The EMBO Journal vol.10 no.8 pp.2279-2289, 1991

Murine genes related to the *Drosophila AbdB* homeotic gene are sequentially expressed during development of the posterior part of the body

Juan-Carlos Izpisúa-Belmonte, Hildegard Falkenstein, Pascal Dollé, Armand Renucci and Denis Duboule

European Molecular Biology Laboratory, Postfach 10.2209, D-6900 Heidelberg, Germany

Communicated by G.Morata

The cloning, characterization and developmental expression patterns of two novel murine *Hox* genes, *Hox-4.6* and *Hox-4.7*, are reported. Structural data allow us to classify the four *Hox-4* genes located in the most upstream (5′) position in the HOX-4 complex as members of a large family of homeogenes related to the *Drosophila* homeotic gene *Abdominal B* (*AbdB*). It therefore appears that these vertebrate genes are derived from a selective amplification of an ancestral gene which gave rise, during evolution, to the most posterior of the insect homeotic genes so far described. In agreement with the structural colinearity, these genes have very posteriorly restricted expression profiles. In addition, their developmental expression is temporally regulated according to a cranio-caudal sequence which parallels the physical ordering of these genes along the chromosome. We discuss the phylogenetic alternative in the evolution of genetic complexity by amplifying either genes or regulatory sequences, as exemplified by this system in the mouse and *Drosophila*. Furthermore, the possible role of 'temporal colinearity' in the ontogeny of all coelomic (metamerized) metazoans showing a temporal antero-posterior morphogenetic progression is addressed.

Key words: colinearity/homeobox/morphogenesis

Introduction

The mammalian genome contains ~40 members of the '*Hox*' gene family (for references see Kessel and Gruss, 1990; Simeone *et al.*, 1990a,b). These genes are clustered in four complexes (HOX-1, -2, -3 and -4) and encode proteins (homeoproteins) harbouring a homeodomain related to that present in the *Drosophila* homeotic genes (see Scott *et al.*, 1989 for references and review). Homeoproteins are likely transcription factors (e.g. Desplan *et al.*, 1988; Hoey and Levine, 1988; Otting *et al.*, 1988) involved in key control morphogenetic events such as instructing various cells regarding their positions and fates along the major body axes (e.g. Gaunt *et al.*, 1988; Holland and Hogan, 1988; Kessel and Gruss, 1990). Though the target genes of such homeoproteins are not yet known, it is believed that various combinations of such proteins could differentially affect the regulation of the same or different genes located 'downstream' in developmental pathways. Recent results obtained by introducing perturbations in such homeoprotein combinations support this hypothesis (Balling *et al.*, 1989; Kessel *et al.*, 1990).

DNA and protein sequence comparisons between *Hox* genes suggest that gene duplication generated, during evolution, the ancestral *Hox* complex which was in turn duplicated to give rise to the multicomplex organization reported in most of the vertebrate species studied so far (e.g. Acampora *et al.*, 1989; Duboule and Dollé, 1989; Graham *et al.*, 1989; Kappen *et al.*, 1989). These large-scale duplication steps may not have occurred in arthropods since *Drosophila* contains only one copy of such an ancestral complex. In this case, it appears that this complex was split to give rise to the two known *Bithorax* and *Antennapedia* homeotic gene complexes (BX-C; ANT-C; Lewis, 1978; Kaufman *et al.*, 1983), whereas other arthropods such as beetles may still have a unique cluster of homeotic genes (Beeman, 1987). It thus becomes clear that this family of genes has been strictly conserved during evolution in most, if not all animal species that show an antero-posterior (AP) asymmetry (reviewed in Akam, 1989).

A detailed study of the expression domains of these various genes during the development of both *Drosophila* and mammals further demonstrates their phylogenetic linkage and suggests a possible conservation of some of their functional features. The order of the expression domains of the rodent *Hox* genes along the developing AP axis (Gaunt *et al.*, 1988; Duboule and Dollé, 1989; Graham *et al.*, 1989), like the order of the structures specified by the *Drosophila* homeotic genes (Lewis, 1978; Harding *et al.*, 1985), is colinear with the arrangement of these genes along the clusters. Consequently, the most 3′-located genes are those expressed most anteriorly in the vertebrates or those determining the most anterior-located structures in *Drosophila* (Akam, 1989). These genes are members of the labial-like family (*lab* in *Drosophila*, Mlodzick *et al.*, 1988; *Hox-1.6* and *-2.9* in the mouse, Baron *et al.*, 1987; Murphy *et al.*, 1989; Wilkinson *et al.*, 1989; Frohman *et al.*, 1990). Similarly, it was expected that those vertebrate genes which are cognates of the most posteriorly expressed *Drosophila* homeotic gene, the *Abdominal B* gene (*AbdB*; Sanchez-Herrero *et al.*, 1985; Tiong *et al.*, 1985; Casanova *et al.*, 1987), would be expressed in the most posterior parts of the developing body. Interestingly, however, these genes (the *AbdB*-like genes; *Hox-4.4*, Duboule and Dollé, 1989; *Hox-2.5*, Bogarad *et al.*, 1989; Graham *et al.*, 1989; *Hox-1.7*, Rubin *et al.*, 1987; and *Hox-3.2*, Erselius *et al.*, 1990) did not reveal an extreme posterior expression pattern but rather suggested an involvement in the ontogeny of 'intermediate' structures. Therefore it seemed likely that additional genes, exhibiting a more posterior expression pattern, should be present in the 5′ extremity of the HOX-4 complex. This proved to be the case with the identification of the *Hox-4.5* gene (Duboule and Dollé, 1989). Upon further characterization and cloning of this 5′ region, we have now identified two additional genes (*Hox-4.6* and *-4.7*) located at more upstream positions

in the HOX-4 complex, and there are thus at least three genes in this complex which could not be aligned with the *Drosophila* homeotic gene complexes (see Figure 2). We have addressed the origin of these genes, their possible functions as the vertebrate most 'posterior' *Hox* genes, and their counterparts in *Drosophila*.

In this paper we show that these genes are indeed very posteriorly expressed in agreement with the structural colinearity. We also demonstrate that these genes are temporally and sequentially activated shortly after the establishment of the embryo AP axis, which allows us to extend the principle of 'temporal colinearity', observed during limb morphogenesis (Dollé *et al.*, 1989), to the development of structures along the major body axis. With respect to phylogeny, we show that all these genes are derived from an amplification of an original vertebrate gene cognate of the *Drosophila AbdB* homeotic gene. Thus, in the course of evolution, additional duplications of the most posterior gene in the ancestral HOX/HOM complex occurred in the vertebrates, most likely to increase the information required in parallel with the differential specification of the most posterior metameres. This may represent a phylogenetic alternative to the presence, in *Drosophila*, of multiple posterior parasegment-specific *cis*-regulatory elements which control the expression of the *AbdB* gene.

Results

Cloning and primary structures of Hox-4.6 and Hox-4.7

The *Hox-4.6* and *-4.7* genes (formerly called *Hox-5.5* and *-5.6*, see Duboule *et al.*, 1990) were isolated by chromosome walking along the 5′ part of the HOX-4 complex. Their locations within this complex and direction of transcription are depicted in Figure 1. The positions of these two novel genes relative to the other members of the Hox network (as judged by the alignment of the various complexes) are shown in Figure 2. This scheme illustrates the fact that the *Hox-4.6* and *-4.7* genes (as well as *Hox-4.5*) are located upstream of the *AbdB*-like subfamily and therefore cannot be aligned with a *Drosophila* counterpart, unlike the other *Hox* genes. Though no murine gene cognates of these 'upstream' genes have yet been reported in the HOX-1, -2 or -3 complexes, their presence in the human HOX1 and HOX3 (e.g. Acampora *et al.*, 1989) strongly suggests their existence in rodents.

Our interest was raised concerning the phylogenetic origin of these novel genes and thus their complete coding sequences were established from various cDNA and genomic clones. Only partial cDNA clones were obtained and the sequences were thus completed with genomic DNA. In both cases, a unique mRNA species was seen by Northern blot analysis using fetal mRNAs from either limbs or posterior carcass; ~2 kb and ~2.5 kb in size for *Hox-4.6* and *-4.7*, respectively (not shown). The general structures of the *Hox-4.6* and *-4.7* proteins fit well with those of the classical vertebrate homeoproteins. In both cases, the proteins are encoded by two exonic sequences separated by rather small introns (Figures 1 and 3A and B). The homeodomain is located in the second exon and protein termination occurs soon after the C-terminal end of the homeodomain (+8 for *Hox-4.7*; +12 for *Hox-4.6*). Both proteins terminate after a leucine residue (Leu) followed by an aromatic tyrosine (Tyr) or phenylalanine (Phe) residue. The upstream exons are very GC rich and therefore encode a high number of proline (Pro), glycine (Gly) and alanine (Ala) residues (Figure 3A and B). The *Hox-4.6* protein is slightly larger than that of *Hox-4.7* (323 amino acids versus 275) due to the presence in the former of large stretches of poly(Gly) and poly(Ala) (21 Gly, 13 Ala; boxed in Figure 3B). Homologous proteins have been cloned from the newt (*Hox-4.6*, J.Brockes, personal communication) and from the chicken (*Hox-4.7*, S.Mackem and K.Mahon, personal communication; *Hox-4.7* and *-4.6*, our unpublished data). In the case of *Hox-4.7*, two donor splice sites are present. The use of the first site (first black triangle in Figure 3A) was illustrated by the isolation of a chicken cDNA clone

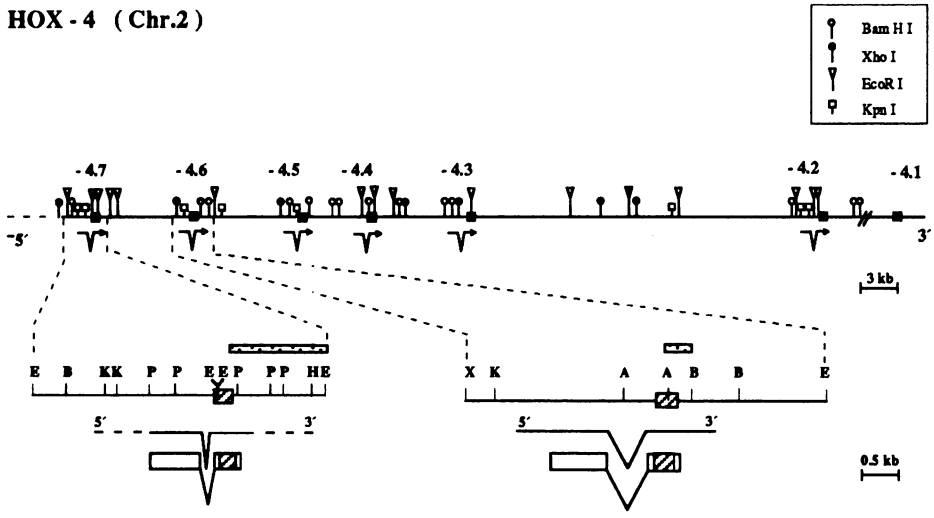

Fig. 1. Molecular organization of the murine HOX-4 complex on mouse chromosome 2 (Featherstone *et al.*, 1988). About 70 kb of DNA are shown containing six genes (from *Hox-4.2* to *-4.7*). The *Hox-4.1* gene is located further downstream. These genes are transcribed from left to right (5′−3′) and some of the restriction sites are indicated. The bottom lines are enlargements of genomic subregions containing the *Hox-4.7* and *-4.6* genes, from left to right, respectively. The RNA (solid lines) and protein coding (open rectangles) sequences are indicated below with the positions of intronic regions. The stippled boxes are the homeoboxes (homeodomains) whereas the rectangles above indicate the clones used for *in situ* hybridizations (see Materials and methods). E, *Eco*RI; B, *Bam*HI; K, *Kpn*I; P, *Pst*I; X, *Xho*I; A, *Acc*I.

(S.Mackem and K.Mahon, personal communication). The second site, also occurring at a glycine residue (second black triangle in Figure 3A), if used, would allow an extended homology with the *Drosophila AbdB* proteins (see later and Figure 5).

The *-4.6* and *-4.7* homeodomain protein sequences are rather divergent from that of the *Antp* homeodomain classically used as a reference (Scott *et al.*, 1989) since only 52% and 45% of the amino acids are identical, respectively. This is expected when one considers the fact that increasing divergence is observed with respect to a central (*Antp*-like) homeobox, when compared with homeobox sequences located at the extremities of the complexes (see e.g. Hart *et al.*, 1987). This is particularly apparent in the case of the HOX-4 complex where a comparison with the *Antp* homeodomain (HD) revealed 83, 68, 63, 52 and 45% identity for the *Hox-4.3*, *-4.4*, *-4.5*, *-4.6* and *-4.7* HD, respectively. Interestingly, the *Hox-4.7* HD contains a phenylalanine (F) residue at position 26 (according to Scott *et al.*, 1989) instead of the tyrosine (Y) residue found so far in all HDs belonging to this class. This is also observed in human (E.Boncinelli, personal communication) and chicken (Izpisúa-Belmonte *et al.*, 1991). A comparison of these HD sequences revealed that they are all highly related to each other. Figure 4A shows an alignment of the four sequences from the previously defined *AbdB*-like subfamily (*Hox-1.7*; *-2.5*; *-3.2*; *-4.4* and *AbdB*) and sequences from

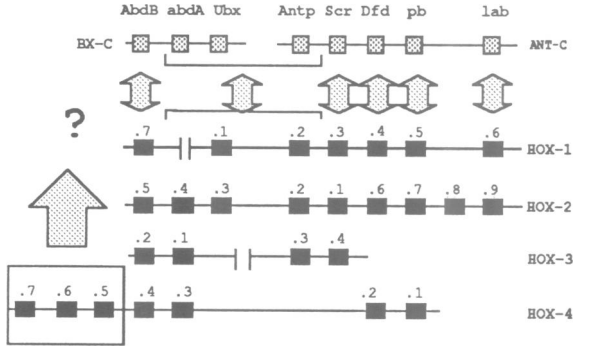

Fig. 2. The alignment of the *Drosophila* homeotic gene complexes *Bithorax* (BX-C) and *Antennapedia* (ANT-C) and the four murine HOX complexes is shown. Arrows phylogenetically link the two gene families. The locations of the murine genes *Hox-4.5*, *-4.6* and *-4.7* are emphasized by the black box. The corresponding large arrow and the question mark illustrate the fact that these genes are lying outside the possible alignment between these families raising the question of their evolutionary origin as well as their possible existence in insects.

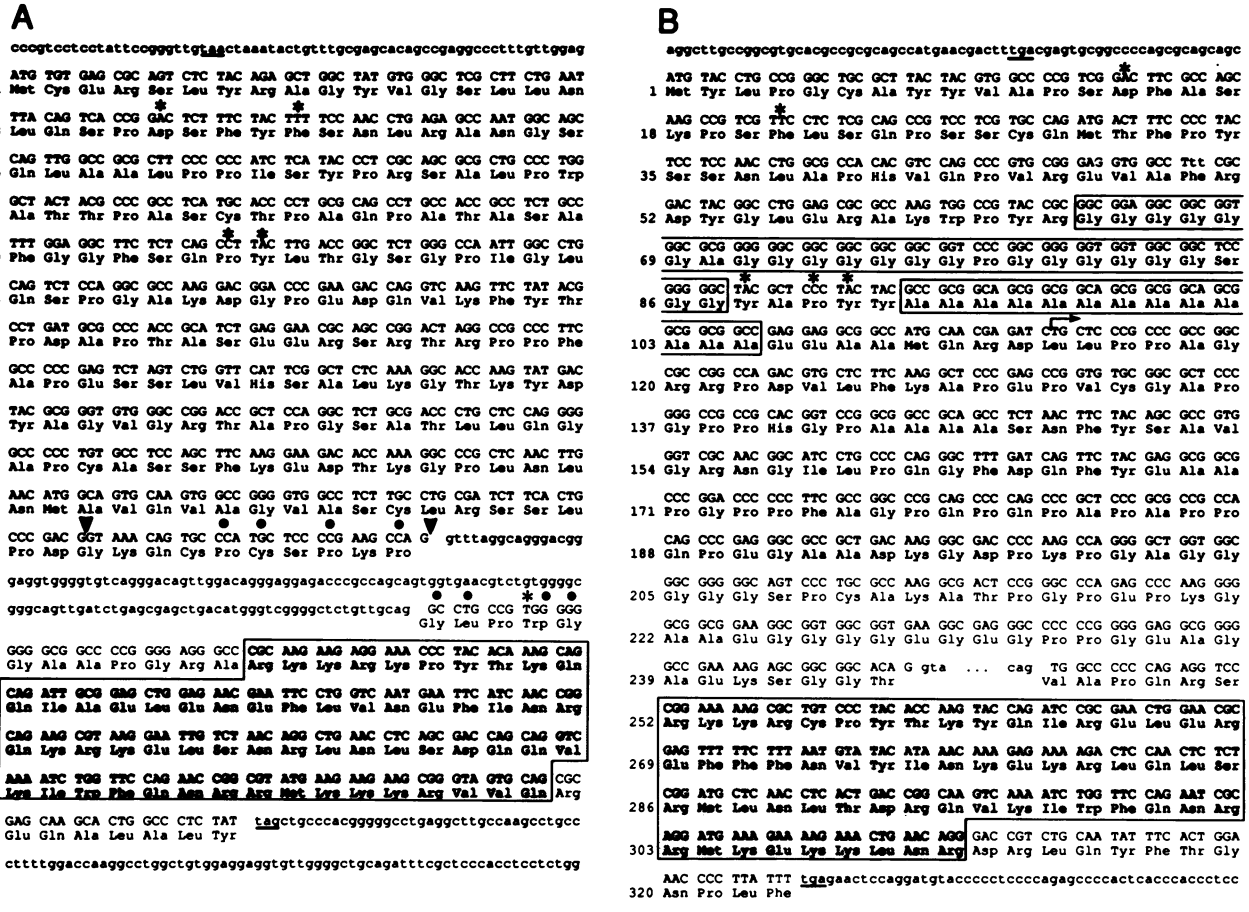

Fig. 3. DNA and protein sequences of the *Hox-4.7* (**A**) and *-4.6* (**B**) coding regions. Partial leader and trailer regions and intronic sequences are shown in small letters. The size of the *Hox-4.6* intron (not shown) is ~750 bp. The homeobox (homeodomain) sequences are boxed and in-frame termination codons are underlined (both in 5′ and 3′). The stars in both sequences indicate consensus amino acids (see Figure 5 and the text) and the black dots in *Hox-4.7* designate an extended homology with the corresponding region of the *Drosophila AbdB* protein sequence, provided the second potential donor splice site is used in the mouse (the second arrowhead). The first arrowhead indicates the donor site present in a chicken *Hox-4.7* cDNA clone (see text). The arrow in *Hox-4.6* marks the 5′ end of our cDNA clone. These sequence data are available from EMBL/GenBank/DDBJ under accession numbers X58848 and X48849.

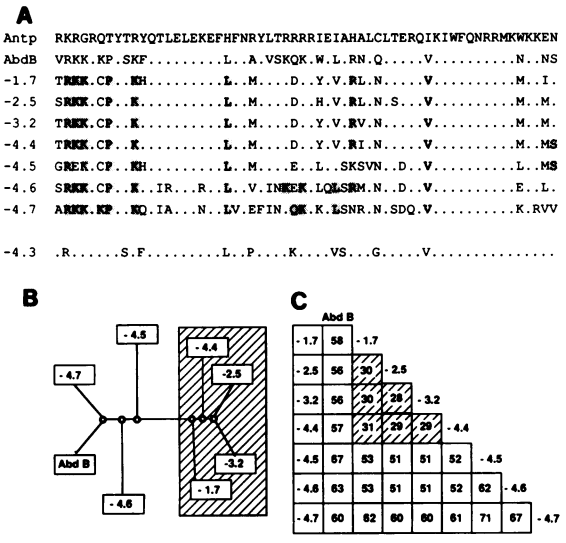

Fig. 4. (**A**) Protein sequence comparisons between the homeodomains of all the murine *AbdB*-like proteins reported to date. The murine and the *AbdB* homeodomain sequences are shown only for residues which differ from those of the *Drosophila Antennapedia* (*Antp*) homeodomain. The non-*AbdB*-like mouse *Hox-4.3* gene (the immediate 3′ neighbouring gene, see Figure 1) is shown as a control. The stippled residues are those found in the murine sequences which differ from the *Antp* reference but are identical to the *AbdB* sequence (identical substitutions). Note that more *Antp*-divergent sequences (like *Hox-4.6* or *-4.7*) also have more identical substitutions with the *AbdB* sequence. Sequences are from: *Antp* (McGinnis *et al.*, 1984), *AbdB* (Regulski *et al.*, 1985), *Hox-1.7* (Rubin *et al.*, 1987), *Hox-2.5* (Bogarad *et al.*, 1989; Graham *et al.*, 1989), *Hox-3.2* (Breier *et al.*, 1990; Erselius *et al.*, 1990), *Hox-4.3* (Izpisúa-Belmonte *et al.*, 1990), *Hox-4.4* and *Hox-4.5* (Duboule and Dollé, 1989), *Hox-4.6* and *-4.7* (this work). (**B**) Unrooted evolutionary tree as obtained with a parsimony algorithm from the program package PHYLIP (Felsenstein, 1985). This algorithm calculates the minimum number of possible intermediate steps required to evolve these various sequences from a common sequence. (**C**) Matrix representation of the protein similarities expressed in percent of mismatches within the homeodomain sequences shown under (A). The stippled areas in (B) and (C) emphasize the short distances (B) or higher homologies (C) between the sequences belonging to the paralogous genes (the four genes from the *Hox-4.4*-like subgroup, see Figure 2).

the three upstream genes *Hox-4.5*, *-4.6* and *-4.7*. The *Hox-4.3* sequence is shown as a control. The N-terminal part of all these HDs (except *Antp* and *Hox-4.3*) is clearly of the '*AbdB* type' (rich in lysine) and many residues present in the *AbdB* homeobox but not in *Antp* are also found in these 'upstream' HDs (boxed in Figure 4A). Interestingly, it is difficult to assess which one of these sequences is the most closely related to that of the *AbdB* gene. For example, the *Hox-4.7* HD does not contain the arginine residue at position 36, typical for the *AbdB* and *AbdB*-like HD (*Hox-1.7*, *-3.2*, *-2.5* and *-4.4*) but instead, contains the glutamine−lysine (QK) doublet which is found at position 29−30 within the *AbdB* sequence. Several such examples can be seen in Figure 4A.

To clarify the evolutionary distances between these novel sequences, we used a parsimony algorithm taking as a control the four paralogous sequences (*Hox-1.7*, *-2.5*, *-3.2* and *-4.4*). The computer clearly recognizes these four latter sequences as originating from a late duplication event (stippled areas in Figure 4B and C) but does not give any indication as to the evolutionary relationships of the other sequences. Moreover, the *AbdB* sequence is not considered significantly closer to this '*AbdB*-like' group but seems equidistant from these HD sequences (Figure 4B). We conclude from this that the 'upstream' located *Hox-4* genes (*Hox-4.4*, *-4.5*, *-4.6* and *-4.7* are equally related to the *Drosophila AbdB* gene. A careful analysis of the complete *Hox-4.6* and *-4.7* protein sequences and their comparison with either the full murine *Hox-4.4* and *-4.5* protein sequences (A.Renucci and D. Duboule, in preparation) or previously published members of the *AbdB*-like family (*Hox-1.7*, Rubin *et al.*, 1987; *Hox-3.2*, Erselius *et al.*, 1990; or the *Xenopus Xlhbox6*, Wright *et al.*, 1990) revealed the presence of consensus sequences shown in Figure 5A. All proteins in Figure 5A (as well as the various *AbdB* proteins, cf. DeLorenzi *et al.*, 1988; Celnicker *et al.*, 1989; Zavortink and Sakonju, 1989) lack the classical hexapeptide consensus (Mavilio *et al.*, 1986) found so far in all other *Hox* proteins. Instead, most of the *AbdB*-like proteins isolated (except the mouse, chicken and newt *Hox-4.6*) have a tryptophan (Trp) residue at position −6 or −7 from the start of the homeobox. A Trp residue is never found at this position in other *Hox* genes. However, it is present in a similar position upstream of the *AbdB* HD as well as in a sea-urchin cognate gene (HB4, Dolecki *et al.*, 1988). The most posterior (5′-located) gene of the HOX-4 complex (*Hox-4.8*) was recently cloned. Its structure and expression pattern confirm the observations reported in this paper (Dollé,P., Izpisúa-Belmonte,J.-C., Boncinelli,E. and DuBoule,D., in press).

When compared with the *AbdB* PS14 protein form ('r') which is thought to achieve a 'regulatory' function in parasegment PS14 (revealed by the class II mutants from Casanova *et al.*, 1987 and encoded by the 3.4 kb transcript P2 in Celnicker *et al.*, 1989; class B in Zavortink and Sakonju, 1989 or β transcripts in Kuziora and McGinnis, 1988), two additional regions of consensus can be seen which are depicted in Figure 5. First, an acidic (Asp) and a cyclic (Phe) residue, separated by two to seven amino acids, are systematically found between positions +15 and +25 in vertebrate proteins and at position +38 in the *AbdB* PS14 ('r') form. Secondly, a highly conserved proline−tyrosine (Pro−Tyr) doublet is found in the central part of the proteins. These consensus sequences (also shown by stars in Figure 3A and B) further suggest the common origin of these genes by successive duplications of an ancestral *AbdB*-like gene.

Expression of the Hox-4.6 and -4.7 genes during development

The extreme 5′ location of these two novel genes in the HOX-4 complex should result, according to the structural colinearity rule, in very posteriorly restricted expression domains. This is indeed the case. At 12.5 days p.c., *Hox-4.6* and *-4.7* are expressed in the structures which usually express *Hox* genes (see e.g. Gaunt *et al.*, 1988; Holland and Hogan, 1988), i.e. within the spinal cord, sclerotome prevertebral condensations, gut mesenchyme, excretory system, etc, but starting at very posterior levels along the rostro-caudal axis. For example, *Hox-4.6* is expressed from prevertebra (pv) 25 whereas *Hox-4.4* starts at pv 20 (Duboule and Dollé, 1989 and Figure 6A). In fact, the limits of expression can vary slightly following the plane of section (more or less paramedian or median as illustrated in Figure 6A versus 6A′). However, for a determined plane of section, colinearity is always observed in the positions of the *Hox-4* gene expression boundaries. This is further exemplified by the *Hox-4.7* gene whose transcripts start to appear from

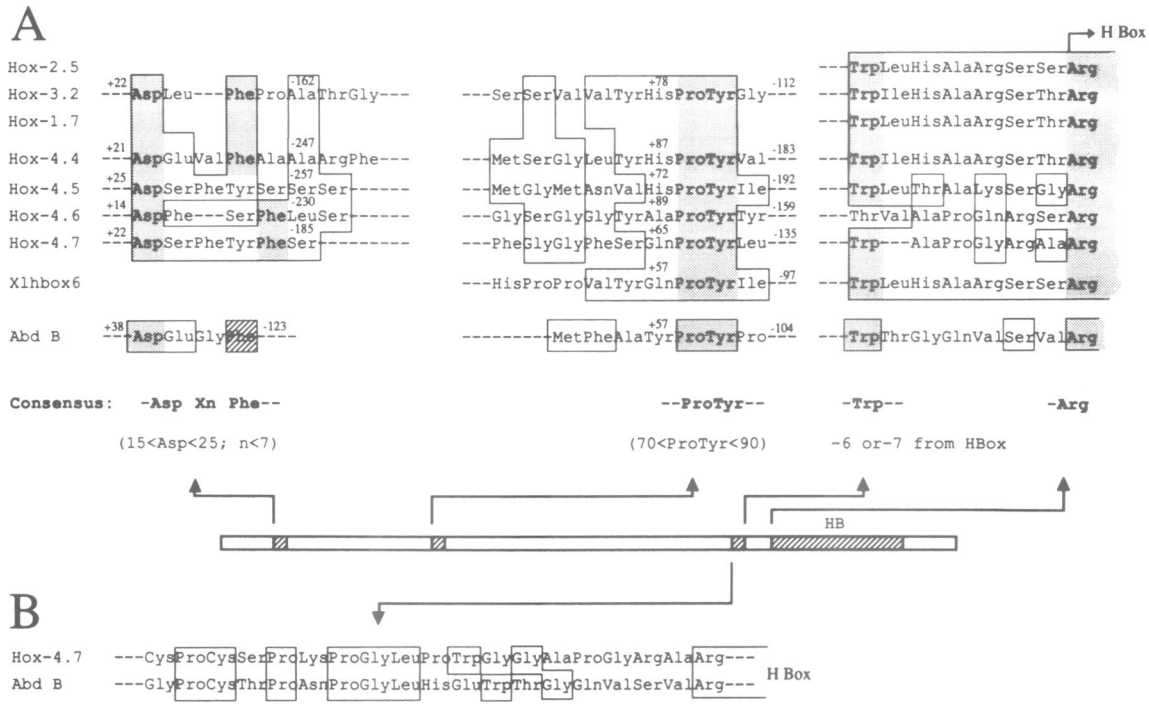

Fig. 5. (**A**) Consensus amino acids between all the *AbdB*-like proteins. Three regions strongly conserved in the various proteins are shown and their positions are depicted by the stippled boxes in the schematic protein shown below. The shadowed residues are those highly conserved whereas boxed amino acids are those present more than once at a fixed position within different proteins. The arrow on the top represents the start of the homeodomain (HB in the scheme below). The strict consensus positions are shown in bold. Positive numbers indicate the positions relative to the initiation codon whereas negative numbers are relative to the start of the HB. This double numbering system is used to illustrate the respective similar positions of these consensus sequences. In the case of *AbdB*, the numbers refer to the protein thought to achieve the regulatory (r) function (see text). References are as in Figure 4 and *Xlhbox6* (Wright *et al.*, 1990). (**B**) Extended amino acid homology between the *Drosophila AbdB* and the *Hox-4.7* sequences, upstream of the HB. This alignment is made possible only if the second consensus splice donor site is used (second arrowhead in Figure 3A).

pv 29−30 and more posteriorly (Figure 6B) which corresponds to the beginning of the tail. The comparative positions of the *Hox-4.4* to *-4.7* expression boundaries in prevertebrae is shown in the bright field panel of Figure 6 (open arrows). This colinearity is also observed in mesenchymal compartments of some internal structures or organs as illustrated by a section through a piece of intestine shown in Figure 6C. The entire histological section is positive for *Hox-4.4* but negative (too anterior) for *Hox-4.7*. This piece of intestine corresponds to the exact position of the *Hox-4.6* expression boundary since only half of it (the posterior part) is positive for the *Hox-4.6* probe (Figure 6C). This distribution of *Hox-4* gene transcripts is already visible at an earlier stage in development, e.g. in the day 9.0 p.c. embryo section shown in Figure 6D, where clear differences can be seen in the positions of the *Hox-4.6* and *-4.7* expression boundaries along the lateral-plate mesoderm (open arrows in Figure 6D). In contrast, *Hox-4.4* is expressed in all the posterior part of the fetus shown in this section (the head is, of course, negative in all three cases).

Sequential appearance of the HOX-4 transcripts

We and others previously reported that the HOX-4 genes were expressed in limbs during development (Dollé and Duboule, 1989; Oliver *et al.*, 1989), following a sequential temporal colinear activation (Dollé *et al.*, 1989). Because of the cranio-caudal progression of morphogenetic events observed along the developing vertebrate AP axis (see Discussion), it was critical to confirm this observation in the case of the trunk axis. A careful time course of *in situ* hybridization shows that only genes anterior to (located 3′ from) *Hox-4.4*, as well as *Hox-4.4* itself, are transcribed up to detectable levels before day 8.25 p.c. Transcripts encoded by the *Hox-4.5* and *-4.6* genes appear between days 8.25 and 8.75 p.c. whereas the *Hox-4.7* gene apparently becomes functional at about day 9 (Figure 7). As expected, the earlier expression of these genes is restricted to very posterior areas (see e.g. the posterior lateral-plate mesoderm in Figure 7A′ or the posterior mesoderm in Figure 7B). Several embryos at different ages were analysed in systematic serial sections giving comparable results. No further attempt was made to try to separate temporally the appearance of the *Hox-4.5* from the *Hox-4.6* transcripts.

Discussion

The two major observations reported in this paper are that vertebrates, or at least mammals, have preferentially amplified an ancestral gene member of the *AbdB*-like family and that the genes produced by such gene duplication events, in the HOX-4 complex, are activated according to a structural and temporal sequence during the development of the trunk and limb axes.

Multiple genes versus multiple regulatory units

The molecular analysis of the *Hox-4.4*, *-4.5*, *-4.6* and *-4.7* genes allowed us to compare their protein coding sequences. Analysis of these sequences revealed the presence of consensus motifs. When compared with the *Drosophila AbdB* gene, the consensus sequences were also conserved although

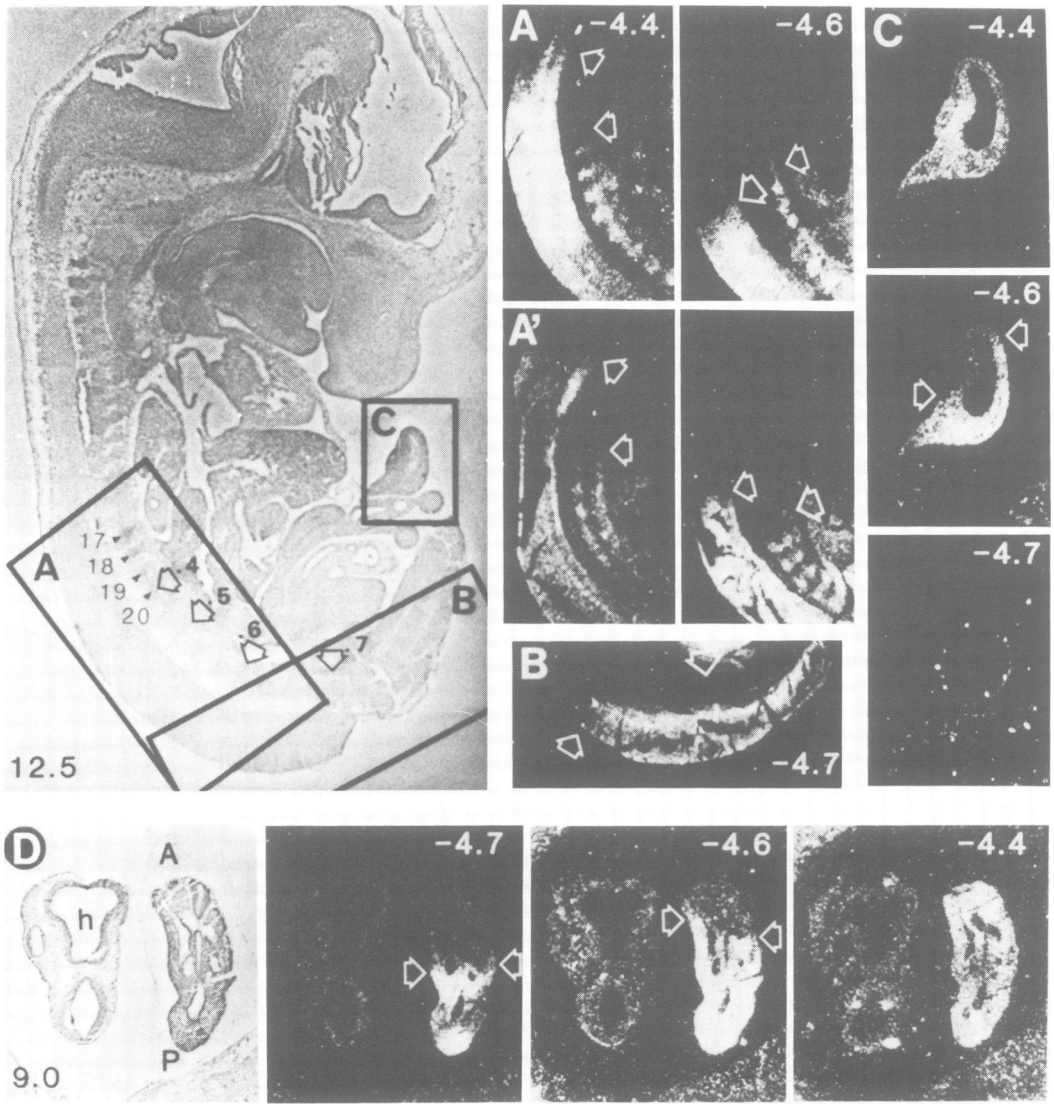

Fig. 6. Restricted expression of the *Hox-4.6* and *-4.7* genes along the antero-posterior (AP) axis of the developing mouse fetus. A, A′, B and C show the various boundaries of expression, in a 12.5-day-old mouse fetus, along the spinal cord and two mesoderm derivatives that have an AP continuity: the prevertebral column and the gut mesenchyme. The *in situ* hybridization signals seen with the *Hox-4.4*, *-4.6* and *-4.7* probes are viewed under dark field illumination and the magnified areas are boxed in the adjacent bright field picture of a fetus. (**A**) and (**A′**) show the boundaries of expression of *Hox-4.4* and *-4.6* in the spinal cord and prevertebrae (pv). The section in (A) is slightly paramedian whereas (A′) is a parallel median section (crossing the spinal cord lumen). The boundaries are more anterior for *Hox-4.4*. Along the pv column, the *-4.6* expression extends more anteriorly in lateral sections of the sclerotome (pv 23 in A) than in mid-sagittal sections (pv 25 in A′). (**B**) Position of the *Hox-4.7* expression boundaries (pv 29−30). Panel (**C**) shows the differential labelling within the mesenchyme of the herniated gut. The open arrows (black in A and white elsewhere) indicate the expression boundaries of *Hox-4.4*, *-4.5*, *-4.6* and *-4.7* (pv 17−20 are shown as reference). (**D**) Distinct AP domains of expression of the *Hox-4.4*, *-4.6* and *-4.7* genes in a 9-day-old embryo. h, head.

absent from all other murine *Hox* genes studied so far. In addition, none of these proteins contain the conserved hexapeptide (Mavilio *et al.*, 1986) which is found upstream of the homeobox of all class I homeobox genes reported to date. This peptide sequence is not found in any *AbdB* protein (see also Erselius *et al.*, 1990). Instead we found at about equivalent distance, 5′ from the homeobox, a conserved tryptophan residue in all *AbdB*-like proteins with the exception of *Hox-4.6*. Interestingly, the 'core' of this conserved hexapeptide is composed of the doublet Trp−Met, present in all *Hox* proteins whereas other positions can show some variation (Duboule *et al.*, 1989). This amino acid is also found in an early coelomate metazoan *AbdB*-related gene (position −7; Dolecki *et al.*, 1988).

Computer analysis using a parsimony algorithm confirmed the difficulty in linking preferentially any of these *AbdB*-like genes to the *Drosophila* gene with respect to evolution. From these data it therefore appears that all of these genes are likely to be derived from the same *AbdB*-like ancestral gene. Apart from the HDs and the short consensus sequences no extended similarities with the *AbdB* gene can be found except for the presence in the *Hox-4.6* of monotonic stretches of amino acids. In particular, a very large stretch of alanine residues (Ala) is observed which is also found in the *AbdB* proteins (see Celnicker *et al.*, 1989; Zavortink and Sakonju, 1989).

Akam *et al.* (1988) have proposed that an ancestral *AbdB*-like gene could have determined the most posterior

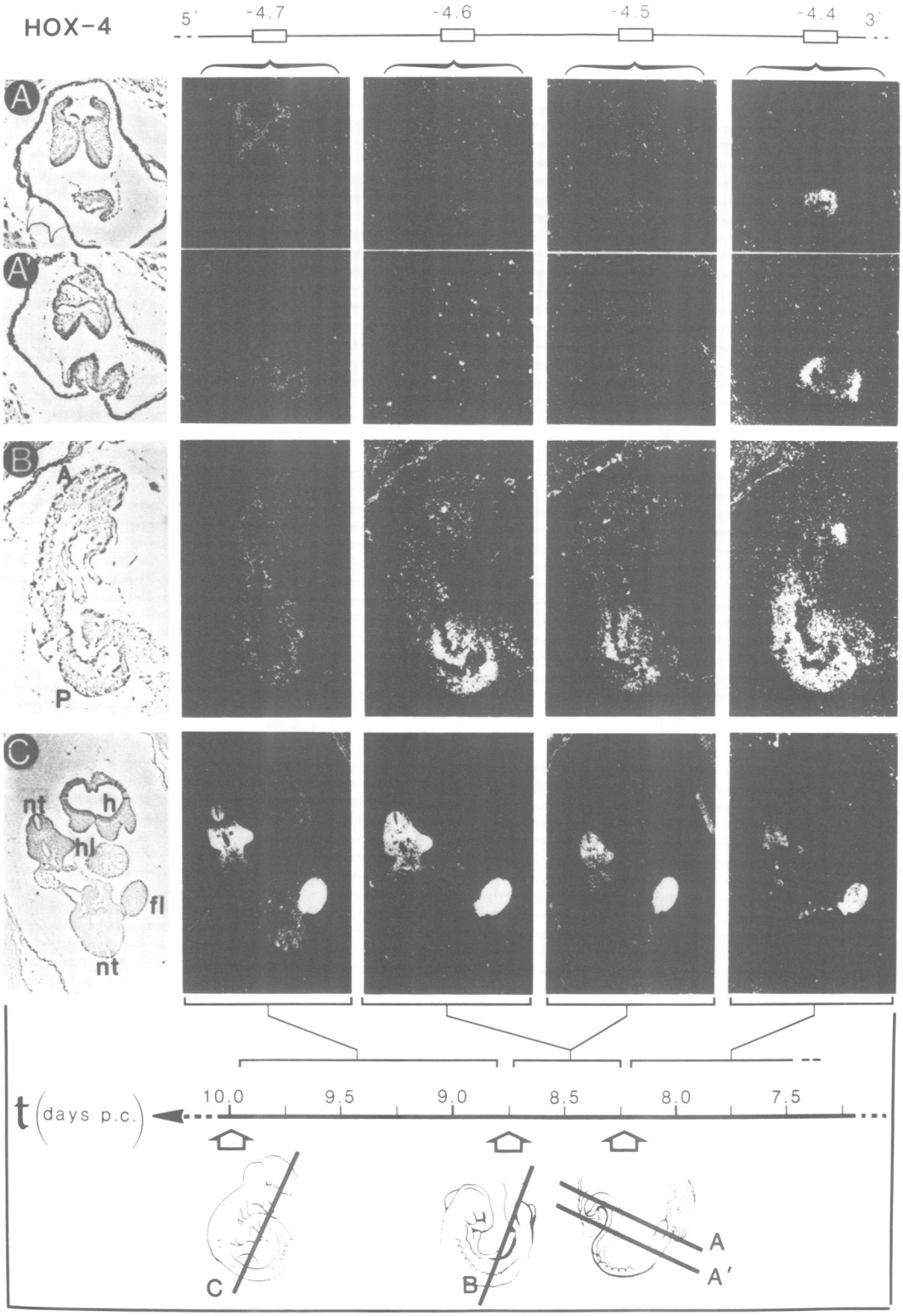

Fig. 7. Sequential appearance of the *Hox-4* transcripts during mouse development. (**A**) and (**A'**) Two parallel sections of a 8.25-day-old embryo (before rotation). The plans of sectioning are shown in the corresponding drawing at the bottom of the figure, below the time axis. In this embryo, only *Hox-4.4* transcripts are detected in lateral mesoderm (A') as well as more posterior regions (A). *Hox-4.5* to *-4.7* transcripts are not visible at this developmental stage, even in the most posterior areas (A). (**B**) Section through an 8.75-day-old fetus (after rotation) crossing the posterior extremity and the allantois. At this stage, *Hox-4.4*, *-4.5* and *-4.6* transcripts are readily detected in posterior regions while *Hox-4.7* is negative. Note that only *Hox-4.4* is transcribed in cells of the forelimb field. (**C**) Section through a 10.0-day-old fetus showing the high expression of the four genes in posterior parts of the body and in the forelimb bud. A, anterior; P, posterior; h, head; hl, hindlimb; fl, forelimb; nt, neural tube.

part of the abdomen in the ancestor of the myriapod−insect lineage. In *Drosophila*, the *AbdB* gene has multiple mRNAs transcribed from different promoters (DeLorenzi *et al.*, 1988; Kuziora and McGinnis, 1988; Celnicker *et al.*, 1989; Zavortink and Sakonju, 1989). The *AbdB* regulatory regions are separated and genetically defined by a set of recessive

loss of function mutations which transform a given parasegment into a copy of the more anterior parasegment. These classes of mutations are called *iab-5, iab-6, iab-7* and *iab-8,9* (Lewis, 1978; Karch *et al.*, 1985; Sanchez-Herrero, 1986; Tiong *et al.*, 1986; and Duncan, 1987 for a review). It has been shown genetically that these mutations affect regions of the *Bithorax* complex (BX-C) arranged in a way which is colinear to the ordering of the adult structures they determine in the *Drosophila* embryo (Lewis, 1978). It was further proposed that this would reflect the colinear arrangement, along the chromosome, of *cis*-regulatory domains conferring various parasegmental specificities to a few transcription units (Peifer *et al.*, 1987). Thus, the mouse and the fly could represent two pathways of evolving genetic complexity from a rather simple system (Figure 8). In *Drosophila*, Akam *et al.* (1988) proposed that the evolution of the myriapod−hexapod lineage involved both the acquisition of new homeotic genes and, in the case of *AbdB*, the elaboration of additional regulatory sequences in parallel with the novel individualization of posterior parasegments. Alternatively, Bender, Karch and colleagues suggest that the ancestral animal found at the origin of this lineage had a number of homeotic genes roughly corresponding to its number of segments (discussed in Karch *et al.*, 1990). In this view, the various *AbdB cis*-regulatory sequences (enhancers) would have been conserved and 'concentrated' for the regulation of a common transcription unit. This phylogenetic conversion from a gene to an enhancer sequence could occur, for example, by deletion of transcribed regions after convergence of multiple regulatory sequences to a restricted number of transcription units. Such a process of 'consolidation' (see Karch *et al.*, 1990) can be envisaged only if the original functions of each reiterated unit were rather comparable, in other words, if the ancestral abdominal segments were poorly functionally individualized as was likely the case in a putative annelid-like ancestor. The situation we observed in mammals can be considered as another form of evolution, by complete gene duplications rather than by increasing the regulatory capacity of one transcription unit. The poor overall similarity between these 5′-located *Hox-4* proteins and their insect cognate (when compared e.g. with the *Hox-1.4/Deformed* gene subfamily; Graham *et al.*, 1988; Galliot *et al.*, 1989) may thus illustrate this particular way to evolve diversity in parallel with the amplification of this subgroup (Figure 2). Alternatively, this particular multi-*AbdB* arrangement in mammals could reflect the structure of a primitive HOX/HOM complex. In this case, consolidation could have 'fixed' the *Drosophila* sequence (by evolving the regulatory units) and therefore be responsible for the high divergence observed between *Drosophila* and vertebrates (see Karch *et al.*, 1990).

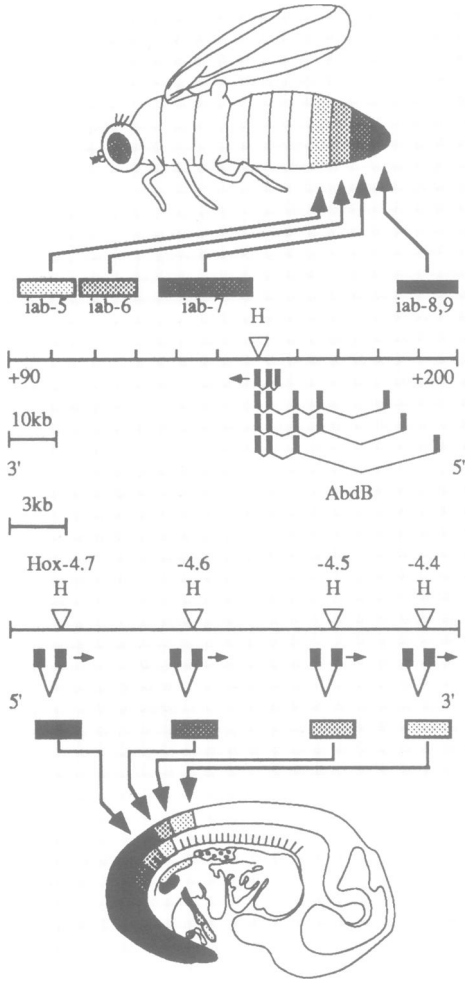

Fig. 8. Schematic representation of the functional homologies between the *AbdB* gene, its set of regulatory sequences (*iab-5* to *iab-8,9*) and the various murine *AbdB*-like *Hox-4* genes. The top half shows the *Drosophila* transcription unit containing several regulatory regions which control, in a parasegmental-specific manner, the expression (from right to left) of at least four different transcripts in the most posterior abdominal segments. The different intensities of grey reflect in which abdominal parasegment (PS) a given *iab* regulatory sequence seems to exert its maximal effect (e.g. *iab-5* in PS10, etc.). The scheme was redrawn from Peifer *et al.* (1987) with additional data from Celnicker *et al.* (1989), Kuziora and McGinnis (1988) and Zavortink and Sakonju (1989). The bottom half illustrates the arrangement of the *AbdB*-like *Hox-4* genes in mammals, with their respective transcription units (from left to right) as well as the positions of their expression boundaries in the CNS, prevertebrae and internal organs. All expression domains extend very posteriorly (they overlap in posterior areas). The position of homeobox sequences is shown (H).

Two hypotheses can be envisaged to describe the evolution of the murine HOX complexes with respect to these multiple *AbdB*-like genes which have been found so far in the HOX-4 (Duboule and Dollé, 1989; this work) and HOX-1 (P.Gruss, personal communication) complexes. In humans, such *AbdB*-like genes are present in the HOX-1, -3 and -4 complexes but, as in the mouse, are probably absent from the HOX-2 complex (Acampora *et al.*, 1989; Simeone *et al.*, 1990a). Consequently, either these multiple copies existed in the ancestral complex and were therefore lost in HOX-2, or they were the result of *AbdB*-like gene duplications after the large-scale duplication of two HOX-2-like complexes. In this last case, the HOX-2 complex would therefore be the representation of the ancestral mammalian complex. Our results do not provide clear-cut evidence to support either possibility but the loss of *Hox* genes from various complexes and at various positions (or their evolution towards a highly divergent unit) is well documented in human and mouse where only few subfamilies contain the four members (Simeone *et al.*, 1990a; Kessel and Gruss, 1990). For example, it appears that the HOX-1 complex contains only some of the *AbdB*-related genes whereas the HOX-4 complex does have the five *AbdB*-like genes (*Hox-4.4*, *-4.5*, *-4.6*, *-4.7* and *-4.8*). It is intriguing, though probably coincidental, that in *Drosophila* the *AbdB* functions

are required in five parasegments, from PS10 to PS14 (see also Acampora *et al.*, 1989).

Temporal colinearity

The *Hox-4.6* and *-4.7* genes are expressed during fetal development in a very posteriorly restricted manner. This is not surprising and fits well with the strict 'structural colinearity' observed within the murine HOX complexes (Gaunt *et al.*, 1988; Duboule and Dollé, 1989; Graham *et al.*, 1989). Indeed, they are expressed in the CNS, sclerotomes and in a few internal organs, slightly more posteriorly than the *Hox-4.4* and *-4.5* genes (Dollé and Duboule, 1989). We have previously shown that four of these genes (*Hox-4.4*–*4.7*) are expressed during the morphogenesis of the limbs according to rules similar to those governing their expression in the trunk. Furthermore, these genes are sequentially activated in the limbs (or their transcripts sequentially stabilized) and we proposed that this was an essential process to establish correctly their expression domains (Dollé *et al.*, 1989). We now show that sequential activation of these *AbdB*-like genes also occurs during development along the AP body axis. The first transcripts to appear are those encoded by the genes located 3′, and the last ones those encoded by the upstream (5′)-located *AbdB*-like genes (e.g. *Hox-4.7*). We refer to this process as 'temporal colinearity'. Temporal colinearity may be difficult to observe among the 3′-located (non-*AbdB*-like) genes since many of these genes seem to be activated during a very short period of time. We believe that this is due to a non-linear timing of activation of the *Hox* genes, many 'anterior' genes being activated almost simultaneously in the case where the structures which are formed are very poorly time-delayed (e.g. the hindbrain neuromeres, reviewed in Lumsden, 1990; Wilkinson and Krumlauf, 1990). This 3′–5′ sequential appearance of the *Hox* gene transcripts is an important observation since it combines the temporal molecular dynamics of a genetic system on the one hand, with the cranio-caudal morphogenetic progression observed during vertebrate ontogeny, on the other. This means that we may not only have, in our chromosomes, a direct molecular representation of our body axes (the 'structural colinearity') but also, linked to it, a linear molecular representation of the temporal sequence which is required to correctly realize this structural colinearity (the 'temporal colinearity').

Such a cranio-caudal (AP) morphogenetic progression is not apparent in *Drosophila*, where the homeotic genes do not seem to be transcribed according to a colinear temporal sequence. However, the very high developmental speed of *Drosophila* as well as the very variable sizes of the homeotic transcription units makes it difficult to address this point (see e.g. Gubb, 1986; Kornfeld *et al.*, 1989; discussed in Karch *et al.*, 1990). In fact, according to the modification proposed by Bender and colleagues (Peifer *et al.*, 1987) of the combinatorial model of Lewis (1978), a corresponding mechanism acting on the *Drosophila AbdB* transcription unit would involve a progressive temporal accessibility ('opening') of *cis*-acting sequences (from *iab-5* to *iab-8,9*; see Figure 8) to regulatory factors. Since there is no known correlation between such *cis*-acting sequences and particular *AbdB* transcripts, there is no evidence so far which would support this idea in the fruit-flies. However, such a mechanism should exist in more primitive short germ band insects such as the locust *Schistocerca gregaria* since they determine and produce their most posterior abdominal segments after the blastoderm stage, in an AP sequential manner (for references and discussion, see Akam *et al.*, 1988; Tear *et al.*, 1988; French, 1990). Among the arthropods, the Onychophoran as well as myriapods such as the Pauropodae and crustaceans also show a clear sub-terminal addition of the posterior segments, or, at least, an AP progression in the development of the segments (see Raff and Kaufman, 1983). In fact, this AP dynamics in the establishment of meristic features is observed in many metamerized invertebrates (see e.g. Dawydoff, 1928) including annelids (oligochaete and polychaete) which suggests that the ancestral organism present at the origin of these various lineages had a similar type of ontogeny (discussed for the arthropods by Raff and Kaufman, 1983). Consequently, long germ band insects could represent a rather peculiar, highly evolved, group of arthropods. The apparent non-involvement of temporal colinearity in these animals might be illustrated (reflected), at the molecular level, by the splitting of the homeotic complexes and the presence of an ANT-C homeotic gene (*Dfd*) on the opposite DNA strand, a situation which contradicts that found in the vertebrate HOX complexes (reviewed in Kaufman *et al.*, 1990). In this context, the analysis of less evolved arthropods could reveal whether in such an evolutionary pathway, a complete heterochrony (the disruption—or tolerance to disruption—of a temporal programme of development) can be envisaged.

As far as the *AbdB* functions are concerned, one would expect, in short germ band insects, other arthropods or annelids, that the equivalent of the *iab-5* function may be active before that of the *iab-6*, etc, in a way somehow homologous to what we observe during the vertebrate trunk and limb axes development. However, this possible temporal colinearity would occur clearly only for 'abdominal' metamers and the question would thus remain open with respect to more 'anterior' homeotics. This possible difference, in regulatory mechanisms, between *AbdB* (*AbdB*-like) and the other homeotic (homeo) genes, both in insects and vertebrates, is reinforced in vertebrates by several other structural and functional observations (see above). Interestingly, the first *AbdB*-like gene found when moving towards the 5′ end of the vertebrate complexes (the *Hox-4.4*-like subgroup) seems also to represent a functional 'boundary' in cells cultured *in vitro*, since genes located either upstream or downstream have very different behaviour in response to retinoid treatment (Simeone *et al.*, 1990a,b). However, the fact that the 3′-located genes *Hox-1.6* and *Hox-2.9* are expressed very early and during a short period of time (Duboule and Dollé, 1989; Wilkinson *et al.*, 1989; Frohman *et al.*, 1990; Sundin *et al.*, 1990; Wilkinson and Krumlauf, 1990) is strongly suggestive that temporal colinearity, like structural colinearity, is a concept applicable to the entire HOX network. The analysis of the HOM/HOX complexes in other species of either arthropods or vertebrates should reveal whether the temporal colinearity is intimately linked to the presence of an AP morphogenetic progression. The first example of such a juxtaposition of an observable, morphogenetic progression with a temporal sequence of molecular events was previously shown for the same genes during the development of the limbs (Dollé *et al.*, 1989). The vertebrate limb progressively develops along its

proximo-distal axis so that proximal patterns are formed before distal ones. We proposed that such a sequential activation could provide an explanation of the mechanisms by which structures derived from proximal blastema are different from those produced by distal blastema. Because of the now apparent uniformity in the patterning mechanisms involved along the different body axes, we suggest in this paper that temporal colinearity could also be the basis for differential specification of the vertebrate metamers according to their AP positions.

Materials and methods

Isolation and mapping of the mouse Hox-4.6 and Hox-4.7 genes

These two novel mouse genes were isolated during our walk on the 5′ part of the HOX-4 complex (Featherstone *et al.*, 1988; Duboule and Dollé, 1989; Izpisúa-Belmonte *et al.*, 1990). The homeobox sequences were revealed by cross hybridization with probes derived from the *Hox-4.4* and *-4.5* genes. The positions and orientations of these genes on the *HOX-4* complex were determined by classical mapping using end-labelled genomic DNA fragments and partial restriction enzyme digests. Various DNA fragments were subcloned into pUC, pPolyIII, pEMBL, pGEM or pBluescript vectors.

Sequencing of the cDNA and genomic DNA clones

The cDNA inserts were isolated from the positive phages and subcloned into pUC18 or pGEM7. Their nucleotide sequences and those of various genomic fragments were determined using the dideoxy method (Sanger *et al.*, 1977). Difficult regions (high GC content) were sequenced with the chemical method (Maxam and Gilbert, 1980).

Screening of cDNA libraries

We used both oligo(dT) and randomly primed cDNA libraries from 11 day p.c. mouse embryonic mRNA (Galliot *et al.*, 1989). In both cases, $\sim 1.7 \times 10^6$ phages were screened under high-stringency conditions (50% formamide, 5 × SSC, 1% SDS, 50 mM Tris−HCl, pH 7.5, 0.1 mg/ml denatured salmon sperm DNA, 42°C) by hybridization with three genomic probes: the 5′−3′ *Eco*RI−*Pst*I homeobox-containing fragment and the 5′−3′ *Pst*I−*Eco*RI immediately upstream fragment, from the *Hox-4.7* gene and the *Acc*I−*Bam*HI homeobox-containing fragment corresponding to the *Hox-4.6* gene (see Figure 1).

Preparation of RNA probes

The *Hox-4.6* and *-4.7* antisense RNA probes were synthesized using a T7 *in vitro* transcription reaction as described previously (Melton *et al.*, 1986). The DNA templates were a 300 bp *Acc*I−*Bam*HI fragment containing part of the *Hox-4.6* homeobox and a 1400 bp *Eco*RI fragment containing part of the *Hox-4.7* homeobox (see Figure 1). Both fragments were subcloned into pGEM. The control sense RNA probes were simultaneously synthesized using the opposite strands as templates. None of these control probes displayed any detectable signals. The *Hox-4.4* and *-4.5* probes were as described by Duboule and Dollé (1989).

In situ hybridization

Mouse embryos and fetuses were obtained from natural matings between (C57/BL6 × SJL) F1 mice. Mid-day on the day of the vaginal plug was designated as day 0.5 p.c. Embryo recovery, embedding, sectioning and *in situ* hybridization were performed as previously described (Dollé and Duboule, 1989), except that the pre-hybridization step was omitted.

Acknowledgements

We would like to thank Drs A.Baron for her help in the initial part of this work, B.Galliot for the embryonic cDNA libraries and M.Vingron for the computer analysis. We thank Dr R.Krumlauf for communicating data prior to publication, and Drs M.Akam, W.Bender, F.Karch, G.Morata and C.Murphy for their comments, corrections and suggestions on the manuscript. We also thank H.Davies and H.Fry for preparing the manuscript. A.R. is the recipient of a postdoctoral fellowship from EMBO.

References

Acampora,D., D'Esposito,M., Faiella,A., Pannese,M., Migliaccio,E., Morelli,F., Stornaiulo,A., Nigro,V., Simeone,A. and Boncinelli,E. (1989) *Nucleic Acids Res.*, **17**, 10385−10427.

Akam,M. (1989) *Cell*, **57**, 347−349.

Akam,M., Dawson,I. and Tear,G. (1988) *Development*, **104** (suppl.), 123−133.

Balling,R., Mutter,G., Gruss,P. and Kessel,M. (1989) *Cell*, **58**, 337−347.

Baron,A., Featherstone,M.S., Hill,R.E., Hall,A., Galliot,B. and Duboule,D. (1987) *EMBO J.*, **6**, 2977−2986.

Beeman,R.W. (1987) *Nature*, **327**, 297−299.

Bogarad,L.D., Utset,M.F., Awgulewitsch,A., Miki,T., Hart,C.P. and Ruddle,F.H. (1989) *Dev. Biol.*, **133**, 537−549.

Breier,G., Dressler,G.R. and Gruss,P. (1988) *EMBO J.*, **7**, 1329−1336.

Casanova,J., Sanchez-Herrero,E., Busturia,A. and Morata,G. (1987) *EMBO J.*, **6**, 3103−3109.

Celniker,S.E., Keelan,D.J. and Lewis,E.B. (1989) *Genes Dev.*, **3**, 1424−1436.

Dawydoff,C. (1928) *Embryologit des Invertébrés*. Masson, Paris.

DeLorenzi,M., Ali,N., Saarai,G., Henry,M., Wilcox,M. and Bienz,M. (1988) *EMBO J.*, **7**, 3223−3231.

Desplan,C., Theis,J. and O'Farrell,P.H. (1988) *Cell*, **54**, 1081−1090.

Dolecki,G.J., Wang,G. and Humphreys,T. (1988) *Nucleic Acids Res.*, **16**, 11543.

Dollé,P. and Duboule,D. (1989) *EMBO J.*, **8**, 1507−1515.

Dollé,P., Izpisúa-Belmonte,J.-C., Falkenstein,H., Renucci,A. and Duboule,D. (1989) *Nature*, **342**, 767−772.

Dollé,P., Izpisúa-Belmonte,J.-C., Boncinelli,E. and Duboule,D. (1991) in press.

Duboule,D. and Dollé,P. (1989) *EMBO J.*, **8**, 1497−1505.

Duboule,D., Boncinelli,E., DeRobertis,E., Feathertone,M., Lonai,P., Oliver,G. and Ruddle,F.H. (1990) *Genomics*, **7**, 458−459.

Duncan,I. (1987) *Annu. Rev. Genet.*, **21**, 285−319.

Erselius,J.R., Goulding,M.D. and Gruss,P. (1990) *Development*, **110**, 629−642.

Featherstone,M.S., Baron,A., Gaunt,S.J., Mattei,M.G. and Duboule,D. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 4760−4764.

Felsenstein,J. (1985) *Evolution*, **39**, 783−791.

French,V. (1990) *Sem. Dev. Biol.*, **1**, 89−100.

Frohman,M.A., Boyle,M. and Martin,G. (1990) *Development*, **110**, 589−607.

Galliot,B., Dollé,P., Vigneron,M., Featherstone,M.S., Baron,A. and Duboule,D. (1989) *Development*, **107**, 343−359.

Gaunt,S.J., Sharpe,P.T. and Duboule,D. (1988) *Development*, **104**, (suppl.), 71−82.

Graham,A., Papalopulu,N., Lorimer,J., McVey,J.-H., Tuddenham,E.G.J. and Krumlauf,R. (1988) *Genes Dev.*, **2**, 1424−1438.

Graham,A., Papalopulu,N. and Krumlauf,R. (1989) *Cell*, **57**, 367−378.

Gubb,D. (1986) *Dev. Genet.*, **7**, 119−131.

Harding,K., Weddeen,C., McGinnis,W. and Levine,M. (1985) *Science*, **229**, 1236−1242.

Hart,C.P., Fainsod,A. and Ruddle,F.H. (1987) *Genomics*, **1**, 182−195.

Hoey,T. and Levine,M. (1988) *Nature*, **332**, 858−861.

Holland,P.W.H. and Hogan,B.L.M. (1988) *Genes Dev.*, **2**, 773−782.

Izpisúa-Belmonte,J.-C., Dollé,P., Renucci,A., Zappavigna,V., Falkenstein, H. and Duboule,D. (1990) *Development*, **110**, 733−746.

Izpisúa-Belmonte,J.-C., Tickle,C., Dollé,P., Wolpert,L. and Duboule,D. (1991) *Nature*, **350**, 585−589.

Kappen,C., Schughart,K. and Ruddle,F.H. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 5459−5463.

Karch,F., Weiffenbach,B., Pfeifer,M., Bender,W., Duncan,S., Celniker,M., Crosby,M. and Lewis,E.B. (1985) *Cell*, **43**, 81−96.

Karch,F., Bender,W. and Weiffenbach,B. (1990) *Genes Dev.*, **4**, 1573−1587.

Kaufman,T.C., Lewis,R. and Wakimoto,B.T. (1980) *Genetics*, **94**, 115−133.

Kaufman,T.C., Seeger,M.A. and Olsen,G. (1990) In Scandalios,J.G. and Wright,T.R.F. (eds), *Advances in Genetics: Genetic Regulatory Hierarchies in Development*. Academic Press, Vol. 27, pp. 309−362.

Kessel,M. and Gruss,P. (1990) *Science*, **249**, 374−379.

Kessel,M., Balling,R. and Gruss,P. (1990) *Cell*, **61**, 301−308.

Kornfeld,K., Saint,R.B., Beachy,P.A., Harte,P.J., Peattie,D.A. and Hogness,D.A. (1989) *Genes Dev.*, **3**, 243−258.

Kuziora,M.A. and McGinnis,W. (1988) *EMBO J.*, **7**, 3233−3244.

Lewis,E.B. (1978) *Nature*, **276**, 565−570.

Lumsden,A. (1990) *Trends Neurosci.*, **13**, 329.

Mavilio,F., Simeone,A., Giampaolo,A., Faiella,A., Zappavigna,V., Acampora,D., Poiana,G., Ruso,G., Peschle,C. and Boncinelli,E.(1986) *Nature*, **324**, 664−666.

Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499−560.

McGinnis,W., Levine,M.S., Hafen,E., Kuroiwa,A. and Gehring,W.J. (1984) *Nature*, **308**, 428−433.

Melton,D., Krieg,P., Rebagliati,M., Maniatis,T., Zinn,R. and Green,M.R. (1984) *Nucleic Acids Res.*, **12**, 7035.
Mlodzik,M., Fjose,A. and Gehring,W.J. (1988) *EMBO J.*, **7**, 2569–2578.
Murphy,P., Davidson,D. and Hill,R.E. (1989) *Nature*, **341**, 156–159.
Oliver,G., Sidell,N., Fiske,W., Heinjmann,C., Mohandas,T., Sparkes,R.S. and DeRobertis,E.M. (1989) *Genes Dev.*, **3**, 641–650.
Otting,G., Qian,Y.Q., Müller,M., Affolter,M., Gehring,W.J. and Wüthrich,K. (1988) *EMBO J.*, **7**, 4305–4309.
Peifer,M., Karch,F. and Bender,W. (1987) *Genes Dev.*, **1**, 891–898.
Raff,R.A. and Kaufman,T.C. (1983) *Embryos, Genes, and Evolution.* Macmillan Publishing Co., Inc., New York.
Regulski,M.H.K., Kostriken,R., Karch,F., Levine,N. and McGinnis,W. (1985) *Cell*, **43**, 71–80.
Rubin,M.R., King,W., Toth,L.E., Sawczuk,I.S., Levine,M.S., D'Eustachio,P. and Nguyen-Huu,M.C. (1987) *Mol. Cell. Biol.*, **7**, 3836–3841.
Sanchez-Herrero,E., Vernos,I., Marco,R. and Morata,G. (1985) *Nature*, **313**, 108–113.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.
Scott,M.P., Tamkun,J.W. and Harzel,I.G.W. (1989) *Biochim. Biophys. Acta*, **989**, 25–48.
Simeone,A., Acampora,D., Arcioni,L., Andrews,P.W., Boncinelli,E. and Mavilio,F. (1990a) *Nature*, **346**, 763–766.
Simeone,A., Acampora,D., Nigro,V., Faiella,A., D'Esposito,M., Stornaiulo,A., Mavilio,F. and Boncinelli,E. (1990b) *Mechanisms Dev.*, **33**, 215–228.
Sundin,O.H., Busse,H.G., Rogers,M.B., Gudas,L.J. and Eichele,G. (1990) *Development*, **108**, 47–58.
Tear,G., Bate,C.M. and Arias,A.M. (1988) *Development*, **104**, 135–145.
Tiong,S.Y.K., Bone,L.M. and Whittle,J.R.S. (1985) *Mol. Gen. Genet.*, **200**, 335–342.
Wilkinson,D.G. and Krumlauf,R. (1990) *Trends Neurosci.*, **13**, 335.
Wilkinson,D.G., Bhatt,S., Cook,M., Boncinelli,E. and Krumlauf,R. (1989) *Nature*, **341**, 405–409.
Wright,C.V.E., Morita,E.A., Wilkin,D.J. and DeRobertis,E.M. (1990) *Development*, **109**, 225–234.
Zavortink,M. and Sakonju,S. (1989) *Genes Dev.*, **3**, 1969–1981.

Received on March 5, 1991; revised on April 9, 1991