An amphioxus homeobox gene: sequence conservation, spatial expression during development and insights into vertebrate evolution

PETER W. H. HOLLAND1, LINDA Z. HOLLAND2, NICOLA A. WILLIAMS1 and NICHOLAS D. HOLLAND2

1Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK
2Scripps Institution of Oceanography, University of California, San Diego, CA 92093, USA

Summary

The embryology of amphioxus has much in common with vertebrate embryology, reflecting a close phylogenetic relationship between the two groups. Amphioxus embryology is simpler in several key respects, however, including a lack of pronounced craniofacial morphogenesis. To gain an insight into the molecular changes that accompanied the evolution of vertebrate embryology, and into the relationship between the amphioxus and vertebrate body plans, we have undertaken the first molecular level investigation of amphioxus embryonic development. We report the cloning, complete DNA sequence determination, sequence analysis and expression analysis of an amphioxus homeobox gene, AmphiHox3, evolutionarily homologous to the third-most 3′ paralogous group of mammalian Hox genes. Sequence comparison to a mammalian homologue, mouse Hox-2.7 (HoxB3), reveals several stretches of amino acid conservation within the deduced protein sequences. Whole mount in situ hybridization reveals localized expression of AmphiHox3 in the posterior mesoderm (but not in the somites), and region-specific expression in the dorsal nerve cord, of amphioxus neurulae, later embryos and larvae. The anterior limit to expression in the nerve cord is at the level of the four/five somite boundary at the neurula stage, and stabilises to just anterior to the first nerve cord pigment spot to form. Comparison to the anterior expression boundary of mouse Hox-2.7 (HoxB3) and related genes suggests that the vertebrate brain is homologous to an extensive region of the amphioxus nerve cord that contains the cerebral vesicle (a region at the extreme rostral tip) and extends posterior to somite four. This proposed homology implies that the vertebrate brain probably did not evolve solely from the cerebral vesicle of an amphioxus-like ancestor, nor did it arise entirely de novo anterior to the cerebral vesicle.

Key words: amphioxus, homeobox, vertebrate evolution, comparative embryology.

Introduction

The phylum Chordata comprises three subphyla: the Vertebrata, the Urochordata (or Tunicata) and the Cephalochordata. The last comprises about 25 anatomically similar species referred to as amphioxus or lancelets. The Chordates all share several derived embryological and anatomical features, notably the presence of a notochord and a dorsal hollow nerve cord. Vertebrates and amphioxus share additional developmental characters, the most obvious of which is the domination of the anteroposterior axis by a metameric series of bilaterally paired somites, which differentiate to give segmented muscle blocks. These similarities in body organisation suggest that vertebrates and amphioxus share fundamental developmental mechanisms employed during formation of the early body plan. They also suggest that amphioxus and vertebrates are sister taxa which diverged after the separation of the lineage(s) leading to the Urochordata (Schaeffer, 1987; Brusca and Brusca, 1990; for discussion of alternatives see Cripps, 1990).

Despite these similarities, there are some anatomical differences between the chordate subphyla. For example, all vertebrates have an elaborate craniofacial region, including a complex subdivided brain, a conserved arrangement of cranial nerves and ganglia, three pairs of special sense organs and extensive mesenchyme derived from migratory neural crest cells; all species of amphioxus lack these features. It has been proposed that these differences represent major evolutionary innovations, which may have required the evolution of new developmental programs (Northcutt and Gans, 1983; Holland, 1988, 1992).

Thus amphioxus are probably the closest living relatives of the vertebrates; they have a similar body organization, and yet they lack (presumably as a primitive not a derived condition) several important vertebrate-specific characters. As such, amphioxus represent living descendants from a critical intermediate stage in vertebrate evolution.

We are undertaking a molecular analysis of amphioxus development. The initial strategy that we selected was to clone amphioxus genes that are true evolutionary homologues of vertebrate genes implicated in the control of embryonic development. By comparing the structure, organization and expression of these genes between
amphioxus and vertebrates, it should be possible to infer how development-regulating genes have been modified during chordate evolution.

Of the many gene families that have been implicated in the control of early vertebrate development, we chose to concentrate on the Antennapedia (Antp) class of homeobox genes (also referred to as the A-class, Class I or Hox/HOM homeobox genes). Mouse and human each have at least 38 of these genes, arranged in four tightly linked clusters thought to have arisen by gene cluster duplication during evolution (Kappen et al., 1989; Simeone et al., 1991). The genes can be classified into 13 ‘cognate groups’ or ‘paralogous subgroups’, each containing two to four genes related by cluster duplication. To date each Hox gene has been ascribed two numbers (e.g: Hox-2.7): the first denoting the gene cluster (Hox-1, -2, -3 or -4), the second reflecting the order of discovery and not necessarily chromosomal order. A nomenclature revision is under discussion at present, whereby the clusters are renamed HoxA, B, C and D, and each gene is ascribed a number denoting the paralogous subgroup (1 to 13 from 3' to 5': M. Scott and W. Bender, personal communication). To avoid confusion, in this paper, we refer to vertebrate Hox genes by the established nomenclature, and give the proposed synonym in parentheses.

The vertebrate Hox-1, -2, -3 and -4 (HoxA, B, C and D) gene clusters are evolutionarily related to the homeotic gene complexes (HOM-C) of Drosophila melanogaster, and accumulating evidence suggests their expression patterns are used as molecular codes for anteroposterior position in vertebrate embryos (reviewed by Holland, 1990; Wright, 1991; McGinnis and Krumlauf, 1992).

We chose to investigate amphioxus Antp class Hox genes, rather than other developmental genes, for two reasons. First, the genomic organization and spatial expression of these genes has been studied in several vertebrates and preliminary evidence suggests these properties are conserved between vertebrate species (Holland, 1992). Second, we suggest that comparison of these genes between amphioxus and vertebrates could yield insight not only into the evolution of developmental mechanisms, but also into ‘spatial homology’ between these taxa. That is, if Hox gene expression patterns have been evolutionarily conserved, we suggest their expression limits could be used as ‘molecular landmarks’ to identify equivalent body regions. Such information would be particularly significant at the anterior end of the body, where amphioxus and vertebrate body organization differs significantly.

Here we report the complete genomic DNA sequence of an amphioxus Hox gene, AmphiHox3, and demonstrate close evolutionarily relationship to a paralogous subgroup of three anteriorly expressed mouse genes: Hox-1.5 (HoxA3), Hox-2.7 (HoxB3) and Hox-4.1 (HoxD3). In situ hybridization to amphioxus embryos and larvae, obtained by in vitro fertilization, revealed region-specific expression of AmphiHox3 in the developing dorsal nerve cord. Comparing the anterior expression limits of AmphiHox3 and the related mouse genes allows us to propose spatial homology between regions of the amphioxus nerve cord and the vertebrate brain. We suggest the vertebrate head evolved from an extensive pre-existing region of the body, not by modification of the extreme anterior body region or by terminal anterior addition.

**Materials and methods**

**Amphioxus collection and in vitro fertilization**

Two species of amphioxus were utilized in this study. We began with adults of Branchiostoma lanceolatum, kindly provided by Dr Quentin Bone FRS (Marine Biological Association, Citadel Hill, Plymouth, UK). These were mated in LB (5% SDS, 250 mM EDTA, 50 mM Tris-HCl pH 8) and stored at 4°C prior to DNA extraction. B. lanceolatum DNA was used for cloning of homeobox sequences from genes related to Hox-2.7 (HoxB3), by PCR amplification. Because of the difficulty in obtaining developmental stages of B. lanceolatum, we used Branchiostoma floridanae for isolation of more extensive genomic DNA clones, prior to gene expression analysis. Adults, juveniles and planktonic larvae of B. floridanae were collected from Old Tampa Bay, Florida, USA in August 1990 and 1991. Earlier developmental stages of B. floridanae were obtained by in vitro fertilization of gametes as previously described (Holland and Holland, 1989, 1992). Embryos and larvae were reared at 24°C in filtered sea water and fixed at times ranging from 13 hours (neurula) to 15 days postfertilization.

**PCR amplification**

Oligonucleotide primers AG3 and AG4 were kindly provided by Dr Anthony Graham (Guy’s Hospital Medical School, London, UK; AG3 = 5’ CTCGACGCTACACCTCCGCCACCTGTGGA 3’ and AG4 = 5’ CTCGACCTTTGCCTTGTTCCTTTT 3’). Amplification and cloning was performed as described by Holland (1991).

**Amphioxus genomic library construction and screening**

High molecular genomic DNA was isolated from fresh frozen adult B. floridanae by standard methods, and further purified by CsCl gradient centrifugation. Genomic DNA was partially digested with Sau3A to give a significant proportion of fragments between 8 and 20 kb. A genomic library was constructed in the Xhol site of bacteriophage vector Lambda Fix II (Stratagene, La Jolla, CA, USA) using a partial fill-in strategy to avoid multiple inserts, following the suppliers protocols. A total of 4x10^6 recombinants from the unamplified library were screened by the method of Benton and Davis (1977) using a short PCR-derived Hox clone at high stringency. The three most strongly hybridizing plaques shared several common restriction fragments; one of these clones, Bfg2, was selected for sequence and expression analysis.

**Sequence determination and analysis of AmphiHox3**

The hybridizing region of Bfg2 was contained within a 5 kb HindIII fragment, which was isolated, subcloned and restriction mapped. A contiguous sequence of 3477 base pairs was derived from within the clone, following further subcloning. Sequence determination was performed on double-stranded plasmid DNA using T7 DNA polymerase, 7-deaza dGTP sequencing mixes and vector primers following the suppliers’ protocols (USB, Cleveland, Ohio, USA; Promega, Madison, WI, USA). Internal primers were also synthesized and used where necessary; Mn buffer was used for sequencing close to primers (USB, Cleveland, Ohio, USA). 90% of the sequence was determined on both strands, including any ambiguous bases and all putative protein-coding regions, even though no compressions were found. All predicted restriction sites were sequenced across to ensure nucleotides were not lost in subcloning. Sequence editing was performed manually; primer design was aided by the OLIQO 2.0 program (MedProbe, Oslo, Norway); protein sequence alignments were aided by CLUSTAL (Higgins and
Amphioxus homeobox gene

Whole mount in situ hybridization

Preliminary experiments indicated that previously published whole mount in situ hybridization protocols, designed for insect, mammalian or amphibian embryos, were unsuitable for amphioxus embryos and larvae. We developed the following protocol (a hybrid of the methods of Tautz and Pfeifle, 1989, Harland, 1991 and Wilkinson, 1992) which proved successful for localization of homeobox and other gene transcripts in amphioxus embryos and larvae ranging from 13 hours to 7 days postfertilization. All steps are at room temperature (approximately 25°C) and pH 7.5 unless otherwise stated.

Embryos and larvae were fixed in 4% paraformaldehyde in 0.5 M NaCl, 0.1 M MOPS pH 7.5, either at room temperature for 1 hour or at 4°C for 12 hours, prior to storage in 70% ethanol at −20°C. After several washes in NaPBSTw (20 mM NaPO₄ buffer, 0.9% NaCl, 0.1% Tween 20), specimens were partially digested with 7.5 µg/ml (embryos younger than 24 hours) or 10 µg/ml (embryos and larvae older than 24 hours) proteinase K in NaPBSTw for 10 minutes. Digestion was stopped with 2 mg/ml glycin in NaPBSTw (10 minutes) and specimens refixed for 1 hour in 4% paraformaldehyde in NaPBS (20 mM NaPO₄ buffer, 0.9% NaCl). After washing in 0.1 M triethanolamine pH 8.0 (10 minutes), specimens were acetylated with 0.25% acetic anhydride in 0.1 M triethanolamine pH 8.0 (2 × 5 minutes), washed in NaPBSTw (10 minutes) followed by prehybridization solution (50% formaldehyde, 5 × SSC, 5 mM EDTA, 0.02% BSA, 0.02% polyvinyl pyrolydine, 0.02% Ficoll 400, 1 mg/ml yeast RNA, 100 µg/ml heparin, 0.1% Tween 20 pH 7).

After transfer to 200 µl fresh prehybridization solution and incubation at 60°C for 1 hour, the volume was reduced to 50 µl and 1 ng/µl probe added. The probe was a 700 nucleotide digoxygenin (DIG)-labelled antisense transcript, synthesized following the kit suppliers instructions (Boehringer Mannheim DIG RNA Labelling kit) and used unhydrolysed after DNAase I treatment, phenol/chloroform extraction and ethanol precipitation. The subclone used for probe preparation consisted of an AmphiHox3, subcloned into pKS+ (Stratagene, La Jolla, CA, USA).

The hybridization tube was placed in a waterbath at 70°C which was allowed to cool slowly to 60°C, before overnight incubation at 60°C. Embryos and larvae were washed in 50% formalamide, 5 × SSC, 1% SDS (3 × 15 minutes, 60°C), then 50% formalamide, 2 × SSC, 1% SDS (5 minutes, 60°C), then the tube allowed to cool to room temperature for 10 minutes; then 2 × 15 minutes at room temperature), and 2 × SSC, 0.1% Tween 20 (2 × 2 minutes); treated with 20 µg/ml preboiled RNAase A and 10 u/ml preboiled RNAase T1 in the same solution (20 minutes, 37°C); and washed in 2 × SSC, 0.1% Tween 20 (2 × 2 minutes) followed by 0.2% SSC, 0.1% Tween 20 (1 × 20 minutes). For immunological detection of the DIG label, the specimens were blocked (30 minutes) in 2 mg/ml BSA, 10% normal sheep serum (pretreated at 55°C for 30 minutes) in NaPBSTw (30 minutes), before incubation with 1:3000 Boehringer Mannheim alkaline-phosphatase-conjugated anti-DIG (preabsorbed to ‘amphioxus powder’ prepared from adult B. floridæ using the ‘embryo powder’ protocol of Wilkinson, 1992) in 2 mg/ml BSA, 5% sheep serum in NaPBSTw (48 hours, 4°C). Washes were in NaPBSTw (4 × 20 minutes), then alkaline phosphatase buffer (100 mM NaCl, 50 mM MgCl₂, 100 mM Tris-HCl pH 9.6, 0.1% Tween 20) before signal detection using NBT and BCIP following the suppliers instructions (Gibco BRL ‘ImmunoSelect’), except that 2 mM levamisole was included in the staining buffer. Reactions were stopped in PBS after 1 to 12 hours, embryos and larvae cleared in 80% glycerol in PBS and viewed and photographed using Nomarski optics.

Results

Cloning and sequence analysis of AmphiHox3 from B. floridæ

We constructed a genomic DNA library from the amphioxus Branchiostoma floridæ and screened this for homologues of the third-most 3’ paralogous subgroup of vertebrate Hox genes. This paralogous subgroup, which includes Hox-1.5 (HoxA3), Hox-2.7 (HoxB3) and Hox-4.1 (HoxD3), is one of the two anteriorly expressed vertebrate Hox subgroups putatively homologous to the proboscipedia gene of Drosophila (Krumlauf, 1992). B. floridæ was selected for library construction and expression analysis since embryos can be reproducibly obtained from this species. The probe used for library screening was a 181 base pair clone from a Hox-2.7 homologue cloned by PCR amplification from B. lanceolatum genomic DNA using primers AG3 and AG4.

Overlapping genomic clones from a B. floridæ Hox gene were obtained and restriction mapped; a 3477 nucleotide contiguous DNA stretch from this gene was completely sequenced (Fig. 1). The sequence contains a homeobox that is clearly related to the third-most 3’ paralogous subgroup of vertebrate Hox genes. We therefore designate the gene AmphiHox3.

Two non-overlapping long open reading frames are present in the sequence, separated by a region containing stop codons, which we propose reflects a single intron separating two exons. This organization is also suggested by sequence comparison to vertebrate Hox genes. In common with many vertebrate Hox genes, the AmphiHox3 sequence predicts (a) the homeobox is contained within the most 3’ exon, (b) downstream of the homeobox is a stop codon before a long 3’ untranslated region, (c) there is an intron 5’ of the homeobox, and (d) a conserved hexapeptide sequence is encoded just 5’ of this intron. Since we have not isolated a cDNA clone, the precise intron/exon junctions have not been defined experimentally; they are predicted from the positions of consensus donor and acceptor splice junctions upstream of the homeobox.

The deduced translation product of AmphiHox3 cannot be compared in detail with all the closely related vertebrate Hox genes, since only one member of the subgroup, mouse Hox-2.7 (HoxB3), has been studied in detail (Sham et al., 1992). Optimal alignment of the deduced AmphiHox3 translation product with the deduced protein product of Hox-2.7 (HoxB3) reveals several conserved features (Fig. 2). The two deduced proteins are similar in their N-terminal sequences (15 identities over 31 sites), including conservation of an unusual Met-Gln-Lys sequence at the translation start site (noted in Hox-2.7 by Sham et al., 1992) as divergent from the Hox consensus of Met-Ser-Ser. The mouse and amphioxus-deduced protein sequences share another conserved region (13/16 identities) N-terminal of the homeodomain, encompassing a consensus hexapeptide sequence characteristic of most Hox proteins (Krumlauf et al., 1987).

Interestingly, the mouse Hox-2.7 sequence predicts a stretch of 26 glycine residues (interrupted by 2 serine residues) between the hexapeptide and the homeodomain. This unusual feature is not present in AmphiHox3; although a stretch of seven contiguous glycine residues is present near
Amphioxus homeobox gene

the N-terminal conserved domain. It is unlikely that these regions have a common descent from an ancestral glycine repeat due to their differing locations in the proteins. The homeodomain is well conserved between the two proteins (55/60 identities) and in each case is followed by a long carboxy terminal domain: longer in AmphiHox3 than in Hox-2.7. At the extreme C terminus, the proteins share another conserved domain (10/14 identities), making a total of four distinct patches of sequence conservation between the proteins (five if the glycine repeats are considered).

Spatial expression of AmphiHox3 during amphioxus development

Initial attempts to detect AmphiHox3 RNA using non-radioactive whole mount in situ hybridization were unsuccessful. By modification of standard protocols, we developed a procedure that clearly detected AmphiHox3 expression in B. floridae embryos and larvae ranging from 13 hours to 7 days postfertilization. The sites, and spatial patterns, of expression are described below. In larvae older than 7 days, signals were not consistently detected in whole mounts. We believe this reflects a technical difficulty (perhaps associated with an increase in extracellular matrix material) rather than lack of AmphiHox3 expression, since radioactive in situ hybridization to sections gave signals at these stages (not shown).

Amphioxus embryonic and larval development is described by Hatschek (1881), Lankester and Willey (1890) and Conklin (1932); here we give only a brief outline to aid interpretation of our in situ hybridization results. At the neurula stage (13 hours postfertilization), the earliest embryonic stage that we analyzed in this study, B. floridae embryos are approximately 150 µm in length and already motile using numerous cilia. Dorsally, the neural plate has sunk completely beneath the epidermis and forms an enclosed furrow, the neural folds having not yet fused dorsally to form a hollow tube. The neural canal is open to the exterior anteriorly at the neuropore, such that there is a continuous channel leading from the exterior, along the neural canal, to the posterior opening of the archenteron cavity. Five pairs of somites are visible in whole mounts using Nomarski illumination, and lie in an anteroposterior sequence lateral to the dorsal wall of the archenteron and the neural plate.

Whole mount in situ hybridization to neurulae (fixed at...
13 to 17 hours postfertilization) reveals Amphi\textit{Hox3} expression in a domain that encompasses the posterior part of the neural furrow and a region of mesoderm surrounding the posterior archenteron, including the archenteron/neural canal connection (Fig. 3A,B). Expression in the neural furrow has a defined anterior cut off but no posterior restriction; the anterior limit is not uniform along the dorsoventral axis, being more anterior dorsally. No expression is detected in the somites, hence expression cannot be described precisely relative to a segmental series. However, the anterior expression limit in the neural furrow is approximately at the level of the boundary between the fourth and fifth clearly visible somite pairs (Figs 3, 4). There is disagreement in the literature as to whether these correspond to the fourth and fifth (Conklin, 1932), or fifth and sixth (Jefferies, 1986), members of the somitic series. We favour the former interpretation, since by optically sectioning whole mounts using Nomarski optics, we see no paired, cavitated epithelial structures anterior to the somite pair ventral and just posterior to the neuropore.

By 20 to 24 hours of embryogenesis, the body is more elongate; approximately four additional pairs of somites have been added posteriorly, and the neural tube and other tissues similarly extended caudally. Whole mount in situ hybridization at 20 and 24 hours postfertilization reveals \textit{AmphiHox3} expression in the posterior half of the neural tube and the most posterior mesoderm; but not in the somites (Fig. 3C). Between 20 and 24 hours, the first pigment spot appears in the neural furrow, at the level of somite five. This is the first clearly visible positional marker in the developing nerve cord. The anterior limit of \textit{Amphi-Hox3} expression at 24 hours is approximately 10 µm (one half somite length) anterior to this pigment spot (Fig. 3C).

During the next two days of embryonic development, the neural tube closes dorsally, the notochord differentiates, the intersomitic furrows become indistinct and the phasing of the somitic series becomes asymmetrical. In the branchial (anterior ventral) region, the club-shaped gland differentiates on the right side, and the first primary gill slit and the mouth break through the ectoderm on the right and left sides, respectively. We follow Willey (1894) and consider the origin of the mouth to mark the transition from embryo to larva. In situ hybridisation revealed that \textit{AmphiHox3} RNA has the same distribution in late embryos and early larvae; Fig. 3D shows expression at the early larval stage (3.5 days postfertilization). \textit{AmphiHox3} RNA is detected in a subset of cells in the neural tube, located in a domain extending posteriorly from a limit just anterior to the most distinct pigment spot, and spatially continuous with expression at the base of the newly developed tail fin. The same expression pattern was detected in 5 day postfertilization larvae (not shown).

The pattern of \textit{AmphiHox3} gene expression is remarkably similar between the wide temporal range of developmental stages that we studied. In the mesoderm, the expression domain does not extend as the body axis elongates; RNA is always restricted to a small patch of posterior tissue. In contrast, the domain of expression in the presumptive nerve cord dramatically increases in length during development, since the expression domain apparenently respects only an anterior boundary (Figs 3, 4). One developmental change in neural expression that is evident relates to the proportion of expressing cells. Staining detected in the neural groove of 13 hour embryos is uniform; in contrast, staining is patchy in the \textit{AmphiHox3}-expressing domain of the neural tube at 3.5 and 5 days, indicating that only a subset of cells contain \textit{AmphiHox3} RNA.

**Discussion**

The proposal that vertebrate \textit{Hox} gene clusters are evolutionarily homologous to the homeotic gene complexes (HOM-C) of \textit{Drosophila} generated much excitement in developmental biology, since it implies that mechanisms for positional specification along the major body axis may be conserved between extremely divergent taxa (Duboule and Dollé, 1989; Graham et al., 1989). The suggested functional similarity is also supported by comparison of the embryonic expression patterns of \textit{Hox} and HOM-C genes and by their mutant phenotypes (reviewed by McGinnis and Krumlauf, 1992). However, it has been difficult to undertake more detailed comparisons between vertebrate \textit{Hox} genes and \textit{Drosophila} HOM-C genes for two principal reasons. First, few aspects of body organization (and hence embryology) are shared between insects and vertebrates. Therefore, comparing \textit{Hox} and HOM-C genes can give few clues as to how the complexities of \textit{Hox} or HOM-C gene deployment arose during evolution. Second, the number of \textit{Hox} or HOM-C genes, and their organization, is considerably different between \textit{Drosophila} and vertebrates; not all \textit{Hox} genes have direct homologues in the HOM-C and vice versa. Indeed, there are only two or three cases where a direct relationship between a single \textit{Drosophila} HOM-C gene and a single paralogous subgroup of vertebrate \textit{Hox} genes can be confidently suggested (\textit{Dfd}, \textit{lab} and possibly \textit{Scr}; Krumlauf, 1992; Holland, 1992).

These limitations could be overcome if \textit{Hox} genes were compared between animals with a closer phylogenetic relationship and more similar body plans. We therefore reasoned that analysis of amphioxus \textit{Hox} genes would be useful in understanding how the vertebrate \textit{Hox} genes evolved their organizational and regulatory complexity. We also anticipated that our analyses could give insight into two additional questions: (1) how are regulatory genes deployed during the embryology of amphioxus? and (2) what are the spatially equivalent regions of the vertebrate and amphioxus body plans?

Using the PCR plus genomic library screening, we identified an amphioxus \textit{Hox} gene, \textit{AmphiHox3}, that is clearly homologous to the third-most 3′ paralogous subgroup of vertebrate \textit{Hox} genes (subgroup 3 following the proposed nomenclature revision), which in turn may be related to the \textit{proboscipedia} gene of \textit{Drosophila}. Our data do not reveal the total number of genes related to subgroup 3, or the number of \textit{Hox} gene clusters, in the amphioxus genome. From the number of clones obtained, however, we suggest that amphioxus has fewer \textit{Hox} clusters than do higher vertebrates.

Comparison of the deduced protein sequences of Amphi -
*Amphioxus homeobox gene* 659

Amphioxus homeobox gene

Hox3 and one of its mouse homologues, Hox-2.7 (HoxB3), reveals several regions of sequence and organizational similarity. These include the sequences of homeodomains, hexapeptides and their flanking regions, the sequences of the N-terminal and C-terminal regions, and the position of the homeodomain relative to the carboxy terminus. These shared features suggest that selection pressures have been acting to constrain sequence change during the 500 million years since the lineages leading to amphioxus and to vertebrates diverged. One feature that differs between the AmphiHox3 and Hox-2.7 (HoxB3) deduced proteins is the position and extent of the homopolymeric tract of glycine residues. It has been suggested that such repeats may serve a hinge function separating distinct functional domains within a flexible protein (Beachy et al., 1985; Sham et al., 1992); however, it seems unlikely to us that two related proteins would both require such a hinge, but in different regions of the protein. Hence the functional significance of the glycine repeats, like many of the conserved sequence features, escapes us. A first step towards elucidating potential functions must be to determine whether those features shared by AmphiHox3 and mouse Hox-2.7 (HoxB3) are also shared by the other two vertebrate genes in the paralogous subgroup: Hox-1.5 (HoxA3) and Hox-4.1 (HoxD3).

The spatial expression pattern of AmphiHox3 has implications for understanding the evolution of Hox gene regulation, amphioxus developmental biology and the origin of vertebrates. We detected AmphiHox3 RNA within a spatially restricted domain of the developing amphioxus dorsal nerve cord, which respected a developmentally static anterior limit. An anterior, but no posterior, limit to expression is also seen in most vertebrate Hox genes, including the homologues of AmphiHox3. We conclude that this property is an ancient feature of Hox gene regulation which predated the divergence of the amphioxus and vertebrate lineages. We also suggest, by analogy with vertebrates, that the amphioxus embryo utilises spatially restricted patterns of Hox gene expression to control region-specific development in the central nervous system (CNS).

Many vertebrate Hox genes are expressed within antero-

![Fig. 4. Homologous Hox gene expression limits in amphioxus and vertebrate. A to C show how the anterior limit to AmphiHox3 neural expression (indicated by the arrowheads) remains static from 13 hours (A) to 3.5 days (C) postfertilization, despite elongation of the body axis. A, B, C (but not D) drawn to the same scale. D shows a chick embryo for comparison (to represent a generalised vertebrate; modified from Lumsden, 1990) showing the major segmental series of the head and brain. The double-headed arrow relates the anterior neural expression limit of AmphiHox3 with that shown by its vertebrate homologues (rhombomere 4/5 boundary). We propose that these positions in the two body plans are homologous. a, archenteron; b, branchial arch; c, club-shaped gland; d, diencephalon; g, first gill slit; m, metencephalon; mo, mouth; n, nerve cord; no, notochord; np, neuropore; p, pigment spot; r1-8, rhombomeres; s, somite; sc, spinal cord; t, telencephalon.](image)
posterior restricted domains outside the nerve cord, including in neural crest derivatives and somitic mesoderm. The neural crest expression cannot be compared to amphioxus (which appears to lack neural crest cells); but extrapolation from vertebrates might predict expression of amphioxus Hox genes in somites. However, we did not detect AmphioHox3 expression within the segmented mesoderm; indeed the only embryonic expression that we observed outside the neural tube did not conform to the rule of anteroposterior limitation, and was always confined to the most-posterior mesoderm despite elongation of the body axis.

What might the significance be of the similarities and differences between AmphioHox3 and vertebrate Hox gene expression? The similarity in CNS expression, but difference in mesodermal expression, may indicate that Hox gene involvement in spatial patterning of the CNS is more ancient than Hox gene deployment in the segmented mesoderm. This conclusion is dependent on three assumptions: (1) that regulation of amphioxus Hox genes represents the primitive condition with respect to vertebrate Hox genes, (2) that AmphioHox3 is typical of other amphioxus Hox genes, and (3) that lack of RNA detection reflects lack of gene expression. Interestingly, the above hypothesis gains indirect support from two other sources. First, mouse Hox genes related by cluster duplication often have similar expression patterns in the CNS, but differ in other aspects of regulation. If functional divergence of Hox genes following cluster duplication was dominated by gain of genespecific controls (as opposed to secondary loss), this also suggests the ancestral site of Hox gene deployment is the CNS (Gould, 1992; Holland, 1992). Second, it has recently been shown that a regulatory element from upstream of the Drosophila Dfd gene can drive reporter gene expression in mouse embryos within an anteroposterior restricted domain (Awgulewitsch and Jacobs, 1992). This spatial domain is limited to the CNS, again suggesting that Hox gene regulatory elements controlling spatial expression in the CNS may be extremely ancient. These diverse lines of evidence for ancestral Hox gene functions in the CNS may imply an earlier evolutionary origin for CNS segmentation than for mesodermal segmentation in the chordate lineage, as recently suggested by Gould (1992).

As vertebrates and amphioxus are reasonably closely related and, since they have similar developmental strategies and body plans, it may be valid to use the anterior expression limits of Hox genes as clues to spatial homology. For example, it is not known which regions of the amphioxus CNS are homologous to the vertebrate brain; resolving this question will give insight into how the vertebrate brain and head evolved. A popular hypothesis, first proposed by Stieda in 1873, is that the entire vertebrate brain is homologous to the ‘cerebral vesicle’ of amphioxus, located in the most rostral region of the nerve cord (reviewed by Willey, 1894). Alternatively, it is possible that part or all of the vertebrate brain is homologous to the cerebral vesicle plus anterior spinal cord of amphioxus (eg: Huxley, 1875; Balfour, 1885). We suggest these hypotheses may be evaluated by comparing the expression patterns of Hox genes between amphioxus and vertebrates, if the function and regulation of these genes has been conserved. In Fig. 4 we compare the anterior CNS expression boundary of AmphioHox3 with that of its three homologues in mouse or chick (Hunt et al., 1991). The three vertebrate genes each have an expression boundary at the anterior limit of rhombomere 5 in the hindbrain. The equivalent Hox gene expression limit in amphioxus is posterior of the cerebral vesicle; instead, the anterior limit is at the level of the four/five somite pair in neurulae, or just anterior to the first CNS pigment spot after this develops. We conclude that the vertebrate brain is spatially homologous to an extensive region of the amphioxus anterior CNS and that, contrary to the proposal of Stieda, it is not the direct homologue of the amphioxus cerebral vesicle. This suggestion is also consistent with the recent discovery of serotonergic neurons posterior to the amphioxus cerebral vesicle (N. Holland and L. Holland, unpublished data), since these may be homologous to the supraspinal serotonergic neurons concentrated in the anterior hindbrain of vertebrates (van Dongen et al., 1985).

Spatial homology between the body plans of two animals can be used to infer the process of evolutionary change, if the polarity of transformation between the types of body plan is first assumed. A popular hypothesis is that the vertebrates evolved from an animal with similar body organization to amphioxus (Brusca and Brusca, 1990); whilst Jeffries (1986) argues that amphioxus is secondarily degenerate from an ancestor with a well-developed brain. Interpreted under Jeffries’ scenario, our data suggest that the most rostral body region has been modified, but not been completely deleted, in the lineage leading to amphioxus. If instead we accept the more popular hypothesis, then the spatial homology that we infer above implies that the vertebrate head did not evolve solely by terminal addition to, or elaboration of, the extreme anterior body region, as has previously been suggested (Northcutt and Gans, 1983). Instead, we suggest that the vertebrate head evolved by elaboration and expansion of an extensive preexisting anterior body region. Attempts to elucidate further what course these evolutionary transformations followed, and how they resulted in the complexities of the vertebrate head, will clearly be aided by comparative analysis of additional conserved genes and gene families.

We thank Quentin Bone FRS, Anthony Graham, Philip Ingham, Dick Jeffries, Trevor Jowett, Robb Krumlauf, Frank Ruddle and David Wilkinson for helpful discussions. We are extremely grateful to Julian Lewis and Suresh Jesuthasan for generosity with their time and microscopical advice, Jon Bartlett and Lynn Westcott for technical assistance, J. M. Lawrence for providing laboratory space during part of the study, Anthony Graham for collaborating with preliminary experiments, John McVey for primer synthesis and Barbara Luke for performing scanning electron microscopy which added to our understanding of amphioxus embryology. This research was funded by SERC grants GR/F1507.6 and GR/H06774 (to P. W. H. H. and N. A. W.), The Queen’s College, Oxford (P. W. H. H.); The Royal Society (P. W. H. H.); a travel grant from the American Philosophical Society (to L. Z. H.); and UCSD Academic Senate Research Grant RR-127-S (to N. D. H.).

References

Fig. 3. Expression of *AmphiHox3* during amphioxus development revealed by whole mount in situ hybridization. The alkaline phosphatase substrate used for detection of the probe gives a blue positive signal. Around the periphery of the embryos a non-specific patchy blue staining is also detected, this varies between specimens. In all photographs, anterior is to the left; dorsal to the top. (A) Hybridization to neurula stage embryo (13 hours postfertilization) showing expression in a spatial domain of the forming neural tube, and in a region surrounding the most posterior part of the archenteron. (B) A neurula stage embryo photographed with the focal plane lateral to the neural tube, showing the anterior limit of *AmphiHox3* neural expression relative to somites 4 and 5. (C) Hybridization to an embryo 24 hours postfertilization, showing expression in the forming neural tube, from a limit just anterior to the single pigment spot (arrow) to the extreme posterior, and continuous with the posterior mesoderm. (D) *AmphiHox3* RNA detected within a spatial domain of the nerve cord, and posterior tail bud, of a 3.5 day amphioxus larva. Scale bars: 50 µm. Arrowheads indicate the anterior limits to *AmphiHox3* neural expression. a, archenteron; n, nerve cord; np, neuropore; p, first pigment spot; s, somite.