Primary structure and embryonic expression pattern of the mouse Hox-4.3* homeobox gene

JUAN-CARLOS IZPISÚA-BELMONTE, PASCAL DOLLÉ, ARMAND RENUCCI, VINCENZO ZAPPAVIGNA, HILDEGARD FALKENSTEIN and DENIS DUBOULE

EMBL, Meyerhofstrasse 1, 6900 Heidelberg, FRG

* According to a newly proposed update in the Hox nomenclature (Duboule et al. 1990), the HOX-5 complex should be renamed HOX-4 and the Hox-5.4 gene Hox-4.3. This new nomenclature is used throughout this paper with the old names in brackets

1 Author for correspondence

Summary

We report the cloning, genomic localization, primary structure and developmental expression pattern of the novel mouse Hox-4.3 gene. This gene is located within the HOX-4(5) complex, at a position which classifies it as a member of the Hox-3.1 and -2.4 subfamily, the DNA and predicted protein sequences further confirmed this classification. Hox-4.3 has a primary structure characteristic of a Hox gene but, in addition, contains several monotonic stretches of amino acids, one of the 'paired'-like type. As expected from its presence and position within the complex, Hox-4.3 is developmentally expressed in structures of either mesodermal or neuroectodermal origin located or derived from below a precise craniocaudal level. However, a very important offset between anteroposterior boundaries within neuroectoderm versus mesoderm derivatives is observed. Like other genes of the HOX-4(5) complex, Hox-4.3 is expressed in developing limbs and gonads, suggesting that 'cluster specificity' could be a feature of the HOX network.

Key words: homeobox, limb development, Hox-4.3, expression, primary structure, mouse.

Introduction

The mammalian genome contains a number of genes harboring a homeobox related to that present in Drosophila homeotic genes (e.g. Acampora et al. 1989; Schughart et al. 1989). These genes – the Hox genes – are dispersed in four clusters and sequence analysis suggested that these complexes arose during evolution by large scale duplication of an original complex. Comparison of adjacent genes of the same cluster also suggests that this ancestral complex was produced by gene duplication and divergence (for reviews and refs, see e.g. Kappen et al. 1989; Duboule and Dollé, 1989; Graham et al. 1989). Consequently, each Hox gene can be defined with respect not only to its presence within a given complex but also by its relative position therein.

The proteins encoded by such Hox genes are candidate transcription factors believed to be involved in giving positional cues to many different cell types along the developing body axis (e.g. Holland and Hogan, 1988; Gaunt, 1988). However, the mechanism of action of these proteins and their target genes remains to be found. It is generally assumed that different combinations, in the same cell, of Hox proteins could differentially affect the regulation of downstream genes (e.g. Gaunt et al. 1988; Dressler and Gruss, 1989; Holland and Hogan, 1988 for refs.). Indeed, it was recently shown that artificial changes in such protein combinations can induce drastic morphogenetic effects in transgenic mice (Wolgemuth et al. 1989; Balling et al. 1989; Kessel et al. 1990). Therefore, the spatial and temporal developmental expression pattern of these genes is of major importance and needs to be accurately defined. In vertebrates, clearest evidence supports the colinearity concept proposed by Lewis (1978), i.e. that the expression domain of a given Hox gene is determined by its position within the Hox structural network. (Gaunt et al. 1988, 1989, 1990; Duboule and Dollé, 1989; Graham et al. 1989, reviewed by Akam, 1989). Thus, genes located 3' have very anterior boundaries of expression whereas 5'-located genes are expressed very posteriorly. It was further shown that paralogous genes from anterior subfamilies display an apparent redundancy in their expression domains (even though slight differences are clearly observed) whereas the expression domains of 'posterior paralogs' look much less redundant (Gaunt et al. 1989; Erlesius et al. 1990). This very strict correlation between the structural and functional organization of this gene family probably illustrates a fundamental
process by which various cells are instructed regarding their positions or identities.

The mouse HOX-4(5) complex (Featherstone et al. 1988; Stubbs et al. 1990), located on chromosome 2, contains at least 6 homeogenes (-4.1, -4.2(-5.1); -4.4(-5.2); -4.5(-5.3); -4.6(-5.5) and -4.7(-5.6); Lonai et al. 1987; Featherstone et al. 1988; Duboule and Dollé, 1989; Dollé et al. 1989) and has the structural peculiarity, like its human cognate complex (e.g. Acampora et al. 1989), of apparently lacking paralogous genes of three subfamilies (Duboule and Dolle, 1989).

Here, we report the cloning, position, structure and developmental expression of the newly isolated Hox-4.3 gene, a member of the Hox-2.4 and Hox-3.1 subgroup. The results of the expression studies confirmed our expectations based on the position of Hox-4.3 within the HOX-4(5) complex and suggest that the expression domains of some Hox genes could reflect cluster-specificity.

Materials and methods

Isolation and mapping of the mouse Hox-4.3 gene

The Hox-4.3 homeobox was isolated by cross hybridization with probes containing either the mouse Hox-1.6 (Baron et al. 1987) or Hox-4.7(-5.6) (our unpublished work) homeobox. Previous experiments using other mouse probes yielded negative results and the report of the human cognate gene (Oliver et al. 1989) prompted us to investigate once more this region in our original cosmide clones (Duboule and Dollé, 1989).

cDNA library screening

Approximately 1.5×10⁶ phages of a λgt10 cDNA library prepared from 8.5 day p.c. mouse embryo RNA (a gift of B. Hogan and K. Fahrner) were screened under high-stringency conditions (50% formamide, 5×SSC, 1% SDS, 50 μM Tris–HCl pH 7.5, 0.1 mg ml⁻¹ denaturated salmon sperm DNA, 42°C) by hybridization with 2 genomic probes (EcoRI–Aval and EcoRI–PvuII fragments shown in Fig. 1).

In vitro transcription and translation

For in vitro transcription, a hybrid clone (clone A), containing the cDNA linked to 528 bp of 5' sequences (from the ATG to position 528 in Fig. 3) was transcribed using the T7 polymerase according to Melton et al. (1984). The RNA was translated using the rabbit reticulocyte system (Promega) in the presence of 35S-labelled methionine (Amersham). The proteins obtained were separated using SDS–PAGE.

cDNA and genomic DNA sequencing

The cDNA insert was isolated from the positive phage and subcloned into pUC18. The nucleotide sequence of the cDNA clone and genomic fragments were determined using the dideoxy method (Sanger et al. 1977). Difficult regions (high GC content) were sequenced with inosine substituted for the guanosine and/or using the chemical Maxam and Gilbert method (1980).

RNA analysis

Developing forelimbs and hindlimbs of day 12.5 post coitum fetuses were dissected and frozen in liquid nitrogen. Following homogenization in concentrated guanidinium thiocyanate and shearing of the DNA by repeated passage through a syringe needle, total RNA was prepared according to Chirgwin et al. (1979). Poly (A)⁺ RNA was selected by oligo(dT)-chromatography (Aviv and Leder, 1972). For Northern blot analysis 3 μg of poly(A)⁺ RNA was size-fractionated on a 1.2% agarose 2.2 M formaldehyde gel, transferred to nitrocellulose and hybridized to a Hox-5.4 32P-labeled 5' located cDNA probe prepared by random priming.

Fig. 1. Organization of the murine HOX-4(5) complex and of the Hox-5.4 gene. (A) 52 kb of genomic DNA is shown, containing 4 different homeogenes (black boxes; Hox-4.2, -4.3, -4.4 and -4.5.). The genes have the same orientations, 5' to 3' from left to right. (B) Enlargement of a 10 kb genomic region containing the Hox-4.3 gene. The two hatched boxes above represent two probes which were used for screening the cDNA library and for in situ hybridization (see later). (C) The Hox-4.3 open reading frame is shown by the two rectangles, the hatched areas delineate the three regions of 'consensus'; from left to right; the N-terminus, the conserved hexapeptide in the first exon, and the homeodomain in the second exon. 1, 2 and 3 indicate the positions of the monotonic stretches of amino acid (see the text). (D) position of a cDNA clone isolated from a 8.5 day old mouse embryonic library. E, EcoRI; X, Xhol; B, BanHI; P, PstI; A, Aval; PV, PvuII.
Hybridization was carried out for 20h at high stringency (50% formamide, 5×SSC, 1% SDS, 50mm Tris-HCl pH 7.5 and 0.1mg ml⁻¹ denatured Salmon sperm DNA at 42°C). The filters were washed in 0.1×SSC, 0.1% SDS at 65°C and exposed for 3 days at −70°C with intensifying screens.

Preparation of RNA probes
The Hox antisense RNA probe was synthesized using an SP6 in vitro transcription reaction according to standard procedures. The DNA template used was a 220bp EcoRI-AvAI fragment containing part of the Hox-4.3 homeobox (see Fig. 1), subcloned into the vector pGEM I (Promega Biotec). The plasmid was linearized with EcoRI prior to transcription. A control sense RNA probe was synthesized in a T7 polymerase reaction. The sense probe displayed no detectable signal. The control Hox-4.2(-5.1) probe was kindly provided by M. S. Featherstone and the Hox-4.4(-5.2) probe was as described previously (Duboule and Dolle, 1989).

In situ hybridization
Mouse embryos and fetuses were obtained from natural matings between (C57/BL6xSJL) F₁ mice. For aging of embryos, midday of the day of vaginal plug was considered as day 0.5 of pregnancy. Preparation of embryos, sections in paraffin and in situ hybridization were performed as previously described (Dolle and Duboule, 1989), with the exception that the prehybridization step was omitted.

Results
Genomic and cDNA cloning and mapping of the Hox-4.3 gene in the HOX-4(5) complex
The Hox-4.3 homeobox clone was originally isolated by weak homology to a Hox-1.6 homeobox containing clone. A careful analysis of this clone was re-initiated following the publication of the linkage, in human, of the cognate gene to the Hox-4,4(-5.2) gene (Oliver et al. 1989) and the observation that the clone cross-hybridized to the murine Hox-4,7(-5.6) homeobox. This subclone was then mapped using a battery of cosmid clones covering the HOX-4(5) complex and Hox-4.3 was found to lie in a position similar to its human cognate (~10kb downstream of Hox-4.4). In the position between Hox-4.3 and Hox-4.2(-5.1), no other homeobox genes have so far been isolated in mouse or in human (see Figs 1, 2). This suggests that the 3 expected genes are absent from the complex, or so far unidentified at the positions predicted by comparison with other complexes (Fig. 2). Thus, the mouse HOX-4(5) complex contains at least 7 genes (Lonai et al. 1987; Featherstone et al. 1988; Dolle et al. 1989) 4 contiguous members being shown in Fig. 1. As for the other clusters, all these genes have the same transcriptional orientation (5' to 3' from left to right in Fig. 1). Genomic subclones were used to screen a mouse embryonic cDNA library and a cDNA clone was isolated. Comparison with genomic DNA showed the expected presence of an intronic sequence (~450 bp in this case) separating the conserved hexapeptide from the homeodomain (Figs 1 and 3, see next section). Because this cDNA clone was not full length (see the arrow in Fig. 3 for the position of its 5' end), the 5' most part of the coding sequence was determined by sequencing the genomic DNA and comparing with the very highly related sequences of the paralogous genes (see Fig. 4).

Predicted protein sequence and comparisons with paralogous genes and human cognate
The Hox-4.3 predicted protein sequences is similar to all Hox protein sequences reported to date (see e.g. the Hox-1.4 family; Galliot et al. 1989). The homeodomain (shadowed in Fig. 3) is not highly related to either one of the Drosophila homeodomains (Scott et al. 1989) and is found at the carboxy terminus of the protein. Upstream of the homeodomain, a conserved hexapeptide sequence is present (also shadowed in Fig. 3), as reported for most Hox genes (Mavilio et al. 1986; Kessel et al. 1987). As expected, this sequence of unknown function, is encoded by an exon located upstream of the homeobox-containing exon (the black triangle in Fig. 3 indicates the position of the intron). The protein N-terminal region (shadowed in Fig. 3) is also highly conserved both with respect to the classical ‘consensus’ sequence (Met-Ser-Ser; Mavilio et al. 1986) and to the presence of a tyrosine (Tyr) residue between position 10 to 13 (Duboule et al. 1989). The overall length of the expected protein is 289 amino-acids which would thus
Fig. 3. DNA and protein sequence of the Hox-4.3 open reading frame. The three regions of consensus, i.e. the N-terminus, the conserved hexapeptide as well as the homeodomain, are shadowed. The stretches of repeated residues are boxed. They all occur within the first half of the protein, encoded by very GC-rich codons. The position of the intron is shown by a black triangle and the arrow points to the 5' end of the cDNA clone.
The mouse Hox-4.3 gene

Fig. 4. Comparison of the Hox-4.3 protein sequence with the mouse paralogous sequences and the human cognate protein. (A) Alignment of the Hox-4.3 protein sequence with that of the Hox-3.1 and -2.4 paralogs (Breier et al. 1988; Le Mouellic et al. 1988). Resolved conservations are boxed and the absence of a residue(s) is indicated by a dash. The two black triangles point to the start and the end of the homeodomain. The three proteins show an overall high degree of conservation, even outside the regions of consensus. Note the absence of reiterated residues in the Hox-3.1 or -2.4 protein (at least to the extent found in Hox-5.4) (B) Alignment of the Hox-4.3 protein sequence with the sequence available from its human cognate (Oliver et al. 1989). The black triangles are as outlined in A.

encode a protein of approximately $35 \times 10^3 M_r$. This is slightly larger than other predicted molecular weights of murine Hox proteins reported so far (e.g. Kessel et al. 1987; Galliot et al. 1989; Odenwald et al. 1987) and is due to the presence of monotonic stretches of amino acids in the amino-terminal half of the protein. These repeats are essentially composed of GC-rich encoded amino acids. The second of these stretches is similar to the previously reported 'paired repeat' (PRD-repeat; Frigerio et al. 1986). Such amino-acid repeats are often found in developmentally related Drosophila genes (see discussion). The carboxy-terminus of the Hox-4.3 protein is rich in acidic residues (glutamic acid, Glu), as for other Hox proteins (e.g. Kessel et al. 1987).

To investigate whether the protein product of the ORF encodes the Hox-4.3 protein, clone A, containing the cDNA linked to 528 bp of 5' sequences (from the ATG to position 528 in Fig. 3), was transcribed and translated in vitro. As expected, a single protein species is translated in vitro. Its molecular weight appears slightly larger ($\sim 40 \times 10^3 M_r$) than that deduced from the sequence. This difference could be due to the electrophoretic properties of this protein. Numerous termination codons prevent this protein being translated in another reading frame.

The position of Hox-4.3 within the HOX-4(5) complex suggested it would belong to the Hox-3.1 (Awgulewitsch et al. 1986; Breier et al. 1988; Le Mouellic et al. 1988) and -2.4 (Hart et al. 1987; Blatt et al. 1987) subfamily (see Fig. 2) since these three genes have similar relative positions within the Hox-2, -3 and -4(-5) complexes. Hox-4.3 nucleotide and protein sequences confirm this expectation since the sequence is most similar to those of Hox-3.1 and -2.4 (Fig. 4). The homology is not restricted to the homeodomains (located between the two black triangles in Fig. 4) but extends widely outside this region. In particular, the 14 N-terminal residues are conserved between the three
paralogs. A striking difference, however, is the absence in the -3.1 and -2.4 paralogs of the various amino-acid repeats found in -4.3 thus leading to shorter proteins (240 and 232 amino-acids for Hox-3.1 and -2.4, respectively). Alignment of the three protein sequences (Fig. 4) reveals a large area of conservation in the centre of the protein (boxed in the middle line of Fig. 4A or position 379 to 483 in Fig. 3). This conserved region contains a high number of cyclic amino-acids (Tyr, Phe) and acidic residues, and could therefore have a specific function common to all three proteins. Interestingly, this region lies in the same position relative to the homeodomain in all three genes, since in the case of Hox-4.3 the amino-acid repeats are not found between these two domains but are located in a more N-terminal portion of the protein. Fig. 4B illustrates the high protein sequence conservation between the human and mouse Hox-4.3 proteins. The mouse sequence is compared with the human sequences so far available (Oliver et al. 1989). The level of similarity suggest that the human and mouse Hox-4.3 proteins are functionally indistinguishable.

**Hox-4.3 expression during mouse development**

We have previously shown that all the other genes from the HOX-4(5) complex are strongly expressed during limb bud outgrowth and limb development (Dolle and Duboule, 1989; Dolle et al. 1989; Oliver et al. 1989). We therefore searched for Hox-4.3 mRNAs separately in developing limb and trunk of 12.5 day old fetuses. Using a Hox-4.3 specific probe on a Northern analysis, we can detect transcripts in both poly(A)+ mRNA populations. In each case, two transcripts are recognized by the -4.3 probe; 5.0 and 2.2 kb (Fig. 5B). Even though a very slight difference can be seen in the ratio of these transcripts in limb versus trunk, it is probable that both mRNA species are present in significant amount in both structures. We then looked at the fine distribution of the Hox-4.3 transcripts using in situ hybridization of various 35S-labelled antisense RNA probes on paraffin sections of mouse embryos and fetuses. Two different probes were initially tested, all synthesized from the very 3' region of the homeobox plus variable lengths of 3' flanking sequence (Fig. 1). Both probes gave identical signals with higher background for the longer one. We therefore used the shorter probe in all subsequent experiments. In situ hybridizations were performed on serial sections of 9.5, 10.5 and 12.5 days p.c. embryos. In addition, adjacent sections of a 12.5 days fetus were hybridized to probes specific for the 3' and 5' neighbouring genes, Hox-4.2(-5.1), and -4.4(-5.2), in order to obtain a direct comparative view of their transcript domains.

Analysis of 12.5 day p.c. fetuses shows that, similar to other Hox genes, Hox-4.3 has an expression domain that is restricted along the anteroposterior (AP) axis in some ectoderm (the nervous system) and mesoderm derivatives (e.g. the prevertebrae, gut, or kidney (Fig. 6A)). However, an unexpected feature was the very large shift between Hox-4.3 anterior expression boundaries in the central nervous system (CNS) and the mesoderm derivatives where Hox-4.3 transcripts are restricted to rather posterior regions. In the prevertebral column, Hox-4.3 expression is detected very weakly in the 10th thoracic prevertebrae (pv) and strongly in all more posterior pv (Fig. 6B). The anterior boundary of expression in the pv therefore corresponds to pv 18 to 19 (Fig. 6B). Hox-4.3 expression in gut mesenchyme is also restricted to posterior regions of the intestine (Fig. 6A and data not shown). Anterior segments of the alimentary tract such as the oesophagus or stomach are not labelled, neither are the lungs (data not shown). In contrast to this posteriorly restricted expression in mesodermal structures, Hox-4.3 transcripts extend far more anteriorly in the CNS, since the boundary of expression lies in the very posterior hindbrain, close to its junction with the spinal cord (Fig. 6A and C). Hox-4.3 expression boundary in the CNS is, however, clearly more posterior that that of Hox-4.2(-5.1) whose transcripts are present well into the hindbrain (Featherstone et al. 1988 and Fig. 6C). Hox-4.3 expression also extends very anteriorly in the dorsal root ganglia, which are elements of the peripheral nervous system of neural crest origin (see Fig. 6A; labelling in the dorsal root ganglia is much more anterior than the expression limit in the pv indicated by an arrow).

Hox-4.3 transcripts are also detected in the limb buds. By day 9.5 p.c., a strong hybridization signal is found in both forelimb buds (Fig. 7A). The signal is restricted to the limb bud mesoderm, and the adjacent axial mesoderm is not labelled. However, labelling is found posterior to the forelimb buds, specifically in the lateral somatopleure mesoderm (data not shown): a restricted expression in the same region has been described for the Hox-4.4(-5.2) gene (Dolle and Duboule, 1989). Strong labelling is found in the hindlimb buds by day 10.5 p.c. At this AP level, labelling is quite widespread in the embryonic body; nevertheless, the hybridization signal is far more intense within the hindlimb bud than in adjacent axial tissues (Fig. 7B). Later in development, by day 12.5 p.c., labelling by the Hox-4.3 probe is restricted to several central regions of the limbs (precartilaginous condensations), in a way similar to that of Hox-4.4(-5.2) (compare labelling by both probes in Fig. 7C). In addition, specific -4.3 labelling is present in superficial mesenchymal layers of the body wall, mainly in thoracic regions, which are not labelled by the Hox-4.4(-5.2) probe (Fig. 7C). Strong Hox-4.3 labelling is seen in superficial mesenchyme at the base of the head, up to very lateral regions (Fig. 7D). This pattern is similar to that obtained with the Hox-4.4(-5.2) probe but extends much more laterally (e.g. the lateral section shown in Fig. 7D, where the head is barely labelled by Hox-4.4(-5.2) although both probes label the humeral head condensation, arrows). Hox-4.3 transcripts are also detected in fasciculated groups of cells in the very proximal area of the forelimb in the vicinity of the brachial blood vessels (Fig. 7E). We identify these structures as probable nerve trunks of the brachial plexus.
Two gene members of the HOX-4(5) complex (Hox-4.2 and -4.4) were shown to be expressed in the genital ridges. Hox-4.2 is later expressed in the somatic cells of both male and female fetal gonads (Dolle and Duboule, 1989; Gaunt et al., 1989). Hox-4.3 also shows a specific expression in the developing gonads. As early as day 10.5 p.c., the mesonephric column is labelled, but a higher grain density is found along the ventral parts of the mesonephros, in areas from which the genital ridges develop (Fig. 8A). Stronger labelling is seen among cells of the mesonephric tubules, which are segmented epithelial tubes (Fig. 8A, arrowheads). In order to detect possible subtle differences in the transcript domains of the contiguous Hox-4.2(-5.1), -4.3 and 4.4(-5.2) genes in the developing gonads, comparative *in situ* analyses were performed on neighbouring sections of the genital ridges and metanephros of a day 12.5 p.c. fetus. All three probes show very comparable hybridization signals in these structures, essentially labelling the cortex as well as parallel bundles of cells within the undifferentiated gonad surrounding the primitive sexual cords (Fig. 8B). However, quantitative differences are detected in the metanephros, the primordium of the definitive kidney, and in the fetal gonads. Indeed, if the three probes label the metanephric stoma and show strong labelling of the epithelium of the metanephric tubules, the Hox-4.2(-5.1) gene is expressed more strongly in the metanephric tubules than in the genital...
Fig. 7. Hox-4.3 expression in developing limbs. (A) Section of a 9.5 days p.c embryo through the head (h) and otic vesicles, the pharyngeal cavity and forelimb buds (fl). In this section, Hox-4.3 transcripts are found only in the forelimb buds. (B) Section of 10.5 days p.c embryo, through the head (h) and a hindlimb bud (hl). Note that Hox-4.3 transcripts are widely distributed in the posterior part of the body present in this section, but that labelling intensity is much stronger among the cells of the hindlimb bud. (C) Sagittal adjacent sections of a 12.5-day-old fetus across the forelimb (fl) and hindlimb (hl), hybridized to the Hox-4.3 (left panel) and Hox-4.4 (right panel) probes. Note the expression of -4.3 and -4.4 in central condensations of both limbs. (D) Section of the same 12.5-day-old fetus in a more median plane, across the lateral region of the head (h) and proximal region of the forelimb. The arrows point to the humeral head condensation, the only structure labelled by both probes in these sections. e: eye. (E) Sagittal section through a 12.5 day-old fetus, the plane of section is more median that that shown under D. The intense specific white spots represent fascicles of cells, labelled by the Hox-4.3 probe, surrounding the brachial vascular trunks (v).
ridge whereas the neighbouring gene, Hox-4.3, is strong in the gonads but weak in the metanephric tubules (compare various panels of Fig. 8B).

Discussion

This paper reports the isolation, primary structure and expression pattern of the murine Hox-4.3 gene. Accumulating data from many laboratories have led to the conclusion that the structure and expression domains, during development, of a given Hox gene can be deduced from its position within the HOX network (see e.g. Gaunt et al. 1988, 1989, 1990; Duboule and Dollé, 1989; Graham et al. 1989; Dressler and Gruss, 1989). Hox-4.3 is not an exception to this rule even though it does present some peculiarities at the level of both its structure and expression pattern.

The Hox-4.3 gene is the third member of a subfamily

The position of Hox-4.3 within the HOX-4(5) complex strongly suggested that it would be related to the previously described Hox-3.1 and -2.4 genes (Awgulevitsch et al. 1986; Breier et al. 1988), even though a Hox-3.1 probe did not allow us to detect this novel gene earlier (Duboule and Dollé, 1989). The sequence comparisons between these paralogous genes revealed high conservation as already reported for other subfamilies (e.g. Galliot et al. 1989; Graham et al. 1988). If the immediately upstream-located gene, Hox-4.4(-5.2), is related to the Drosophila homeotic gene Abdominal B (AbdB; see Duboule and Dollé, 1989), the homeobox of the Hox-4.3 gene is not clearly similar to those of the Abdominal A (AbdA) or Ultrabithorax (Ubx) Drosophila genes (reviewed in Scott et al. 1989) as one would expect from their genomic positions. In

Fig. 8. Hox-4.3 expression in the developing mesonephros and genital ridges. (A) Transverse section across the mesonephros of a 10.5 day p.c. fetus. Hox-4.3 transcripts are detected along the ventral side of the mesonephros, in the coelomic epithelium and in cells immediately subjacent (arrowheads) as well as in the Wollfian ducts (w) and mesonephric tubules (t). ao, aorta. (B) Three neighbouring sagittal sections crossing the genital ridge (gr) and metanephros (m) of a 12.5 day p.c. fetus, hybridized to the Hox-4.3, -4.4 and -4.2 probes. Although the three probes label the genital ridge and metanephros in a similar way, the Hox-4.2 signal is more intense in metanephric tubules.
addition, no particular sequence similarities are found outside the homeobox within these genes, as is the case for other subfamilies containing both vertebrate and insect genes (refs above).

The Hox-4.3 protein contains large monotonic stretches of amino acids

Many repeats of identical amino acids are found in the amino-terminal half of the Hox-4.3 protein. The very high content of GC nucleotides in the first exon (see also e.g. Galliot et al. 1989) results in repeats of alanine (Ala), glycine (Gly) or proline (Pro) residues. Monotonic repeats of such amino acids have been reported for all the Drosophila homeotic genes (see e.g. for Ubx; Beachy et al. 1985; AbdB; Celniker et al. 1989; Antp; Laughon et al. 1984; Dfd; McGinnis et al. 1984) but their functions remain unclear. Similar repeats have been reported, though to a lesser extent (short poly Ala, poly Gin etc) in other vertebrate Hox genes and a very large poly Gly was recently found within the predicted human Hox-2.7 protein (Acampora et al. 1989). In the case of the Hox-4.3 gene, the fact that these stretches are not found in the two paralogous genes whereas the three genes show a highly conserved region with many acidic residues (in the middle of the protein) suggests that these repeats may not be involved in a putative transactivation function of the -4.3 protein. Rather, these particular sequences could play a role in either separating different protein domains, as proposed earlier for the poly Gly (hinge region) of the Ubx proteins (Beachy et al. 1985), or allowing particular spatial conformation and/or protein–protein interactions. Alternatively, these stretches of amino acids may be related to transcriptional regulatory properties of the -4.3 protein as already shown for other transcription factors (e.g. Courey and Tjian, 1988; Mermod et al. 1989). Comparative functional studies of the three paralogous proteins could be interesting in this respect. The Hox-4.3 protein contains a paired-like repeat (PRD-repeat; Frigerio et al. 1986). Such repeats have been found in several Drosophila genes which are involved in early development (e.g. prd or bcd; Frigerio et al. 1986), their function remain to be elucidated.

Hox-4.3 is expressed, during development, according to its position within the HOX-4(5) complex

The genes of the HOX-4(5) complex were shown to be very strongly expressed in developing limbs (Dolle et al. 1989). Hox-4.3 is also expressed in these structures (see below). We therefore investigated whether different Hox-4.3 transcripts are present in these structures. Both the trunk and limbs contains the same two transcripts recognized by the -4.3 probes probably ruling out the possibility that different transcripts would have different spatial or temporal regulation.

Murine Hox genes are generally expressed in a number of tissues of the developing animal (essentially the CNS and PNS as well as some derivatives of the axial, paraxial and lateral plate mesoderm) in partially overlapping domains which differ in their extensions along the AP (rostro-caudal) axis of the embryo (reviewed in Holland and Hogan, 1988). Each gene has a sharp and specific anterior limit of expression and is widely expressed in more posterior regions of the embryo (refs above). In Drosophila, spatially restricted expression domains are a characteristic of the homeotic genes of the ANT-C and BX-C which are expressed along the anteroposterior axis according to their respective positions within the complexes (e.g. Harding et al. 1985). Such a colinearity, originally proposed by Ed Lewis (1978), has been recently demonstrated in the mouse according to a similar rule: a 3' position in a complex gives rise to anterior expression domains whereas a 5' position gives rise to a posterior expression domain along the cranio-caudal axis (reviewed in Akam, 1989). An extension of this observation to vertebrate limb morphogenesis has been recently reported (Dollé et al. 1989).

Most of the features of the Hox-4.3 transcript domains in the mouse embryo fit with the general scheme of Hox gene expression (e.g. a restricted expression along the AP axis, in the CNS, the prevertebral column and trunk mesoderm derivatives), while some points appear to be more specific to the HOX-4(5) complex (e.g. the strong expression in the developing limbs and fetal gonads). A remarkable feature is the very large offset between the Hox-4.3 expression limits in the CNS and PNS on the one hand, which lies very anteriorly and close to the spinal cord–hindbrain junction, and, on the other hand, in the trunk mesoderm and prevertebral column, which is found far more caudal (prevertebra 18). Differences in expression boundary levels in the CNS and pv have been reported for a number of murine Hox genes (Holland and Hogan, 1988) but with the usual difference of 2 to 3 metameres. However, the restricted expression of Hox-4.3 in the CNS as well as in the trunk mesoderm is in full agreement with the colinearity, since its anterior expression boundary in both structures lies between those of the neighbouring genes, Hox-4.2(-5.1) and -4.4(-5.2) (Featherstone et al. 1988; Dollé and Duboule, 1989). Interestingly, this large offset could be due to the fact that Hox-4.3 does not have immediate 3' neighbour genes. It seems, indeed, that the CNS boundary is that expected for a putative 5' neighbour of the Hox-4.2(-5.1) gene whereas the trunk mesoderm boundary is clearly that expected for a 3' neighbour of Hox-4.4. This observation may illustrate the presence of distinct mechanisms (or a different regulation of the same mechanism) to set up the anterior boundaries of the Hox genes expression domains in different structures.

Paralogous Hox genes show a very high conservation of their nucleotide sequence and appear to have similar expression domains in the embryo (Kappen et al. 1989; Gaunt et al. 1988, 1989). However, a detailed analysis reveals that paralogous genes belonging to the 'Deformed-like' family, Hox-1.4, -2.6 and -4.2(-5.1), as well as to the -1.3, -2.1 and -3.4 subgroup display subtle differences in both their respective anterior expression boundaries and their expression domains in mesoderm derivatives (Gaunt et al. 1989, 1990). Table 1 shows a
The mouse Hox-4.3 gene

Table 1. Comparison of the expression boundaries and domains of the Hox-4.3 paralogous genes (see the text for refs.)

<table>
<thead>
<tr>
<th>Anterior limit in CNS</th>
<th>Posterior hindbrain</th>
<th>Cervical spinal cord</th>
<th>Posterior hindbrain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hox-4.3</td>
<td>Hox-3.1</td>
<td>Hox-2.4</td>
<td></td>
</tr>
<tr>
<td>Anterior limit in pv</td>
<td>pv 18</td>
<td>pv 11</td>
<td>nd</td>
</tr>
<tr>
<td>Lung</td>
<td>−</td>
<td>−</td>
<td>nd</td>
</tr>
<tr>
<td>Stomach</td>
<td>−</td>
<td>−</td>
<td>nd</td>
</tr>
<tr>
<td>Metanephric kidney</td>
<td>+</td>
<td>+</td>
<td>nd</td>
</tr>
<tr>
<td>Genital ridges</td>
<td>+</td>
<td>−</td>
<td>nd</td>
</tr>
<tr>
<td>Limb buds</td>
<td>++</td>
<td>−</td>
<td>nd</td>
</tr>
</tbody>
</table>

This table illustrates the shift in the AP boundary observed with the posterior subfamilies (see Gaunt et al. 1989), the posteriorly restricted expression of these genes (e.g. negatives in lung and stomach and their expression in either the same organs (metanephros) or different organs suggesting a 'cluster specificity'.

Comparison of the Hox-4.3 expression features with those of the two paralogs, Hox-3.1 (Awgulewitsch et al. 1986; Breier et al. 1988; Le Mouellec et al. 1988; Gaunt, 1988; Gaunt et al. 1988) and Hox-2.4 (Graham et al. 1989). All three genes have comparable expression boundaries in the CNS, though Hox-3.1 is slightly more posterior. In contrast, there is a large difference in the prevertebral column between the boundaries of Hox-3.1 (pv 11-12) and Hox-4.3 expression. Such clear-cut differences in the expression domains of paralogous genes have already been observed between Hox-4.4(-5.2) and Hox-2.5 (Duboule and Dole, 1989; Graham et al. 1989). Our study demonstrated that this previous case was not an exception and that the 'posterior' (5' located) genes tend to shift their expression boundaries (among paralogs) to finally specify distinct metameric units. It has been proposed earlier (Gaunt et al. 1989) that the apparent 'anterior redundancy' might be due to the accumulation of important structures along (or derived from) a short anterior fragment of the developing embryo. The results now obtained with the Hox-4.3 gene support this idea. Similar observations were recently reported by studying the mouse Hox-3.2 gene (Erselius et al. 1990).

Hox-4.3 expression in limbs

Hox-4.3 is expressed in the forelimb buds, i.e. in cells derived from an AP level located far more anterior than the Hox-4.3 expression boundary in the trunk mesoderm. This is also true for Hox-4.4(-5.2) and -4.5(-5.3) (Dolle and Duboule, 1989; Oliver et al. 1989) as well as for more 5' located genes of the HOX-5 complex (Dolle et al. 1989) and suggests that the expression of such genes in the limb may be controlled by a mechanism slightly different than that establishing the coordinate expression of the Hox genes along the AP axis of the embryo. Alternatively, this could be due to the absence, in the limbs, of some factors involved in this mechanism (e.g. other Hox proteins). We showed previously that four genes of the HOX-4(5) complex (Hox-4.4, -4.5, -4.6 and -4.7; Dolle et al. 1989) have distinct expression domains in the developing limbs, their expression being more distally restricted for 5' located genes. We therefore suggested that the Hox-4(5) gene products may provide positional information to the mesodermal cells of the developing limbs. Hox-4.3 transcripts are detected in the limb buds at an early stage, when the limbs consist of undifferentiated mesenchyme. Within the forelimbs, transcripts extend up to the very proximal and anterior part of the buds and the expression domain appears at least as wide as that of Hox-4.4(-5.2) (Dolle and Duboule, 1989; Oliver et al. 1989). Thus, the Hox-4.3 expression pattern in the early limb buds perfectly fits with what could be expected from our previous work on the expression of HOX-4(5) genes in these structures.

Later in development, by day 12.5 p.c., Hox-4.3 expression in the limbs becomes regionally restricted, similar to that of Hox-4.2 and Hox-4.4 (Dolle et al. 1989). However, two observations distinguish Hox-4.3 from Hox-4.4: first, Hox-4.4 is expressed in a small area of the superficial mesenchyme at the base of the head. Hox-4.3 is also expressed in this area but more intensely and in a domain that extends more laterally and ventrally. Hox-4.3 and -4.4 distinct expression domains in the head lateral superficial mesenchyme may thus correspond to an extreme proximal part of the forelimb field. Labelled cells could have the same origin as those which express these genes in the early undifferentiated forelimb bud. Since Hox-4.3 is located 3' to Hox-4.4(-5.2) in the HOX-4(5) complex, its expression domain in the forelimb region is expected to be slightly more proximal and anterior than that of Hox-4.4(-5.2). By day 12.5 p.c., the Hox-4.3 expression in the head mesenchyme, more widespread than that of -4.4, would thus accompany its more extended expression domain in the forelimb field. Second, Hox-4.3, but not Hox-4.4, is expressed in the superficial mesenchyme of the body wall and in bundles of cells, probably nerve trunks, from the proximal part of the forelimbs.

The expression of Hox-4.3 illustrates possible cluster specificities

Hox-4.3 transcripts are detected in the genital ridges and, later, in the fetal gonads. Two other members of the HOX-4(5) complex, Hox-4.2(-5.1) and Hox-4.4(-5.2), have been shown to be expressed in the fetal gonads in contrast to the genes located 5' to Hox-4.4 (our unpublished results). In this case, the three genes expressed in the genital ridges are those that show
specific expression in the epithelium of mesonephric tubules. Interestingly, the mesonephric tubules will give rise to the genital excretory ducts, while the major part of the mesonephros will degenerate before birth. We have compared the distribution of transcripts encoded by Hox-4.2, -4.4 and -4.3 in the fetal gonads at 12.5 p.c. and could not detect any qualitative differential labelling in these structures. However, a careful analysis revealed that some quantitative differences may occur. Such quantitative differences between Hox genes in the same structure have been previously described as a possibility to create additional differences in the content of various Hox protein combinations (Gaunt et al. 1989). In addition to this specific expression in gonads, the fact that all the Hox-4(5) genes are predominantly expressed in developing limbs suggests that a particular complex could have distinct expression specificities. Gaunt and colleagues (1990) have recently proposed a similar property for the HOX complexes, showing that two genes of the HOX-3 complex (Hox-3.3 and -3.4) are also expressed in the genital ridges whereas genes from two other complexes are not (Gaunt et al. 1990). Thus, expression in the developing gonads could be an exclusive property of gene members of the HOX-4(5) and HOX-3 complexes. Interestingly, they point out that Hox-3.1 is weakly (or not at all) expressed in developing gonads thus being an exception to this 'complex specificity' (in this case for the genital ridges). In both human and mouse, however, the HOX-3 complex appears to lack one gene, precisely between Hox-3.1 and -3.3 (e.g. Acampora et al. 1989; Schughart et al. 1989; see Fig. 2). Thus, in the HOX-3 complex, genes grouped in the 3' part are expressed in the genital ridges whereas genes located upstream are not. We therefore propose that a given specificity could be attributed to clusters of genes within a complex rather than to complexes themselves. The molecular basis for such a 'cluster specificity' is not known but an attractive explanation relies on the presence of 'master' promoter elements (Simeone et al. 1988) giving certain specificities to either a whole complex (the strong expression of the Hox-4 genes in limbs), or part of the complex (the expression of the 3'-located Hox-3 genes in genital ridges). In any case, this 'cluster specificity' would always be combined to the AP restrictions that colinearity imposes (e.g. the non-expression of the posterior Hox-4 genes-like e.g. Hox-4.5(-5.3)- in the genital ridges).

We would like to thank Prof. P. Chambon in whose laboratory this work was initiated. We also thank Brigid Hogan and Karen Fahrner for the cDNA library and Hakan Jansson for help in the cDNA screening. Special thanks to M. S. Featherstone for the generous gift of the Hox-4.2 (-5.1) in situ hybridization probe, S. J. Gaunt for discussing unpublished results, C. Murphy for useful comments and H. Davies for preparing the manuscript. J.-C.I-B. and A.R. are recipients of postdoctoral fellowships from the Ministerio de Educacion y Ciencia of Spain and EMBO, respectively, and V.Z. is on leave from the Instituto Superiore di Sanita, Roma.

References


The mouse Hox-4.3 gene