear in what probably are single neuroblasts and their progeny (Fig. 9a–d). During neuronal differentiation, BSH4 expression is restricted to specific neurons at segmentally repeated homologous locations. Such expression in single neurons at a single-segment repeat has also been observed for the pair-rule genes ftz (Carroll and Scott 1985, Hiromi et al. 1985), en (DiNardo et al. 1985; Ingham et al. 1985; Weir and Kornberg 1985), and eve (Macdonald et al. 1986; Frasch et al. 1987). In addition, many homeotic genes are expressed in the central nervous system (see, e.g., Levine and Wedeen 1985). Recently, cell-surface glycoproteins that are expressed in specific neurons spaced in a segmentally repeated pattern and that are probably required for specific cell–cell recognition during fasciculation of axons have been found in the grasshopper (Bastiani et al. 1987) and in Drosophila (Patel et al. 1987). Hence, it is attractive to speculate that the expression of such cell-surface molecules is regulated by the combinatorial expression of segmentation (Patel et al. 1987) and homeotic genes.

Several observations indicate that BSH9 and BSH4 are expressed in posterior compartments. In whole-mount preparations of embryos, in which ectodermal segmentation has become visible, BSH9 transcripts are found only in the posterior portions of thoracic and abdominal segments (Fig. 8i, j). In addition, in sections of embryos at the extended germ band stage, the regions immediately posterior to the parasegmental grooves are labeled (Fig. 7a). Because these grooves are at the boundary between anterior and posterior compartments (Martinez-Arias and Lawrence 1985), we conclude that BSH9 is expressed in posterior compartments at this stage. Finally, serial sections probed with en, which is expressed in the posterior third or quarter of segments (DiNardo et al. 1985; Fjose et al. 1985; Ingham et al. 1985; Kornberg et al. 1985; Weir and Kornberg 1985; Howard and Ingham 1985), BSH9 or a combination of these two probes demonstrate that expression of BSH9 coincides with that of en in bands 0–14, that is, in the hypopharyngeal, gnathal, thoracic, and abdominal segments A1–A8. However, the band width of BSH9 transcripts (Fig. 10i) is slightly broader than that of en (Fig. 10i). Because there is no indication of BSH9 transcripts in the anterior-most region of segments (i.e., immediately posterior to en expression), BSH9 is probably also expressed slightly anterior to en. The conclusion that BSH9 activity is restricted to posterior regions of thoracic and abdominal segments is in good agreement with the gsb− phenotype in which the posterior parts of segments are deleted and replaced by mirror-image duplications of the remaining anterior portions of these segments (Nüsslein-Volhard and Wieschaus 1980).

At the extended germ band stage, BSH9 expression is also regularly spaced in the tail region and appears to label each of the abdominal segments A9–A11. The observation that BSH9 transcripts do not appear in tail segments before this stage is consistent with the view that development of the tail region is delayed with respect to the other segments and that its subdivision into metameric units occurs only when it has been sufficiently stretched out at the fully extended germ band stage (Jürgens 1987).

Shortly before head involution, when segment boundaries are visible, BSH9 transcripts are not predominantly expressed in posterior regions of the head segments. Thus, BSH9 transcripts appear in the middle of the labial segment, in the portion of the maxillary segment adjacent to the mandibular lobe, and in most of the mandibular segment (Fig. 8i, j). Hence, it was surprising to find that BSH9 and en expression also coincide in the head segments (Fig. 10i, j). However, it is possible that cells expressing BSH9 in the head segments at an earlier stage (e.g., extended germ band) are located in the posterior portions and that their position changes during morphogenetic movements in the head. Considering the regular spacing along the entire embryo of cells expressing BSH9...
or *en* during germ band elongation, it might be assumed, in agreement with the postulated role of the *en* gene (Morata and Lawrence 1975, 1978, 1979; Kornberg 1981a, b; Struhl 1981a, b; Lawrence and Struhl 1982), that *en* (DiNardo et al. 1985; Ingham et al. 1985; Komberg et al. 1985) and BSH9 label posterior compartments also in the larval head segments. However, in contrast to *en*, BSH9 is not transcribed throughout larval development (Fig. 5) and, hence, might be considered to act as an activator gene rather than a selector gene (Garcia-Bellido 1975) in the process of segmentation.

The observed correspondence of spatial expression (Fig. 10) and protein structure (Fig. 4) between *prd* and the two *gsb* genes might suggest a functional link be-

![Figure 9](image-url)
between these three genes, as predicted by the gene network concept (Bopp et al. 1986; Frigerio et al. 1986). Although an obvious possibility is that these genes are regulated in parallel by the same gene product(s), we have observed a regulatory link of hierarchical nature between them. The activation of both gsb genes requires an active prd-gene product in every other segment (S. Baumgartner, unpubl.), indicating that prd participates in the induction of gsb transcription at these locations. The induction of gsb by the prd gene product might sug-

Figure 10. (See facing page for legend.)
gest that prd and the two gsb genes are part of a developmental pathway that determines segmentation and progressively defines cellular fates along the anteroposterior axis of the embryo. Hence, all three genes appear to belong to the same gene network, which establishes segmentation and cellular identity during embryogenesis.

Materials and methods

General procedures

A cDNA library, constructed in λgt10 of poly(A)+ RNA from 3- to 12-hr-old embryos [Poole et al., 1985], kindly provided by Tom Kornberg, was screened for cDNAs of BSH9 and BSH4 according to standard procedures [Maniatis et al. 1982]. A recombinant phage containing the deficiency Df(2R)SB1 was isolated from a genomic library, prepared in EMBL4 [Frischauf et al. 1983] from DNA of Df(2R)SB1/CyO flies, and the proximal breakpoint of the deletion was mapped by restriction analysis and comparison with the map of the corresponding wild-type DNA clone isolated previously [Bopp et al. 1986].

DNA sequencing

Figure 1 shows restriction maps of the gsb region and the various isolated cDNAs. Except for the small first intron of BSH4 and the short regions upstream of the SalI site in BSH9 (Fig. 1b) and of the first exon in BSH4c4 (Fig. 1c), of which only one strand was sequenced, all DNA sequences were analyzed on both strands. For some sequences, the sequencing strategy described by Henikoff [1984] was used. The DNAs were sequenced by the dideoxynucleotide method of Sanger et al. [1977], using the M13 vector mWB3296 (Frigerio et al. 1986) or its counterpart, mWB3226, which contains the M13mp18 polylinker in opposite direction. Both were derived from M13 vectors described by Barnes et al. [1983].

In situ hybridization to embryonic tissue sections and to whole-mount embryos

In situ hybridizations to embryonic tissue sections were carried out essentially as described by Hafen et al. [1983]. The specific activity of all probes, nick translated in the presence of 3H-labeled TTP, dCTP, and dGTP and nonlabeled dATP was ∼2 × 10^6 dpm/µg. Purified DNA fragments were nick translated to generate a mean single-stranded DNA length of 20–50 nucleotides.

For hybridization to whole-mount embryos, embryos were dechorionated and devitellinized according to Mitchison and Sedat [1983]. The following steps were modified from a procedure of Kornberg et al. [1985], and all were carried out at room temperature. After rehydration in phosphate-buffered saline (PBS) (130 mM NaCl, 7 mM Na2HPO4, 3 mM NaH2PO4·H2O), the embryos were fixed again for 20 min in 4% paraformaldehyde in PBS and washed for 5 min in PBS. Excess liquid was removed with blotting paper, and the embryos were transferred to 2.2-ml microtubes (Eppendorf no. 3812). Depending on the number of embryos, 50–150 µl of hybridization buffer and the same volume of radioactive probe (1.6 µg DNA/ml), boiled for 2 min, were added to the tube. Incubation was carried out for 16 hr at 37°C on a turning wheel. The embryos were washed six times with PBS at 35°C during an 8-hr period and dehydrated, first in 70% and then in 100% ethanol. Subsequently, the dehydrated embryos were pipetted onto gelatin-subbed slides, air-dried, and dipped into emulsion for autoradiography.

Acknowledgments

We are indebted to Tom Kornberg for a cDNA library and the c-2.4 en cDNA clone, to Christiane Nüsslein-Volhard for the prd probe, and to Herbert Jäckle for the Df(2R)SB1 mutant stock. We are grateful to our colleagues for many stimulating discussions and to Leslie Pick and Hans Noll for comments on the manuscript. This work was supported by Swiss National Science Foundation grants 3.600-0.84 and 3.348-0.86 and by the Kanton Basel.

References

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

Bastiani, M.J., A.L. Harrelson, P.M. Snow, and C.S. Goodman. 1987. Expression of fasciclin I and II glycoproteins on subsets of axon pathways during neuronal development in the grasshopper. Cell 48: 745–755.

Bopp, D., M. Burri, S. Baumgartner, G. Frigerio, and M. Noll. 1986. Conservation of a large protein domain in the segmentation gene paired and in functionally related genes of Drosophila. Cell 47: 1033–1040.

Boswell, R.E. and A.P. Mahowald. 1985. tudor, a gene required for assembly of the germ plasm in Drosophila melanogaster. Cell 43: 97–104.

Campos-Ortega, J.A. and V. Hartenstein. 1985. The embryonic development of Drosophila melanogaster. Springer-Verlag, Berlin.

Carroll, S.B. and M.P. Scott. 1985. Localization of the fushi tarazu protein during Drosophila embryogenesis. Cell 43: 47–57.

Note that in Figure 10, localization of gsb (BSH9 and BSH4), relative to prd and en transcripts. Sagittal serial sections of embryos at early gastrula (e–h) were hybridized with the following 3H-labeled probes: BSH4 [a], BSH9 [b], and a combination of BSH9 and BSH4 [c,d], with prd [e] in BSH9 [f], and prd and BSH9 [g,h], and with en [i] in BSH9 [i], and a mixture of en and BSH9 [j,k,l]. The probes for BSH9 and BSH4 are the same as those in Figs. 6–9. The prd probe c7340.1 [Frigerio et al. 1985] and the en probe c2.4 [Poole et al. 1985] have been described. Serial sections of each of the three embryos are shown in consecutive order [a,b,c,d; e,f,g,h;i,j,k,l], with their anterior end to the left and, in sagittal sections, with their dorsal side up. All sections are shown under dark-field illumination, the last section in each set is also shown under bright-field illumination. Note that in Figs. a–h, only the odd-numbered bands of BSH9 extend sufficiently laterally to be cut by the horizontal sections (cf. Fig. 8c,d) and that band 0 of en [Ingham et al. 1985], which is not visible in the section shown in i, and of BSH9 also coincide, as evident from autoradiographs of other serial sections. Autoradiographic exposure was for 28 days. (k) Bar represents 0.1 mm. For abbreviations, see legend to Fig. 6.
pression of the fushi tarazu segmentation gene during early Drosophila embryogenesis. Cell 45: 113–126.

Carroll, S.B., G.M. Winslow, T. Schübchach, and M.P. Scott. 1986. Maternal control of Drosophila segmentation gene expression. Nature 323:278–280.

Coleman, K.G., S.J. Poole, M.P. Weir, W.C. Soeller, and T. Kornberg. 1987. The inverted gene of Drosophila: Sequence analysis and expression studies reveal a close kinship to the engrailed gene. Genes Dev. 1: 19–28.

Degelmann, A., P.A. Hardy, N. Perrimon, and A.P. Mahowald. 1986. Developmental analysis of the torso-like phenotype in Drosophila produced by a maternal-effect locus. Dev. Biol. 115: 479–489.

DiNardo, S., J.M. Kuner, J. Theis, and P.H. O’Farrell. 1985. Development of embryonic pattern in D. melanogaster as revealed by accumulation of the nuclear engrailed protein. Cell 43: 59–69.

Fjose, A., W.J. McGinnis, and W.J. Gehring. 1985. Isolation of a homeo box-containing gene from the Drosophila embryo by the maternal-effect locus. Dev. Biol. 115: 479–489.

Garcia-Bellido, A. 1975. Genetic control of wing disc development in Drosophila melanogaster. Wilhelm Roux’s Arch. Dev. Biol. 196: 141–157.

Jürgens, G., R. Lehmann, M. Scharl, and C. Nüsslein-Volhard. 1986. Segmental organisation of the head in the embryo of Drosophila melanogaster. Wilhelm Roux’s Arch. Dev. Biol. 195:359–377.

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

Kornberg, T. 1981a. engrailed: A gene controlling compartment and segment formation in Drosophila. Proc. Natl. Acad. Sci. 78: 1095–1099.

——. 1981b. Compartments in the abdomen of Drosophila and the role of the engrailed locus. Dev. Biol. 86: 363–372.

Kornberg, T., I. Siden, P. O’Farrell, and M. Simon. 1985. The engrailed locus of Drosophila: In situ localization of transcripts reveals compartment-specific expression. Cell 45: 45–53.

Lawrence, P.A. and G. Struhl. 1982. Further studies of the engrailed phenotype in Drosophila. EMBO J. 7: 827–833.

Lehmann, R. and C. Nüsslein-Volhard. 1986. Abdominal segmentation, pole cell formation, and embryonic polarity require the localized activity of oskar, a maternal gene in Drosophila. Cell 47: 141–152.

Lehrach, H., D. Diamond, J.M. Wozney, and H. Boedtker. 1977. RNA molecular weight determinations by gel electrophoresis under denaturing conditions, a critical reexamination. Biochemistry 16: 4743–4751.

Levine, M.S. and C.J. Wedeen. 1985. Homeotic gene expression in Drosophila. Trends Neurosci. 8: 239–245.

Macdonald, P.M., P. Ingham, and G. Struhl. 1986. Isolation, structure, and expression of even-skipped: A second pair-rule gene of Drosophila containing a homeo box. Cell 47: 721–734.

Maniatis, T., E.F. Fritsch, and J. Sambrook. 1982. Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Martinez-Arias, A. and P.A. Lawrence. 1985. Parasegments and compartments in the Drosophila embryo. Nature 313: 639–642.

Mitchison, T.J. and J. Sedat. 1983. Localization of antigenic determinants in whole Drosophila embryos. Dev. Biol. 99: 261–264.

Morata, G. and P.A. Lawrence. 1975. Control of compartment development by the engrailed gene in Drosophila. Nature 255: 614–617.

——. 1978. Anterior and posterior compartments in the head of Drosophila. Nature 274: 473–474.

——. 1979. Development of the eye-antenna imaginal disc of Drosophila. Dev. Biol. 70: 355–371.

Nüsslein-Volhard, C. and E. Wieschaus. 1980. Mutations affecting segment number and polarity in Drosophila. Nature 287: 789–801.

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

Patel, N.H., P.M. Snow, and C.S. Goodman. 1987. Characterization and cloning of fasciclin III: A glycoprotein expressed on a subset of neurons and axon pathways in Drosophila. Cell 48: 975–988.
Poole, S.J., L.M. Kauvar, B. Drees, and T. Kornberg. 1985. The engramed locus of Drosophila: Structural analysis of an embryonic transcript. Cell 40: 37–43.

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

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

Schüpbach, T. and E. Wieschaus. 1986. Maternal-effect mutations altering the anterior-posterior pattern of the Drosophila embryo. Wilhelm Roux’s Arch. Dev. Biol. 195: 302–317.

Scott, M.P. and P.H. O’Farrell. 1986. Spatial programming of gene expression in early Drosophila embryogenesis. Annu. Rev. Cell Biol. 2: 49–80.

Struhl, G. 1981a. Anterior and posterior compartments in the proboscis of Drosophila. Dev. Biol. 84: 372–385.

———. 1981b. A blastoderm fate map of compartments and segments of the Drosophila head. Dev. Biol. 84: 386–396.

Turner, F.R. and A.P. Mahowald. 1979. Scanning electron microscopy of Drosophila melanogaster embryogenesis. III. Formation of the head and caudal segments. Dev. Biol. 68: 96–109.

Weir, M.P. and T. Kornberg. 1985. Patterns of engramed and fushi tarazu transcripts reveal novel intermediate stages in Drosophila segmentation. Nature 318: 433–439.
Structure of two genes at the gooseberry locus related to the paired gene and their spatial expression during Drosophila embryogenesis.

S Baumgartner, D Bopp, M Burri, et al.

*Genes Dev.* 1987, 1:
Access the most recent version at doi:10.1101/gad.1.10.1247

References

This article cites 52 articles, 4 of which can be accessed free at:

http://genesdev.cshlp.org/content/1/10/1247.full.html#ref-list-1

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.