

Conservation of *Brachyury* (*T*) genes in amphioxus and vertebrates: developmental and evolutionary implications

Peter W. H. Holland¹, Birgit Koschorz², Linda Z. Holland³ and Bernhard G. Herrmann²

¹School of Animal and Microbial Sciences, The University of Reading, Whiteknights, Reading, RG6 6AJ, UK

²Max-Planck-Institut für Immunbiologie, Postfach 1169, D-79011 Freiburg, Germany

³Scripps Institution of Oceanography, University of California, San Diego, CA 92093, USA

SUMMARY

Homologues of the murine *Brachyury* (*T*) gene have been cloned from several vertebrates, and are implicated in mesoderm formation and in differentiation of the notochord. In contrast, the roles of the ascidian *Brachyury* gene may be restricted to presumptive notochord. To understand the evolution of *Brachyury* genes and their developmental roles, we have searched for homologues in amphioxus, representing the third chordate subphylum and the probable closest relative of the vertebrates. We report the isolation of two amphioxus cDNA clones with clear homology to *Brachyury* genes, and demonstrate that these derive from separate loci resultant from a recent gene duplication. This finding represents an exception to the emerging consensus of an archetypal prevertebrate genome in amphioxus. The spatial and temporal distribution of

Brachyury transcripts during amphioxus development is remarkably similar to vertebrate *Brachyury*, in presumptive mesoderm, posterior mesoderm and the notochord. Gene expression extends throughout the anteroposterior axis of the notochord, despite the most rostral regions being a more recent specialization; it also persists into larval stages, despite differentiation into contractile tissue. We propose that roles of *Brachyury* in notochord differentiation are more ancient than roles in mesoderm formation, and that the latter are shared by cephalochordates and all vertebrates.

Key words: amphioxus, *Brachyury*, *T* gene, evolution, notochord, mesoderm

INTRODUCTION

The murine *Brachyury* (*T*) gene was originally identified by its semidominant mutant phenotype, which has been well described (Dobrovolskaia-Zavadskaja, 1927; Chesley, 1935; Gruneberg, 1958; for review see Herrmann and Kispert, 1994). In homozygous mutant mouse embryos, posterior mesoderm formation in the primitive streak is arrested early, resulting in loss of the posterior trunk region beyond the seventh somite. Embryos die on day 11 of gestation due to a severe reduction of the allantois (Glückssohn-Schönheimer, 1944). Notochord differentiation is also strongly affected in *T/T* homozygotes, and, although a head process is formed initially, it does not differentiate and is not extended into the trunk (Herrmann, 1991). Heterozygotes are viable and show a variable tail phenotype plus posterior axial skeleton defects (Kispert and Herrmann, 1994).

The *Brachyury* gene has been isolated by positional cloning (Herrmann et al., 1990) and its expression during mouse development studied extensively. *Brachyury* mRNA and protein expression can be detected transiently in nascent and migrating mesoderm originating from the primitive streak, and then continuously in the notochord. At later stages of embryogenesis, *Brachyury* is expressed in the tail bud whilst mesoderm formation proceeds, whereas in the notochord it persists to the

end of the tail bud stage (Wilkinson et al., 1990; Herrmann, 1991; Herrmann and Kispert, 1994).

Recently the *Brachyury* protein has been shown to act as a transcription factor (Kispert, Koschorz and Herrmann, unpublished data), which binds specifically to a palindromic recognition sequence in vitro. DNA-binding of the *Brachyury* protein is achieved by the N-terminal T-domain (Kispert and Herrmann, 1993), which is not unique to *Brachyury*, but characterizes a novel gene-family with representatives in the mouse, fly and nematode (Bollag et al., 1994; Agulnik et al., 1995). Homologues of the *Brachyury* gene have been isolated from other vertebrates: the clawed frog *Xenopus laevis* (Smith et al., 1991), the zebrafish *Brachydanio rerio* (Schulte-Merker et al., 1992) and the chick (Kispert et al., 1995). The deduced protein sequences and the expression patterns of these genes are similar to mouse *Brachyury*, suggesting evolutionarily conserved roles of this gene within vertebrate lineages. In zebrafish, a mutant has been described (*no tail*, *ntl*) lacking notochord and posterior trunk (Halpern et al., 1993); *ntl* has been shown to be allelic with zebrafish *Brachyury* (Schulte-Merker et al., 1994), thus confirming functional similarity to the mouse gene.

Comparisons between vertebrate *Brachyury* genes led Smith et al. (1991) to suggest that "mesoderm formation occurs through similar mechanisms in all members of the subphylum"

(Vertebrata). This contention was expanded by Beddington and Smith (1993) in a bold attempt to define a 'molecular phylogenetic stage' for early vertebrate development, using similarities in expression patterns of several genes. Such conclusions rely on extrapolation, since mouse, *Xenopus*, chick and zebrafish are all gnathostome (jawed) vertebrates; there are more divergent lineages within the vertebrate subphylum (including sharks, lampreys and hagfish) that have not been included in the comparison.

The roles of *Brachyury* in mesoderm and notochord formation, therefore, might not be common to all vertebrates. Alternatively, since neither mesoderm nor notochord are unique to the vertebrates, the conserved roles deduced for *Brachyury* could have an earlier origin. To resolve these ambiguities, the comparative analyses must be broadened. A homologue of *Brachyury* has recently been isolated from *Drosophila* (Kispert et al., 1994); the *Drosophila* homologue, *Trg*, plays a role in developmental specification of the hindgut and anal pads. Of key importance for understanding the evolution of vertebrate *Brachyury* function are non-vertebrate chordates (amphioxus and tunicates), principally because these are the only invertebrate animals to possess a notochord.

Yasuo and Satoh (1993, 1994) reported the cloning of a *Brachyury* homologue (*As-T*) from the ascidian tunicate *Haliocynthia roretzi*. Their precise description of *As-T* expression, at the single cell level, revealed a significant difference to the expression of mouse, *Xenopus*, chick and zebrafish *Brachyury*. *As-T* RNA is restricted to presumptive notochord cells, commencing after their fate is restricted to notochord only; *As-T* transcripts are not detectable in cells fated to give rise to any non-notochordal cells. Hence, expression of *Brachyury* in notochord development preceded vertebrate origins; in contrast, expression in early mesoderm (other than notochord) probably evolved later. The data do not resolve whether this latter expression, and its associated developmental role, is restricted to vertebrates in general, or only to higher vertebrates.

To further characterize the evolution and developmental roles of chordate *Brachyury* genes, we have searched for homologues of *Brachyury* from amphioxus. We report the isolation and sequence of two amphioxus cDNA clones with clear homology and high sequence identity to vertebrate *Brachyury* genes. We find that these derive from two *Brachyury* loci in the amphioxus genome, resultant from a recent gene duplication. The spatial and temporal distribution of amphioxus *Brachyury* transcripts during embryonic and larval development was found to be remarkably similar to mouse, *Xenopus*, chick and zebrafish *Brachyury* genes, in presumptive mesoderm, posterior mesoderm and notochord. This similarity, and the contrast to ascidian *As-T* gene expression, allows us to propose a model for the evolution of *Brachyury* gene function and implies that two principal roles of *Brachyury* are conserved across all vertebrates plus amphioxus.

MATERIALS AND METHODS

Amphioxus DNA extraction and in vitro fertilization

Amphioxus (*Branchiostoma floridae*) were collected from Old Tampa Bay, Florida, USA in August 1993 and 1994. Adults for DNA extrac-

tion were either minced in BLB (5% SDS, 250 mM EDTA, 50 mM Tris-HCl, pH 8) and stored at -20°C before DNA purification as described by Holland (1993), or homogenized in guanidinium isothiocyanate (4 M guanidinium isothiocyanate, 50 mM Tris-HCl pH 7.6, 10 mM EDTA, 2% Sarcosyl, 1% β -mercaptoethanol) and stored at $+4^{\circ}\text{C}$ before DNA extraction following the method of Garcia-Fernández et al. (1993). Ripe adults were used for in vitro fertilization as described by Holland and Holland (1993); embryos and larvae were reared at 24°C in filtered sea water and fixed at developmental stages ranging from 10 hours (gastrula) to 3.5 days (swimming larva).

PCR amplification of *Brachyury*-related fragments

Degenerate primers corresponded to the highly conserved amino acid sequences NSLHKY (HO5; forward, sense primer) and YQNEEI (HO6; reverse, antisense primer), encoded in the third exon of the mouse *Brachyury* gene (Herrmann et al., 1990).

HO5 (5' to 3'): TCGCTCGAG/AA(C/T)(A/T)(G/C)CCT(GC)CA(C/T)AA(A/G)TA(C/T)G

HO6 (5' to 3'): TGCTCTAGA/(G/A)AT(T/C)TC(T/C)TC(G/A)TT(T/C)TG(G/A)TA

Amplification was performed for 40 cycles of 94°C (45 seconds), 54°C (90 seconds), 72°C (30 seconds), after an initial denaturation at 94°C (3 minutes). PCR products were purified by gel electrophoresis and cloned into pBluescript II KS(+) (Stratagene).

Isolation and sequencing of cDNA clones

An amphioxus cDNA library, prepared using RNA extracted from 3-5 days postfertilization laboratory-reared larvae (L.Z. Holland, unpublished data), was screened using PCR-derived clones, using high stringency conditions described by Herrmann et al. (1987). Two hybridizing clones were isolated, the inserts subcloned into pBlue-script II KS(+) and prepared for sequencing using the Erase-a-base kit (Promega) for generating nested deletions. Clones were sequenced from both strands using the Autoread sequencing kit of Pharmacia and the A.L.F. automatic sequencer.

PCR-typing of individuals

The polymerase chain reaction (PCR) was used to resolve alternative hypotheses concerning the origin of the AmBra-1 and AmBra-2 clones (see text). Primer sequences used were:

TCON-1 (5'-3'): TCCGGTGTACCACGACTCGCA;

TVAR-1 (5'-3'): AAGTTCAAATAGCCGCGCAGG;

TVAR-2 (5'-3'): CTTTACATTATGTACAATGGACGGA.

PCR primer sites were selected after sequence alignment using CLUSTAL4 (Higgins and Sharp, 1989), and checked for compatibility using OLIGO 2.0 (MedProbe, Norway). PCR amplification, using TCON-1 with either TVAR-1 or TVAR-2, was as described by Holland (1993), with 30 cycles of 94°C (45 seconds), 57.5°C (60 seconds), 72°C (75 seconds). The cDNA sequences predict band sizes of 704 and 411 base pairs respectively.

Sequence comparisons and molecular phylogenetic analyses

Percentage sequence identities between amphioxus and vertebrate *Brachyury* sequences were calculated from alignments of amphioxus and vertebrate deduced proteins only; gaps were treated as mismatches. Alignments used the CLUSTAL4 program (Higgins and Sharp, 1989) followed by manual adjustment; the T-domain alignment included 234 sites, whilst the complete coding sequence alignment included 458 sites. Molecular phylogenetic analyses used a 235 amino acid alignment restricted to the T-domains of all chordate *Brachyury* proteins, since the ascidian *As-T* sequence (used here as an outgroup) is divergent over its C-terminal half, making identification of homologous sites uncertain. Estimation of molecular phylogeny was performed using the Neighbour-Joining method

(which does not assume a molecular clock), applied to a protein distance matrix estimated using a Dayhoff PAM matrix model of amino acid substitution (implemented from PHYLIP 3.5; Felsenstein, 1993). Gaps were treated as deletions. Confidence in the phylogeny was assessed by bootstrap resampling of the data ($\times 100$); current models suggest bootstrap values are generally underestimates of statistical significance (Sitnikova et al., 1995).

DNA-binding analyses

Protein was produced from the AmBra-1 cDNA by in vitro translation, and used for DNA-binding assays as described by Kispert and Herrmann (1993). Binding sites used were BS.p, BS3-1, 3-5, 3-6 and 4-1 of Kispert and Herrmann (1993).

Whole-mount in situ hybridization

Embryos and larvae of *Branchiostoma floridae* were fixed in 4% paraformaldehyde in 0.5 M NaCl, 1 mM MgSO₄, 2 mM EGTA, 0.1 M morpholinopropanesulfonic acid buffer pH 7.5, at 4°C for 12 hours. After fixation, specimens were transferred to 70% ethanol and stored at -20°C for up to one year before use. Digoxigenin (DIG)-labelled riboprobes were synthesized in vitro from linearized plasmid following the DIG-UTP supplier's instructions (Boehringer Mannheim). Specimens shown generally used a probe from AmBra-1, covering the full coding sequence plus 50 nucleotides of 3' untranslated region. Probes from outside the coding region spanned from 60 nucleotides (AmBra-1) or 20 nucleotides (AmBra-2) 3' of the stop codon to the end of each clone. Whole-mount in situ hybridization was performed as described by Holland et al. (1992) except that the RNase-treatment was omitted, as little background was detectable after overnight staining, and the alkaline phosphatase-conjugated anti-DIG antibody (Boehringer Mannheim) was used at a final concentration of 1:1500 and preabsorbed against 'mouse embryo powder' instead of 'amphioxus powder'. After staining, the embryos were washed in PBST, post-fixed in 4% formaldehyde/PBST for 1 hour at room temperature, washed twice in PBST and then either prepared for whole-mount microscopy or for sectioning. For whole mounts, embryos were transferred to 40% glycerol/PBST for several hours at 4°C until they sank, and then transferred to 80% glycerol/PBST/0.1% sodium azide. They were mounted on slides in 2-3 μ l 80% glycerol, and viewed and photographed using either bright-field or Nomarski optics on an Axioplan microscope (Zeiss). For histological sectioning, embryos were dehydrated (50% ethanol for 5 minutes, before transfer to 100% ethanol) and embedded in Polaron Embedding Medium according to the supplier's directions (BioRad Polaron Instruments). Sections were cut at 7 μ m with a glass knife on a Jung Autocut 2055 microtome and mounted for microscopy using DPX Mountant (Fluka).

RESULTS

Cloning amphioxus homologues of the *Brachyury* gene

The degenerate oligonucleotides, HO5 and HO6, were designed to amplify a fragment of 135 base pairs conserved between *Brachyury* genes. After PCR amplification of amphioxus genomic DNA, and cloning of the amplified band, twelve recombinant clones were sequenced. This yielded two classes of amphioxus clones with high sequence similarity to vertebrate *Brachyury* genes. Screening of an amphioxus larval cDNA-library with the PCR-derived clones yielded two different cDNA clones, corresponding to the two PCR fragments. We designate these clones AmBra-1 and AmBra-2.

AmBra-1 contains an insert of 2277 bp encompassing a potential open reading frame of 1344 nucleotides (correspond-

ing to a deduced protein of 448 amino acids). The clone ends with a poly(A) tract of 17 adenine residues, preceded by a polyadenylation signal at position 2228 (the complete nucleotide sequence of AmBra-1 is accessible in the EMBL/GenBank/DDBJ databases under accession number X91903). AmBra-2 contains an insert of approximately 1.8 kb, with the potential to encode a protein of 440 amino acids (database accession number P80492). Both deduced protein sequences have clear similarity to vertebrate *Brachyury* proteins (see below).

The two amphioxus cDNA clones share extremely high sequence identity throughout their coding regions (94% amino acid identity over 444 sites), but differ markedly over their 5' and 3' untranslated regions (UTRs). The lengths of the 3' UTRs are also quite different, being 825 nt for AmBra-1 and 377 nt for AmBra-2 (excluding stop codon and terminal A residues). Attempts to find an optimal alignment between the two 3' UTR sequences using the CLUSTAL4 program gave different alignments as the gap penalty was varied, indicating there is no significant 3' sequence similarity. The 5' UTRs contain neither stop signals nor potential start codons in any forward reading frame.

We suggest both clones encode the full protein coding sequences for two reasons. First, both deduced proteins include the N-terminal limit of the conserved T-domain (see below); indeed, the AmBra-1 and mouse *Brachyury* proteins share three identical amino acids downstream of the deduced translation start site (Fig. 1). Second, AmBra-2 has a perfect match to the Kozak consensus translation start sequence (Kozak, 1986) and AmBra-1 fulfils the most critical Kozak criteria (data not shown).

Sequence conservation between *Brachyury* proteins

Comparison between the deduced protein sequences encoded by AmBra-1 and AmBra-2 and related genes reveals two levels of sequence conservation. First, both the AmBra-1 and AmBra-2 deduced proteins contain the highly conserved N-terminal T-domain (underlined in Fig. 1), which has been shown to confer sequence specific DNA-binding (Kispert and Herrmann, 1993). This domain is not unique to *Brachyury* genes, but is thought to characterize a more extensive gene family (Bollag et al., 1994; Agulnik et al., 1995). Over the T-domain, AmBra-1 and AmBra-2 show 75-82% sequence identity to vertebrate *Brachyury* proteins; this is comparable to levels between the vertebrate homologues (78-91%), and is higher than the T-domain similarity between *Brachyury* and divergent members of the T-domain gene family. For example, mouse *Tbx-2* and mouse *Brachyury* share 50% identity over the T-domain (cited by Bollag et al., 1994).

Second, the C-terminal halves of the two deduced AmBra proteins show significant similarity with vertebrate *Brachyury* proteins. Although the level of conservation is lower than over the T-domain, it is still sufficient to allow close alignment (Fig. 1). In contrast, divergent members of the T-domain gene family (*omb*, *Tbx*), plus the *Drosophila* *Brachyury* homologue (*Trg*), show no similarity over this region. The C-terminal half of the ascidian homologue, *As-T*, shares some conserved residues, but aligns far less than the amphioxus proteins (Yasuo and Satoh, 1994; our unpublished analyses).

Over the complete coding sequence, AmBra-1 and AmBra-2 proteins share 55-63% sequence identity with vertebrate

Brachyury proteins; this compares favourably with the 59-60% between zebrafish and tetrapod Brachyury proteins. These comparisons indicate that the AmBra cDNA clones derive from true *Brachyury* homologues; they are not simply more distant members of the T-domain gene family.

PCR typing of individual amphioxus

Several alternative hypotheses can be proposed to explain the very different 3' untranslated sequences of AmBra-1 and AmBra-2, despite the close similarity over their coding sequences. First, the two clones could derive from two *Brachyury* homologues in the amphioxus genome. Second, the clones could represent alternatively spliced mRNA transcripts from alleles of a single *Brachyury* gene. Third, the clones could represent two alleles of a single, highly polymorphic, *Brachyury* gene. (Significant polymorphism has been found in amphioxus homeobox genes; unpublished data). A fourth hypothesis is that either AmBra-1 or AmBra-2 is a hybrid or artefactual cDNA clone, containing sequences not normally found together in mRNA.

We used PCR to resolve between these hypotheses. One positive strand primer was synthesized to match coding sequence shared by AmBra-1 and AmBra-2 (= TCON-1); two negative strand primers were synthesized to complement unique 3' non-coding regions of either AmBra-1 (= TVAR-1) or AmBra-2 (= TVAR-2). PCR using TCON-1 with TVAR-1, and TCON-1 with TVAR-2, was used to amplify genomic DNA extracted from ten individual amphioxus. Hypotheses one to four make different predictions for this experiment. We found that each primer pair amplified a band of a size compatible with the cDNA sequences, in every animal (Fig. 2). This is consistent with the first hypothesis only, and demonstrates that AmBra-1 and AmBra-2 derive from two, non-allelic, amphioxus loci related to vertebrate *Brachyury* genes.

Evolutionary origin of amphioxus *Brachyury* genes

We used molecular phylogenetic analysis to assess the relationship between the two amphioxus *Brachyury* genes and the homologues cloned from vertebrates and ascidian. Fig. 3 shows the evolutionary relationships between the chordate *Brachyury* genes, deduced by Neighbour-Joining analysis. The tree is rooted with ascidian *As-T* as the outgroup, since the branch length to *As-*

T is by far the longest; this root position is also consistent with generally accepted phylogenetic relationships between the chordate subphyla. It is clear from Fig. 3 that the branch order deduced for the vertebrate *Brachyury* genes is completely consistent with accepted models of vertebrate phylogeny. The branch tips are approximately contemporaneous, indicating a relatively uniform rate of molecular evolution within the T-domain. Interestingly, the two amphioxus genes form a well-supported grouping, to the exclusion of all other *Brachyury* genes. We conclude that the *Brachyury* gene duplicated in the lineage leading specifically

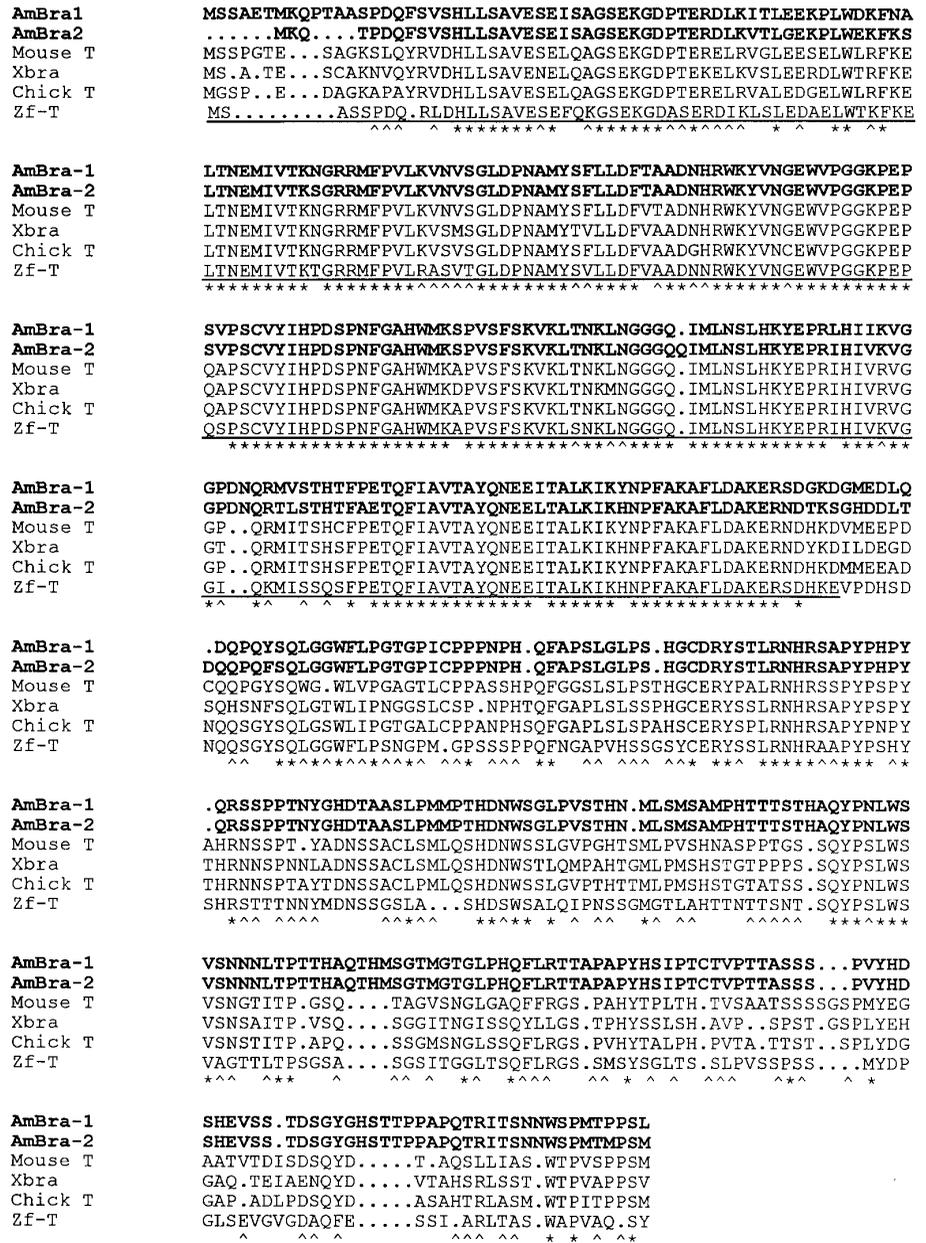


Fig. 1. Complete deduced protein sequences of amphioxus cDNA clones, AmBra-1 and AmBra-2, aligned with vertebrate Brachyury proteins. The highly conserved T-domain is underlined. Gaps introduced to optimize alignment are indicated by dots. Asterisks denote sites conserved between all sequences shown; arrowheads denote partially conserved sites, shared by AmBra-1, AmBra-2 plus at least one vertebrate Brachyury protein. (Database accession nos: AmBra-1 X91903; AmBra-2, P80492)

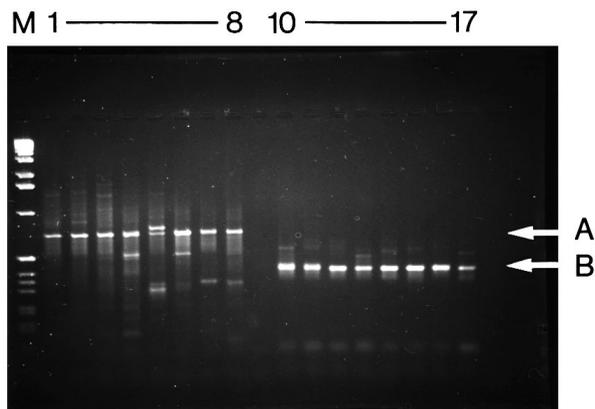


Fig. 2. PCR amplification confirming that *AmBra-1* and *AmBra-2* derive from separate loci. DNA extracted from eight individual amphioxus was amplified using TCON-1 and TVAR-1 (lanes 1-8) or TCON-1 and TVAR-2 (lanes 10-17). Lanes 9 and 18 are control amplifications with no DNA template. Major products of approximately 704 base pairs (lanes 1-8; arrow A) and 411 base pairs (lanes 10-17; arrow B) are seen for every individual, as expected for two loci uninterrupted by introns between the PCR priming sites. Two further individuals gave comparable results. M, molecular mass marker (1 kb ladder, Gibco BRL).

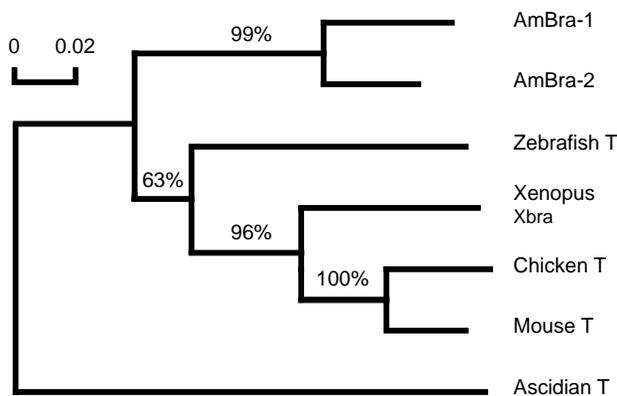


Fig. 3. Evolutionary relationship between Brachyury protein sequences, deduced by molecular phylogenetic analysis using the Neighbour-Joining method and rooted by ascidian As-T. Branch lengths are proportional to evolutionary distance corrected for multiple substitutions; the scale bar denotes 0.02 underlying amino acid substitutions per site. Percentage values are estimates of the robustness of each node, estimated from 100 bootstrap replicates.

to amphioxus; this duplication yielded the loci *AmBra-1* and *AmBra-2*.

DNA-binding properties of *AmBra-1*

The mouse Brachyury protein acts as a transcription factor (Kispert, Koschorz and Herrmann, unpublished data); a palindromic consensus sequence has been identified to which the mouse Brachyury protein binds specifically *in vitro* and when transfected into eukaryotic cells (Kispert and Herrmann, 1993). We therefore tested the capacity of the *AmBra-1* protein to bind DNA *in vitro*. Electrophoretic mobility shift assays (EMSA) were performed using *in vitro* translated *AmBra-1* protein and different Brachyury binding sites. A weak shifted

band, cosegregating with a strong mouse Brachyury protein/DNA complex was observed (data not shown). We conclude that *AmBra-1* can form complexes with the Brachyury binding site *in vitro*, although of reduced stability compared to the mouse Brachyury protein.

Expression of amphioxus *Brachyury* genes

Attempts to detect gene expression using *in situ* hybridisation probes from the 3' untranslated regions of *AmBra-1* and *AmBra-2* were unsuccessful. The reason for this is unclear (although not unprecedented, having been found with 3' UTR probes from some other amphioxus genes; unpublished data). We therefore used probes from the conserved coding region of each clone, which gave indistinguishable signals. We consider these a summation of expression levels from the two genes. Expression was mapped using *in situ* hybridisation to whole-mount preparations of amphioxus embryos and larvae of developmental stages ranging from 10 hours (gastrula) to 2.5 days (swimming larva).

In gastrulae, transcripts were detected in a very broad circumferential ring around the closing blastopore (Fig. 4A). This tissue is thought to invaginate during gastrulation, to contribute to mesoderm and endoderm. In neurula stage embryos (approximately 5 somite pairs; stage N2 of Hirakow and Kajita, 1994), expression is strongest in the dorsal posterior mesoderm, and weaker in more anterior axial mesoderm (Fig. 4B). From its position, we identify this axial tissue as presumptive notochord, although the notochord is not histologically demarcated at this stage (Hirakow and Kajita, 1994). Dorsal views of slightly later neurulae (7 somites), clearly show that the posterior patch of expression is continuous through all mediolateral levels of the mesoderm, with a sharp anterior boundary at all positions lateral to the notochord (Fig. 4C). Axially, expression does not terminate, but extends anteriorly into the forming notochord (Fig. 4C,D,E). Again, we are identifying presumptive notochord from its position. In late neurulae (approximately 10 somite pairs; Hirakow and Kajita stage N3), expression is still detected in a patch of posterior mesoderm, and throughout the forming notochord (now distinguishable using Nomarski optics; Fig. 4F). The posterior patch of expression is still confined to pre-somatic regions, indicating that the expression boundary in non-notochordal tissue does not respect a fixed axial level; expression remains posterior by down-regulation in forming somites.

Histological sections through hybridized embryos, at a series of axial levels, clearly confirm the existence of posterior expression in lateral and axial mesoderm (Fig. 5B,H). This expression domain includes the epithelial cells of the most posterior somite pair, during their formation from the dorsolateral walls of the posterior archenteron (Fig. 5C,I). In more anterior (older) regions, expression persists in presumptive notochord only (Fig. 5D-F, J-L). At the transition between the posterior and anterior expression patterns, *Brachyury* expression is asymmetrical in our sections (Fig. 5D,J); we suggest this may parallel the asymmetry inherent in amphioxus somite formation (right somites are slightly rostral to left somites at this stage; see Jefferies, 1986, for discussion).

At later stages of development (20 hours to 2.5 days), this basic pattern of expression is maintained: transcripts are

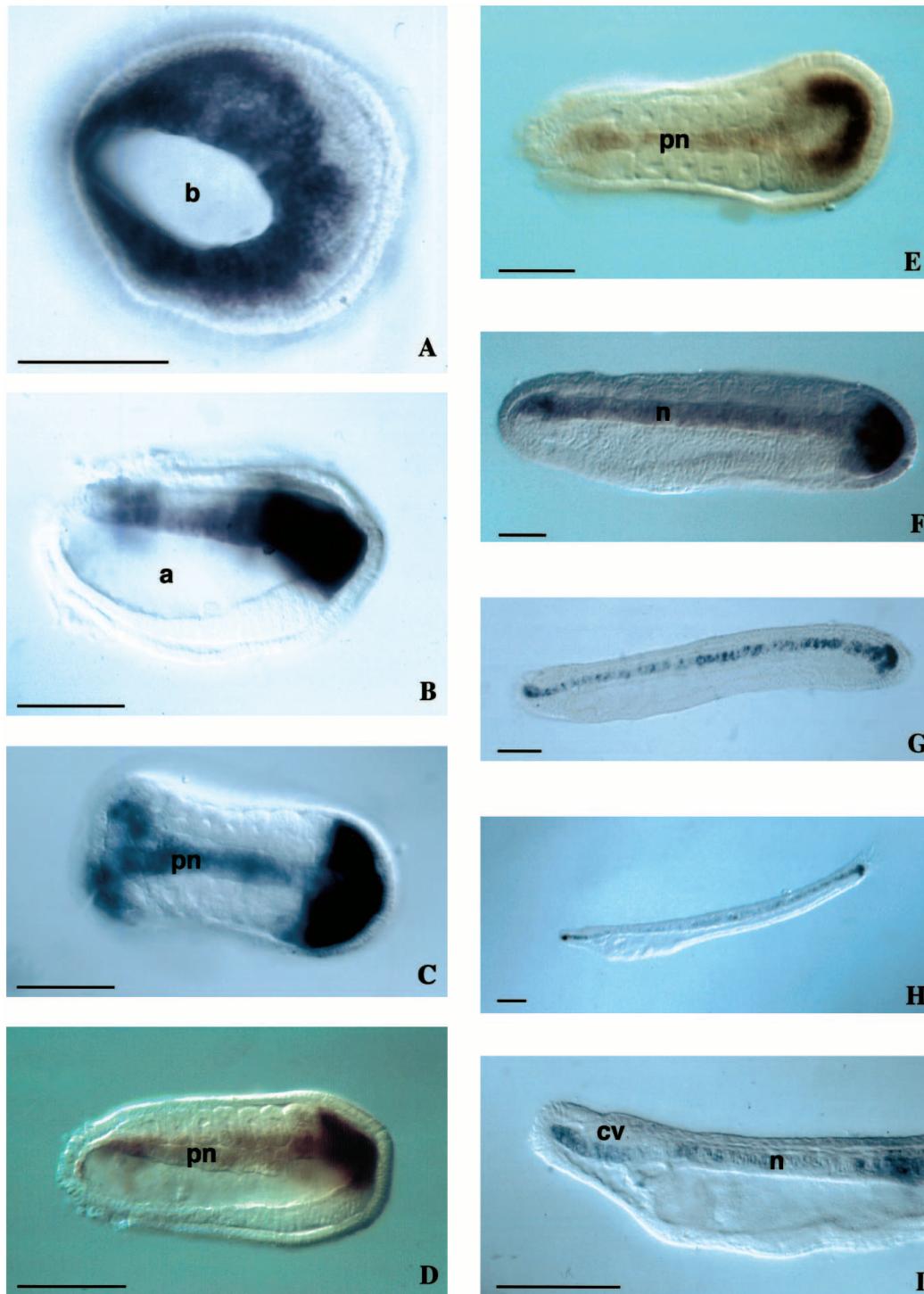


Fig. 4. *Brachyury* RNA expression in amphioxus embryos and larvae, as revealed by whole-mount in situ hybridization. In all specimens from B to I, anterior is to the left; in lateral views (B,D,F-I), dorsal is to the top. (A) Vegetal view of a 10 hour gastrula showing extensive expression around the blastopore. (B) Lateral view of early neurula. (C,D,E) Mid-neurula stage in dorsal (C,E) and lateral (D) perspective, showing expression in dorsal posterior mesoderm and presumptive notochord. (F) 18 hour embryo showing expression in all posterior mesoderm and newly forming notochord. (G) Early, 24 hour, larva and (H,I) later larvae showing expression in posterior mesoderm, plus non-uniform expression throughout the rostrocaudal axis of the notochord to the extreme rostral tip. Abbreviations: a, archenteron; b, blastopore; cv, cerebral vesicle; n, notochord; pn, presumptive notochord. Scale bars, 50 μ m.

detected in a localised patch of extreme posterior mesoderm, and throughout the rostrocaudal axis of the notochord (Fig. 4G,H). The larvae figured correspond to Hirakow and Kajita stages L1 to late L2. The only difference from earlier stages is that expression in the notochord is now somewhat patchy, not being at a uniform level in all cells. It is particularly clear from the larval stages that expression in the notochord extends to the extreme rostral tip of the notochord (Fig. 4I), anterior to the position of the cerebral vesicle (marking the limit of the neural tube).

DISCUSSION

The first clear *Brachyury* homologue to be cloned from an invertebrate was the ascidian gene *As-T* (Yasuo and Satoh, 1993). The sequence similarity between *As-T* and vertebrate *Brachyury* genes, together with phylogenetic considerations, suggested that a homologue of *Brachyury* should also exist in the genome of amphioxus. In this report we describe the isolation of two different cDNA clones related to *Brachyury* from the amphioxus species *Branchiostoma floridae*. Sequence

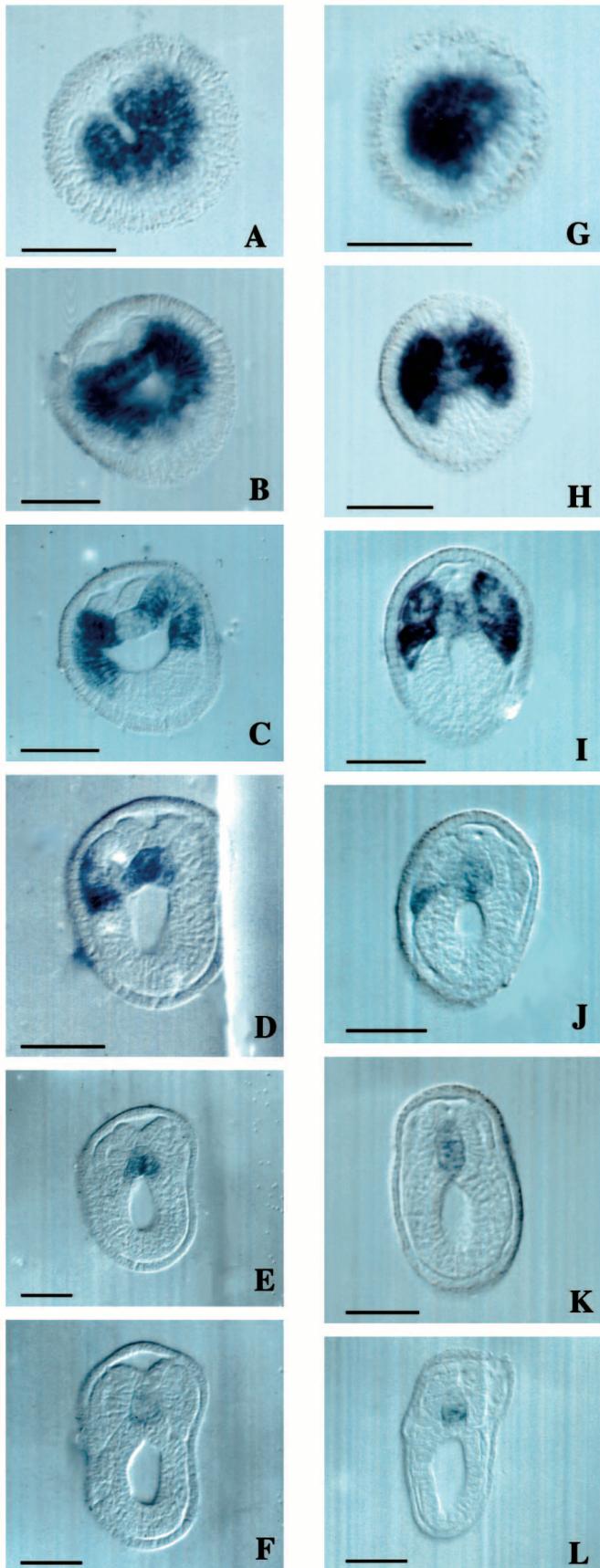


Fig. 5. Transverse histological sections through amphioxus embryos stained in whole mount for *Brachyury* RNA expression. (A-F) early neurula; (G-L) mid-neurula. Each series is arranged from posterior to anterior. Note posterior expression in all mesoderm, including epithelial cells of budding somites (C,I), but anterior expression in notochord only. Scale bars, 25 μ m.

comparisons reveal that the two clones encode proteins with clear similarity to the products of vertebrate *Brachyury* genes, extending over the entire open reading frame. This indicates that the clones are not derived simply from members of an extended gene family sharing the T-domain. Instead, the amphioxus clones represent true homologues of vertebrate *Brachyury* and ascidian *As-T*. Here we use the term homologue to indicate a particularly close evolutionary relationship (clearly, however, all members of a multigene family are related at some level).

The two *Brachyury* cDNA clones derive from two non-allelic genes, as confirmed by PCR analysis of individual amphioxus. This conclusion was unexpected and has several implications. The first is the relevance to current models of genome evolution. We expected a single *Brachyury* gene to be present in the amphioxus genome because (a) multiple *Brachyury* genes have not yet been described in other chordate species, and (b) recent evidence suggests that in many gene families, the amphioxus genome has fewer genes than are present in vertebrates. For example, amphioxus has a single *Hox* gene cluster (Garcia-Fernández and Holland, 1994) and probably single copies of a *Cdx* homeobox gene (Holland et al., 1994), *Msx* homeobox gene (Holland et al., 1994; Sharman and Holland, unpublished data), *Otx* homeobox gene (Williams and Holland, unpublished data), alkaline myosin light chain gene (L.Z. Holland et al., unpublished data) and insulin/IGF gene (Chan et al., 1990); in every case, vertebrates have multiple family members. This contrast between single genes in amphioxus and multiple related genes in vertebrates probably extends to many more gene families, since the genomic location of several of the above genes in mammals suggests they were duplicated as part of large chromosomal regions (Lundin, 1993). Indeed, a strong argument can be made for a duplication of much or all of the genome close to vertebrate origins, with amphioxus retaining the archetypal, preduplication condition (Garcia-Fernández and Holland, 1994; Holland et al., 1994).

Can the existence of two *Brachyury* genes in amphioxus be reconciled with this model? It still remains possible that amphioxus has fewer true *Brachyury* genes than do vertebrates, if the latter possess *Brachyury* genes that have not yet been identified by hybridisation screening or mutagenesis methods. This seems unlikely, hence *Brachyury* seems to represent an exception to the general rule emerging for amphioxus genes. We suggest that a single ancestral *Brachyury* gene was independently duplicated in the lineage leading to amphioxus; in the vertebrate lineage, widespread gene or genome duplications were followed by either subsequent loss or divergence of certain genes, including additional *Brachyury* genes. This scenario is consistent with our molecular phylogenetic analyses of *Brachyury* sequences. The model also predicts that T-box genes that are secondarily divergent from *Brachyury* may be found in vertebrate genomes but not in amphioxus. This cannot be assessed until the full diversity of T-box genes

is known from a range of species, and their DNA sequences examined to reveal which are ancient and which (if any) are secondarily derived from *Brachyury*.

The second implication of the two amphioxus *Brachyury* genes relates to terminology. We outline above why we consider both genes to be true homologues of *Brachyury*. However, we suggest it is incorrect to refer to either gene as an orthologue of vertebrate *Brachyury*, unless more extensive data reveal additional *Brachyury* genes in vertebrates. If, as we suggest, the two amphioxus genes arose by an independent gene duplication, then they represent two paralogous genes that are collectively orthologous to vertebrate *Brachyury* (see Dickinson, 1995, for a general discussion of this point).

The third implication is practical. Repeated attempts to determine gene-specific expression patterns using the 3' UTR probes from AmBra-1 and AmBra-2 were unsuccessful, despite hybridisation to several hundred embryos. The reasons are unclear, although we have encountered similar problems using 3' UTR probes from other amphioxus genes. The principal reason is not technical (as judged using control probes) and possibly not due to excessive polymorphism. We therefore used probes from the protein coding region, which is highly conserved between AmBra-1 and AmBra-2, even at the nucleotide level. The gene expression patterns we describe, therefore, could derive from either gene, or a summation of both (or indeed other, as yet unknown, homologues).

Expression analysis of the amphioxus *Brachyury* homologues is particularly informative for understanding the evolution of *Brachyury* and its developmental roles. One key question was whether amphioxus *Brachyury* genes are expressed like ascidian *As-T* (presumptive notochord only) or like *Brachyury* in mouse, *Xenopus*, chick and zebrafish (notochord and posterior mesoderm). Our in situ hybridisation analyses of amphioxus embryos and larvae revealed that the second alternative is correct: *Brachyury* gene expression is remarkably conserved between vertebrates and amphioxus. In amphioxus, transcripts are first detected around the blastopore of gastrula stage embryos, in presumptive mesoderm. The similarity to the expression pattern of the *Brachyury* homologue in *Xenopus* gastrulae is very striking (Smith et al., 1991). As gastrulation and neurulation proceed, transcripts are seen throughout newly formed posterior mesoderm, and more anteriorly solely in the notochord. The notochord expression persists at least until the swimming larval stage. The implications of this similarity to the homologues in mouse, chick, *Xenopus* and zebrafish are that the dual roles of *Brachyury* in early mesoderm formation and notochord differentiation evolved prior to vertebrate origins and neither role is vertebrate specific. The contention of Smith et al. (1991) that "mesoderm formation occurs through similar mechanisms in all members of the subphylum" (Vertebrata) is strengthened by these conclusions, and extended to cephalochordates. The contrast between amphioxus and vertebrate *Brachyury* on one hand, and ascidian *As-T* on the other, is intriguing. One possibility is that the posterior mesodermal expression seen in amphioxus and vertebrates (but not ascidian) is related to the evolution of axial extension: an aspect of development associated with possession of a true tail?

Although amphioxus retains many primitive chordate developmental features, it has also evolved some unique characters and these deserve consideration. One distinctive feature is that

the notochord extends to the extreme anterior tip of the body, beyond the rostral limit of the neural tube. Willey (1894), following earlier workers, claimed that this rostral region of the amphioxus notochord is slightly retarded in its development, reflecting its later evolution. The peculiar rostral extension of the notochord is unique to amphioxus; indeed, even in the putative fossil cephalochordate *Pikaia*, from the 520 million year old Burgess Shale, the notochord appears to terminate a short way back from the rostral tip (Conway Morris, 1993). It was interesting to find, therefore, that amphioxus *Brachyury* transcripts are distributed throughout the length of the notochord, and quite clearly up to its rostral tip. We detected no evidence for either qualitative differences, or differing timings of expression, between rostral and other notochord. Hence, even though rostral notochord tissue is probably more evolutionarily recent than the rest of the amphioxus notochord, it has come under common developmental control. Our conclusions do not give support to claims that rostral notochord development is histologically and temporally distinguishable (Willey, 1894), although a subtle difference in timing onset would not have been detected.

A second peculiarity of amphioxus notochord development is that it differentiates into a contractile rod, containing horizontal muscle fibres arranged in a series of transverse lamellae (Flood et al., 1969; Guthrie and Banks, 1970). Contractile notochord fibres are clearly present in day 12 larvae (Lacalli et al., 1994), but probably differentiate much earlier before muscular twitching commences (around 24 hours). Indeed, the alkaline myosin light chain gene of amphioxus is expressed in notochord at late embryonic and early larval stages (L. Z. Holland et al., unpublished data). It is interesting, therefore, that *Brachyury* transcripts are still evident in the notochord during larval development, despite differentiation to form contractile fibres. *Brachyury* gene expression is also surprisingly stable in the vertebrate notochord; for example, RNA expression is detected in a repeating pattern in the mouse vertebral column at 17.5 days post coitum, presumably in remnants of the notochord (Wilkinson et al., 1990). We suggest that *Brachyury* has conserved roles in early mesoderm formation (in amphioxus and vertebrates) and in controlling the differentiation and fate of notochordal cells (in all chordates). The latter role appears to have been conserved between chordate subphyla despite evolutionary divergence in the patterns of terminal differentiation.

There is dispute about the phylogenetic relationships between the three chordate subphyla: Vertebrata (here used as synonymous with Craniata), Cephalochordata (amphioxus) and Urochordata (tunicates). Developmental genetic characters should prove useful for resolving this trichotomy. The conservation of *Brachyury* gene expression in the notochord of all three taxa is not phylogenetically informative. In contrast, expression in posterior mesoderm is a character shared by vertebrates and amphioxus only, not being detected in an ascidian. This character can be added to a suite of comparative morphological and molecular data, most (but not all) of which also support a sister grouping of vertebrates and amphioxus (Schaeffer, 1987; Wada and Satoh, 1993). Although we consider this to be the most supported phylogeny, we do not yet consider the data as conclusive on this issue; we urge that the relationships within the Chordata continue to be evaluated in the light of comparative studies

on developmental gene expression, gene sequence and gene family organisation.

We thank Hubert Ortner, Freiburg, for performing a major part of this work. Hubert Ortner was offered first authorship, but declined for reasons of conflict with his religious beliefs. This work was funded by the Deutsche Forschungsgemeinschaft and the BBSRC. We thank Ernst-Martin Fuechtbauer for help with microscopy and photography, John M. Lawrence for kindly providing laboratory space in Florida, and referees for their critical comments.

REFERENCES

- Agulnik, S. I., Bollag, R. J. and Silver, L. M. (1995). Conservation of the T-box gene family from *Mus musculus* to *Caenorhabditis elegans*. *Genomics* **25**, 214-219.
- Beddington, R. S. P. and Smith, J. C. (1993). Control of vertebrate gastrulation: inducing signals and responding genes. *Curr. Opin. Genet. Dev.* **3**, 655-661.
- Bollag, R. J., Siegfried, Z., Cebra-Thomas, J. A., Garvey, N., Davison, E. M. and Silver, L. M. (1994). An ancient family of embryonically expressed mouse genes sharing a conserved protein motif with the T locus. *Nature Genetics* **7**, 383-389.
- Chan, S. J., Cao, Q. -P. and Steiner, D. F. (1990). Evolution of the insulin superfamily: cloning of a hybrid insulin/insulin-like growth factor cDNA from amphioxus. *Proc. Natl. Acad. Sci. USA* **87**, 9319-9323.
- Chesley, P. (1935). Development of the short-tailed mutation in the house mouse. *J. exp. Zool.* **70**, 429-459.
- Conway Morris, S. (1993). The fossil record and the early evolution of the Metazoa. *Nature* **361**, 219-225.
- Dickinson, W. J. (1995). Molecules and morphology: where's the homology? *Trends Genet.* **11**, 119-121.
- Dobrovol'skaia-Zavad'skaia, N. (1927). Sur la mortification sponta-nee de la queue chez la souris nouveau-nee et sur l'existence d'un caractere hereditaire 'non viable'. *C. R. Soc. Biol.* **97**, 114-116.
- Felsenstein, J. (1993). PHYLIP version 3.5c. University of Washington, Seattle.
- Flood, P. R., Guthrie, D. M. and Banks, J. R. (1969). Paramyosin muscle in the notochord of amphioxus. *Nature* **222**, 87-88.
- Garcia-Fernández, J. and Holland, P. W. H. (1994). Archetypal organization of the amphioxus *Hox* gene cluster. *Nature* **370**, 563-566.
- Garcia-Fernández, J., Baguñà, J. and Saló, E. (1993). Genomic organization and expression of the planarian homeobox genes *Dth-1* and *Dth-2*. *Development* **118**, 241-253.
- Glückssohn-Schönheimer, S. (1944). The development of normal and homozygous *Brachyury* (*T/T*) mouse embryos in the extraembryonic coelom of the chick. *Proc. natl. Acad. Sci. USA* **30**, 134-140.
- Gruneberg, H. (1958). Genetical studies on the skeleton of the mouse XXIII: the development of *Brachyury* and *Anury*. *J. embryol. exp. Morph.* **6**, 424-443.
- Guthrie, D. M. and Banks, J. R. (1970). Observations on the function and physiological properties of a fast paramyosin muscle - the notochord of amphioxus (*Branchiostoma lanceolatum*). *J. exp. Biol.* **52**, 125-138.
- Halpern, M. E., Ho, R. K., Walker, C. and Kimmel, C. B. (1993). Induction of muscle pioneers and floor plate is distinguished by the zebrafish *no tail* mutation. *Cell* **75**, 99-111.
- Herrmann, B. G. (1991). Expression pattern of the *Brachyury* gene in whole-mount *T^{wis}/T^{wis}* mutant embryos. *Development* **113**, 913-917.
- Herrmann, B. G. and Kispert, A. (1994). The T genes in embryogenesis. *Trends Genet.* **8**, 280-286.
- Herrmann, B. G., Barlow, D. P. and Lehrach, H. (1987). A large inverted duplication allows homologous recombination between chromosomes heterozygous for the proximal to complex inversion. *Cell* **48**, 813-825.
- Herrmann, B. G., Labelit, S., Poustka, A., King, T. R. and Lehrach, H. (1990). Cloning of the *T* gene required in mesoderm formation in the mouse. *Nature* **343**, 617-622.
- Higgins, D. G. and Sharp, P. M. (1989). Fast and sensitive multiple sequence alignments on a microcomputer. *CABIOS* **5**, 151-153.
- Hirakow, R. and Kajita, N. (1994). Electron microscopic study of the development of amphioxus *Branchiostoma belcheri tsingtauense*: the neurula and larva. *Acta Anat Nippon* **69**, 1-13.
- Holland, P. W. H. (1993). Cloning genes using PCR. In: *Essential Developmental Biology: A Practical Approach* (ed. C. D. Stern and P. W. H. Holland), pp. 243-255. Oxford: IRL Press at Oxford University Press.
- Holland, N. D. and Holland, L. Z. (1993). Embryos and larvae of invertebrate deuterostomes. In: *Essential Developmental Biology: A Practical Approach* (ed. C. D. Stern and P.W.H. Holland), pp. 21-32. IRL Press at Oxford University Press, Oxford.
- Holland, P. W. H., Holland, L. Z., Williams, N. A. and Holland, N. D. (1992). An amphioxus homeobox gene: sequence conservation, spatial expression during development and insights into vertebrate evolution. *Development* **116**, 653-661.
- Holland, P. W. H., Garcia-Fernández, J., Williams, N. A. and Sidow, A. (1994). Gene duplications and the origins of vertebrate development. *Development 1994 Supplement*, 125-133.
- Jefferies, R. P. S. (1986). *The Ancestry of the Vertebrates*. London: British Museum (Natural History).
- Kispert, A. and Herrmann, B. G. (1993). The *Brachyury* gene encodes a novel DNA binding protein. *EMBO J.* **12**, 3211-3220.
- Kispert, A. and Herrmann, B. G. (1994). Immunohistochemical analysis of the *Brachyury* protein in wild-type and mutant mouse embryos. *Dev. Biol.* **161**, 179-193.
- Kispert, A., Herrmann, B. G., Leptin, M. and Reuter, R. (1994). Homologs of the mouse *Brachyury* gene are involved in the specification of posterior terminal structures in *Drosophila*, *Tribolium* and *Locusta*. *Genes Dev.* **8**, 2137-2150.
- Kispert, A., Ortner, H., Cooke, J. and Herrmann, B. G. (1995). The chick *Brachyury* gene: developmental expression pattern and response to axial induction by localised activin. *Dev. Biol.* **168**, 406-415.
- Kozak, M. (1986). Point mutations define a sequence flanking AUG initiator codons that modulates translation by eukaryotic chromosomes. *Cell* **44**, 283-292.
- Lacalli, T. C., Holland, N. D. and West, J. E. (1994). Landmarks in the anterior central nervous system of amphioxus larvae. *Phil. Trans. Roy. Soc. B.* **344**, 165-185.
- Lundin, L. (1993). Evolution of the vertebrate genome as reflected in paralogous chromosomal regions in man and the house mouse. *Genomics* **16**, 1-19.
- Schaeffer, B. (1987). Deuterostome monophyly and phylogeny. *Evolutionary Biology* **21**, 179-235.
- Schulte-Merker, S., Ho, R. K., Herrmann, B. G. and Nüsslein-Volhard, C. (1992). The protein product of the zebrafish homologue of the mouse *T* gene is expressed in nuclei of the germ ring and the notochord of the early embryo. *Development* **116**, 1021-1032.
- Schulte-Merker, S., van Eeden, F., Halpern, M. E., Kimmel, C. B., Nüsslein-Volhard, C. (1994). *No tail (ntl)* is the zebrafish homologue of the mouse *T* (*Brachyury*) gene. *Development* **120**, 1009-1015.
- Sitnikova, T., Rzhetsky, A. and Nei, M. (1995). Interior-branch and bootstrap tests of phylogenetic trees. *Molec. Biol. Evol.* **12**, 319-333.
- Smith, J. C., Price, B. M. J., Green, J. B. A., Weigel, D. and Herrmann, B. G. (1991). Expression of a *Xenopus* homolog of *Brachyury* (*T*) is an immediate-early response to mesoderm induction. *Cell* **67**, 79-87.
- Wada, H. and Satoh, N. (1993). Details of the evolutionary history from invertebrates, as deduced from the sequences of 18S rDNA. *Proc. Natl. Acad. Sci. USA* **91**, 1801-1804.
- Wilkinson, D. G., Bhatt, S. and Herrmann, B. G. (1990). Expression pattern of the mouse *T* gene and its role in mesoderm formation. *Nature* **343**, 657-659.
- Wiley, A. (1894). *Amphioxus and the ancestry of the vertebrates*. New York: MacMillan.
- Yasuo, H. and Satoh, N. (1993). Function of the vertebrate *T* gene. *Nature* **364**, 582-583.
- Yasuo, H. and Satoh, N. (1994). An ascidian homolog of the mouse *Brachyury* (*T*) gene is expressed exclusively in notochord cells at the fate restricted stage. *Dev. Growth Differ.* **36**, 9-18.