ELSEVIER

Gene 287 (2002) 121–128

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

www.elsevier.com/locate/gene

Gene duplications in the prototypical cephalochordate amphioxus ☆

Carolina Minguillón, David E.K. Ferrier[1], Cristina Cebrián[2], Jordi Garcia-Fernàndez*

Departament de Genètica, Facultat de Biologia, Universitat de Barcelona., Avenida Diagonal 645, 08028 Barcelona, Spain

Received 20 June 2001; received in revised form 5 November 2001; accepted 14 November 2001
Received by R. Di Lauro

Abstract

The new discipline of Evolutionary Developmental Biology (Evo-Devo) is facing the fascinating paradox of explaining morphological evolution using conserved pieces or genes to build divergent animals. The cephalochordate amphioxus is the closest living relative to the vertebrates, with a simple, chordate body plan, and a genome directly descended from the ancestor prior to the genome-wide duplications that occurred close to the origin of vertebrates. Amphioxus morphology may have remained relatively invariant since the divergence from the vertebrate lineage, but the amphioxus genome has not escaped evolution. We report the isolation of a second Emx gene (*AmphiEmxB*) arising from an independent duplication in the amphioxus genome. We also argue that a tandem duplication probably occurred in the Posterior part of the Hox cluster in amphioxus, giving rise to *AmphiHox14*, and discuss the structure of the chordate and vertebrate ancestral clusters. Also, a tandem duplication of Evx in the amphioxus lineage produced a prototypical Evx gene (*AmphiEvxA*) and a divergent gene (*AmphiEvxB*), no longer involved in typical Evx functions. These examples of specific gene duplications in amphioxus, and other previously reported duplications summarized here, emphasize the fact that amphioxus is not the ancestor of the vertebrates but 'only' the closest living relative to the ancestor, with a mix of prototypical and amphioxus-specific features in its genome.

Keywords: Hox; Evx; Emx; *Branchiostoma*; Genome duplication; Evolutionary Developmental Biology; Evo-Devo

1. Introduction

Evolutionary biology and developmental biology are finally being married after nearly 150 years of stormy courtship. After an initial cosy relationship in the late 19th and early 20th centuries, they went their separate ways, due to differences of opinion and interests. In the late 20th century the impact of molecular biology has facilitated a reconciliation. The result of this reconciliation has been the rejuvenation of the discipline of Evolutionary Developmental Biology, or Evo-Devo (Raff, 2000). Evo-Devo is concerned with how developmental processes themselves change and evolve. The rationale being that evolution is change in morphology, morphology depends on embryonic development, and development depends on developmental genes, gene networks, and modules. Thus, evolution can be seen as the result of adaptive variation in development (Von Dassow and Munro, 1999). Experimental work in Evo-Devo has produced a major generalization: major genes, gene networks and modules are conserved amongst most animal phyla. This universality of developmental genes has lead to the exciting paradox of Evo-Devo: how can one evolve with conserved genes and modules, if evolution is basically change? Subtle changes in cis-regulatory regions and in hierarchies and connections within and between networks may give clues to the paradox (see Salazar-Ciudad et al., 2001 for a theoretical view). A further mechanism of genetic change that can enable evolution without necessarily compromising previous developmental functions is gene duplication.

2. Gene duplication, amphioxus and the origin of vertebrates

The late Susumo Ohno's heavily cited book, *Evolution By*

Abbreviations: Myr, million years; cDNA, DNA complementary to RNA; TF, transcription factor; STR, structural molecule; SIG, signalling molecule; ENZ, enzyme; bHLH, basic helix–loop–helix; Tyr K, tyrosine kinase; PCR, polymerase chain reaction; PTP, protein tyrosine phosphatase; RDH; retinol dehydrogenase; m, mouse; zf, zebrafish; h, human; Dm, *Drosophila melanogaster*; Ce, *Caenorhabditis elegans*; X, *Xenopus*; Cn, cnidarian; *B.*, *Branchiostoma*

☆ The Emx cDNA sequence has been deposited in GenBank under accession number AY040834.

* Corresponding author. Tel.: +34-93-402-1502; fax: +34-93-411-0969.
E-mail address: jgarcia@bio.ub.es (J. Garcia-Fernàndez).

[1] Present address: School of Animal and Microbial Sciences, University of Reading, Whiteknights, Reading RG6 6AJ, UK.

[2] Present address: Centre d'Investigacions en Bioquímica i Biologia Molecular, Hospital de la Vall d'Hebrón, Pg. Vall d'Hebrón 119–129, 08035 Barcelona, Spain.

0378-1119/02/$ - see front matter
PII: S0378-1119(01)00828-9

Gene Duplication (Ohno, 1970) was a visionary anticipation of the importance of gene duplication in evolution. After duplication, one of the gene copies may be less constrained, allowing innovation and diversification. Ohno also proposed that two rounds of full genome duplication occurred at the origin of vertebrates.

Lundin reported widespread paralogy regions in the human genome, and interpreted them as remnants of duplications of large genomic regions, or ancient polyploidy (Lundin, 1993). This, together with the discovery that cephalochordates, the sister group of the vertebrates, had a single Hox cluster, prototypical to the four Hox clusters of mammals (Garcia-Fernàndez and Holland, 1994), strengthened the hypothesis that two full polyploidizations occurred at the origin of vertebrates (Holland et al., 1994). These genome duplications intriguingly correlate with the origin of vertebrates and their many innovations.

The nature and extent of these duplications have been debated in recent years, but it is now beyond doubt that widespread duplications did occur in the vertebrate lineage, most probably via polyploidizations. The precise locations and extent of these events within a vertebrate phylogeny is now becoming the focus of the debate (Holland, 1999). Amphioxus thus occupies a key phylogenetic position, being the sister group to the vertebrates, but moreover being the lineage closest to these genome-wide duplications before they occurred. Morphologically amphioxus is considered to be not excessively derived from the cephalochordate-vertebrate common ancestor. Cephalochordates possess a hollow dorsal neural tube, a notochord, and lateral muscle blocks, as is typical for chordates (urochordates, cephalochordates and vertebrates), but lack the vertebrate innovations. The origin of vertebrates involved the origin of new cell types and organs, and increased body organization complexity. Vertebrate innovations include neural crest cells, neurogenic placodes, an elaborated and morphologically segmented brain, paired sense organs and endoskeleton (Shimeld and Holland, 2000).

The phylogenetic position of amphioxus as the sister group of vertebrates, and the simple and prototypical body plan of amphioxus compared to vertebrates, has led many to examine their favourite gene (or gene family) in this animal in order to ascertain the ancestral state for vertebrates. This has proven to be a fruitful venture (see Holland, 1999 for a catalogue), but it must always been remembered that amphioxus itself has been evolving from the common ancestor of cephalochordates and vertebrates for over 550 Myr. We will present three examples of the amphioxus genome reflecting both its closeness to the pre-duplication vertebrate ancestor, but also its derived condition. These amphioxus novelties are gene duplications specific to the amphioxus lineage.

3. Posterior Hox genes: a chordate Hox14

Hox genes pattern the anterior–posterior axis of animals and are organized into clusters in the genome. Within a cluster the genes are organized such that they are colinear with their domains of expression and function: the Anterior patterning genes are at one end of the cluster (the 3′ end), the Medial genes are in the centre of the cluster, and the Posterior-patterning genes at the other end (5′) (Fig. 1). This phenomenon of colinearity is seen from flies to humans, including amphioxus (Graham et al., 1989; Wada et al., 1999). The multiple Hox clusters of vertebrates, e.g. four

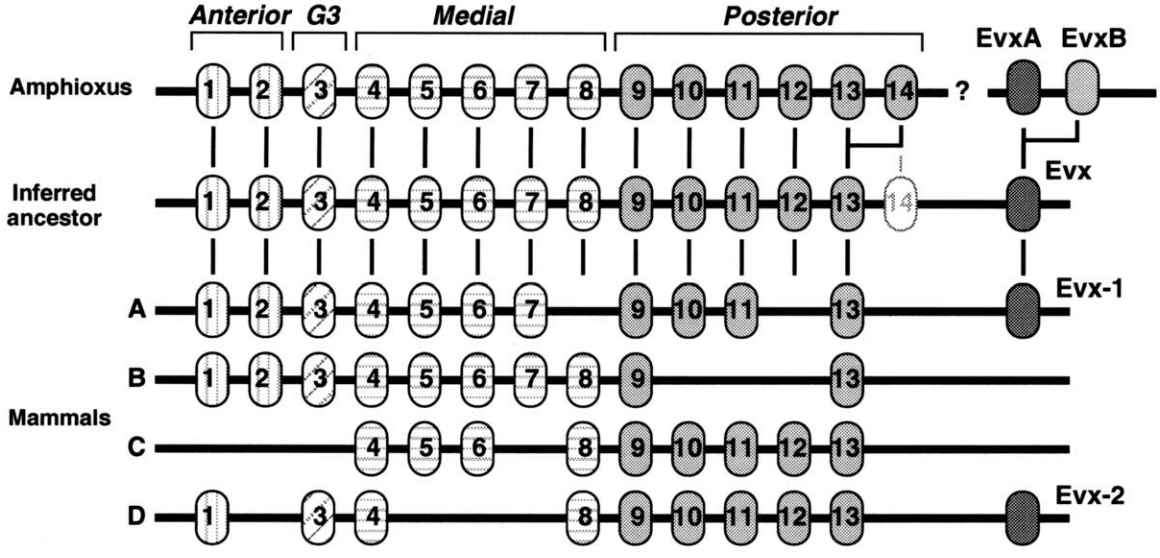

Fig. 1. Organization of the Hox cluster of amphioxus and mammals, and the presumed cluster of the last common ancestor of cephalochordates and vertebrates. Hox paralogous groups are classified as anterior (vertical hatching), group 3 (diagonal hatching), medial (horizontal hatching) and posterior (grey filling) classes. Posterior Hox genes and Evx genes underwent tandem duplications in the amphioxus lineage. Alternatively, the ancestral cluster may had a cluster with 14 Hox genes, as discussed in Section 3.

in mammals and seven in zebrafish (Amores et al., 1998), contain 13 different types of Hox gene (paralogy groups), divisible into the Anterior, Group 3, Medial and Posterior classes (Fig. 1). The Posterior class consists of Hox groups 9–13. Amphioxus, consistent with its lineage diverging from that of vertebrates before the genome-wide duplications at vertebrate origins, has a single Hox gene cluster. In many respects the amphioxus cluster can be considered to be prototypical relative to those of vertebrates. It has no gaps, and contains a pro-orthologue for each of the vertebrate paralogy groups, at least up to the Posterior region of the cluster (Garcia-Fernàndez and Holland, 1994). In the Posterior region the situation is not so straightforward. Amphioxus may not have genes missing, but it does contain an extra gene (Hox 14), and the distinctive evolutionary rates of these Posterior genes reveal the phenomenon of Deuterostome Posterior Flexibility, whereby the rate of evolution of the deuterostome Posterior Hox genes has increased relative to the more anterior Hox genes, and the protostome Posterior Hox genes (Ferrier et al., 2000). This obscures the relationships between the deuterostome Posterior Hox genes.

The 14th amphioxus Hox gene was found during a genomic walk from the Posterior/5′ end of the Hox cluster. Its discovery gives amphioxus the (dubious) honour of possessing the most gene-rich Hox cluster to date. Did *AmphiHox14* arise from a tandem duplication only within the amphioxus lineage, or does it indicate that the vertebrate ancestor also had a fourteenth Hox gene, which has subsequently been lost in at least some vertebrates (only humans, mice, and zebrafish have been examined in sufficient detail to unambiguously determine the absence of a Hox14)? Molecular phylogenetics has not clearly resolved this issue (Ferrier et al., 2000). There is a faint suggestion from some phylogenetic trees that AmphiHox13 and 14 form a sister group, which is consistent with the fourteenth Hox gene being an amphioxus-specific feature. This grouping however is not stable amongst all trees, or with various different molecular phylogenetic approaches. Furthermore since AmphiHox13 and 14 are two of the most divergent sequences in these trees there is the possibility that their grouping is artefactual, due to their longer branches attracting each other.

As for the vertebrate ancestor, we favour the hypothesis that it contained 13 Hox genes orthologous to *AmphiHox 1–13*, with *AmphiHox14* being an amphioxus-specific duplication, and the orthology relationships between the amphioxus and vertebrate genes being obscured by Deuterostome Posterior Flexibility. The possibility that it contained a Hox14 gene, however, cannot yet be completely excluded.

Resolution of the ancestral state would only conclusively come with the discovery of a fourteenth vertebrate Hox gene. However it would also be intriguing to know precisely how many Posterior genes other deuterostome taxa have. This would require a thorough genomic walk from the end of the Hox cluster in the respective animal. The divergent nature of the Posterior-most Hox genes means that degenerate PCR approaches alone are not sufficient. In the sea urchin, genomic walking has been done as far as three Posterior Hox genes (Martinez et al., 1999). There are probably more further 5′, as up to five Posterior Hox genes have been cloned (by PCR) from other species of echinoderm (Méndez et al., 2000; Mito and Endo, 2000). The sea urchin genome project will tidy this up. However, it must also be borne in mind that this echinoderm also exhibits an example of Hox gene loss. It lacks either a Hox 4 or 5 gene (Martinez et al., 1999). Another animal in which a genome project might help is the urochordate *Ciona intestinalis*. Three Posterior Hox genes have so far been isolated in *C. intestinalis* (Di Gregorio et al., 1995), but considerable effort with genomic walking has so far been unsuccessful in linking all of the ascidian Hox genes together (Di Lauro and colleagues, personal communication). Of course, amphioxus may have still more Posterior Hox genes. So far the walk has been extended 80 kb beyond *AmphiHox14*, without uncovering any more Hox genes (C. Minguillón and J. Garcia-Fernàndez, unpublished). However, linkage to the Evx genes has not yet been established as a means to confirm that the end of the cluster has been reached (see Section 4).

4. Prototypical and divergent Evx genes after tandem gene duplication in the amphioxus lineage

Evx homeobox genes are present from cnidarians to vertebrates. They seem to have a basic function in patterning the posterior part of the embryo at the onset of gastrulation in all bilaterians, and also function in the central nervous system. Lineage-specific functions have also been acquired: *even-skipped* is a Pair-Rule gene in *Drosophila melanogaster* segmentation, whilst in vertebrates Evx genes function in development of the tail and limb buds, and are expressed around the Midbrain/Hindbrain Boundary (MHB) (Ferrier et al., 2001 and references therein). The genes are linked to the posterior ends of Hox clusters in vertebrates, and an Evx gene has been reported closely linked to a Hox-like gene in corals, a characteristic probably shared with other cnidarians (Ferrier and Holland, 2001). We isolated Evx representative(s) from amphioxus and traced the evolutionary history of Evx in chordates (Ferrier et al., 2001). We found two linked Evx genes in amphioxus (*AmphiEvxA* and *AmphiEvxB*), only 35 kb apart (schematized in Fig. 1). *AmphiEvxA* and *AmphiEvxB* probably arose via tandem gene duplication. The homeodomain of AmphiEvxA is clearly closer to vertebrate Evx1 and Evx2 than the homeodomain of AmphiEvxB, whose similarity to the vertebrate proteins is distinctly lower (Table 1). This suggests that *AmphiEvxA* is prototypical while *AmphiEvxB* is a divergent member of the family. Molecular phylogenetic analyses of the Evx genes confirmed that *AmphiEvxA* lies at the base of vertebrate genes, while *AmphiEvxB* is a divergent Evx.

AmphiEvxA is expressed during three developmental

Table 1
Percentage of identities to AmphiEvxA and B homeodomains and six amino acid flanks[a]

		% Identities to	
		AmphiEvxA	AmphiEvxB
Vertebrate Evx1	m Evx 1	86	75
	X Xhox3	86	75
	zf Evx 1	85	74
Vertebrate Evx2	m Evx 2	88	76
	zf Evx 2	85	75
Invertebrate Eve	Dm Eve	82	72
	Ce Vab7	69	64
	coral EveC	82[b]	75[b]

[a] Species abbreviations: Ce, *Caenorhabditis elegans*; Dm, *Drosophila melanogaster*; m, mouse; X, *Xenopus*; zf, zebrafish.
[b] No flanking regions of coral EveC are available. Percentages refer only to the homeodomain.

processes: gastrulation, neurogenesis, and tail bud development. During gastrulation, *AmphiEvxA* is expressed in the ectoderm in a ventral-posterior domain (Fig. 2A–C). At neurula and early larval stages *AmphiEvxA* is expressed at the posterior of the embryo, in all three germ layers (data not shown), with clear expression in the central nerve cord. The pairs of cells in the CNS are seen at intervals equivalent to the length of one somite, but never more anterior than the level of the somite 4/5 boundary, with the pair at the level of the somite 5/6 boundary being most persistent (Fig. 2D). Finally, in larval stages *AmphiEvxA* is expressed in the tail bud, in the posterior-most neural tube and endoderm (Fig. 2E). *AmphiEvxB* expression is not detectable until hatching, when a consistent stain all over the ectoderm appears and is maintained until larval stages (Fig. 2F).

Both sequence comparisons and expression patterns clearly indicate that *AmphiEvxA* is a prototypical chordate Evx gene. The basal chordate role of Evx genes was the pan-bilaterian function in gastrulation and neurogenesis, plus a probable pan-chordate role in tail bud development. The expression of Evx around the MHB is probably a vertebrate innovation, perhaps linked to the evolution of the MHB

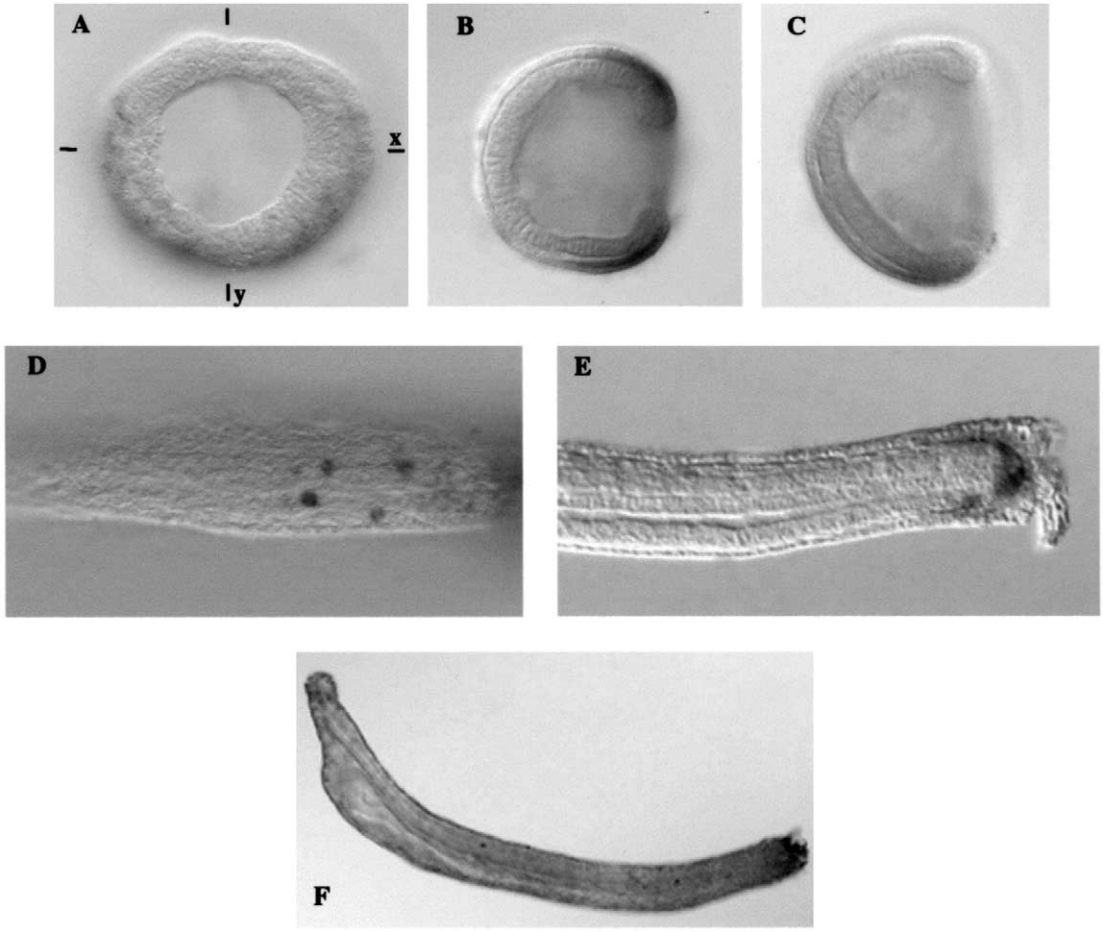

Fig. 2. Expression of *amphiEvxA* and *AmphiEvxB*. Whole-mount in situ hybridizations of gastrula (A–C), 22 h embryo (D) and 48 h (E) larva. Anterior is to the left in all panels except (A). (A) is a posterior (blastopore) view, with dorsal at the top. Panels (B,C) are optical sections of (A) in the *x* and *y* planes, respectively. Expression in the gastrula is in the ventral–posterior ectoderm. (D) Dorsal view showing pairs of cells staining in the central nervous system, the anterior-most of which is level with the somite 5/6 boundary. (E) Lateral view showing tail-bud expression. The ventral gap in the crescent of expression is where the anus breaks through. (F) Expression of *AmphiEvxB* in larvae is seen all over the ectoderm.

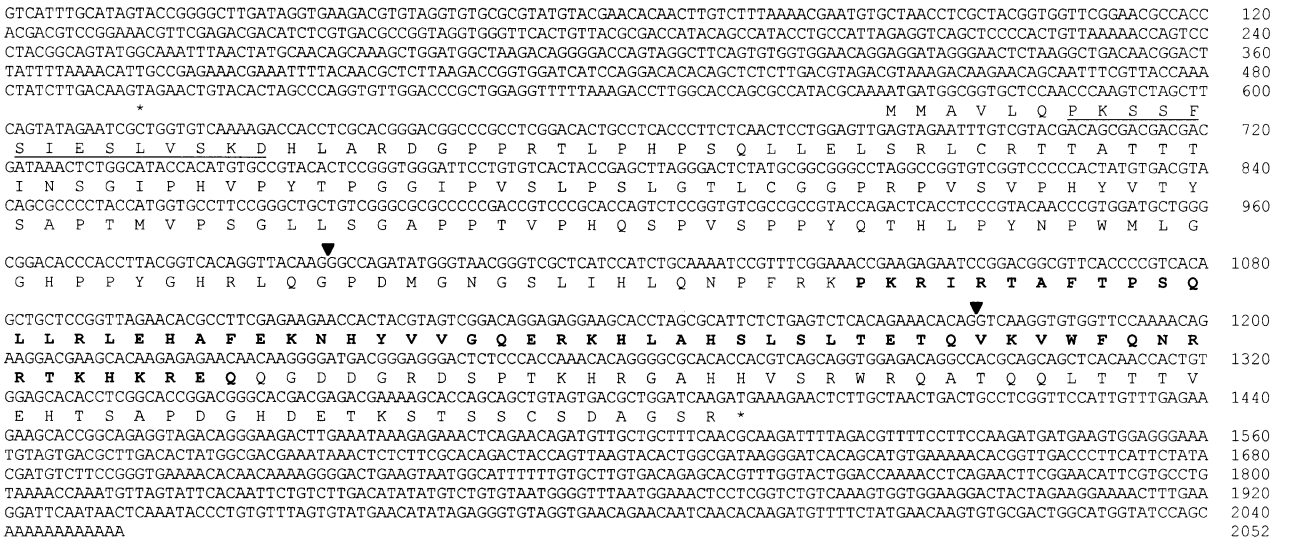

Fig. 3. cDNA nucleotide sequence and predicted amino acid sequence of *AmphiEmxB*. The homeodomain residues are shown in bold, and the conserved Emx peptide domain near the N-terminus is underlined. Asterisks indicate the first 5′ and 3′ in-frame stop codons. Deduced intron positions are indicated with black triangles. The Emx cDNA sequence has been deposited in GenBank under accession number AY040834.

organizer activity in vertebrates. Appendages, and the function of Evx in their development, evolved within the vertebrates. In summary, an ancestral Evx gene was duplicated in the amphioxus genome, giving rise to a prototypical Evx gene (*AmphiEvxA*) and a fast evolving gene (*AmphiEvxB*), that is no longer involved in typical Evx functions.

5. Duplication in amphioxus of an Emx-class homeobox gene

Emx genes are expressed and function in anterior cephalic domains during embryonic development of vertebrates and *D. melanogaster* (Panesse et al., 1998 and references therein). Orthologues have also been found in *Caenorhabditis elegans*, the hydrozoan *Hydractinia symbiolongicarpus*, and recently in the cephalochordate amphioxus (*AmphiEmxA*, Williams and Holland, 2000).

Vertebrates possess two Emx genes (Emx1 and 2) arising from the genome-wide duplications at the origin of vertebrates. Emx genes were also independently duplicated in the lineage leading to higher insects, giving raise to *empty-spiracles* (*ems*) and *E5* genes in *D. melanogaster* (Walldorf and Gehring, 1992).

Here we report the cloning of a second AmphiEmx gene, *AmphiEmxB*. The cDNA sequence and predicted amino acid sequence is shown in Fig. 3. Intron positions were deduced by comparison of the cDNA sequence with genomic clones. *AmphiEmxB* contains two introns, which are shared with *AmphiEmxA*. The first one (also present in mouse and human Emx genes) is located 5′ of the homeobox, and separates it from a region coding for a partially conserved hexapeptide sequence found in several classes of homeobox-containing genes. The second intron is within the homeobox, between homeodomain residues 44 and 45, and is also in the same position in several vertebrate Emx genes, *D. melanogaster E5*, and *C. elegans ceh-12*.

Homeodomain similarities with other Emx proteins range from 80 to 90%, with the exception of the divergent cnidarian protein (only 68% similar to AmphiEmxB) (Fig. 4).

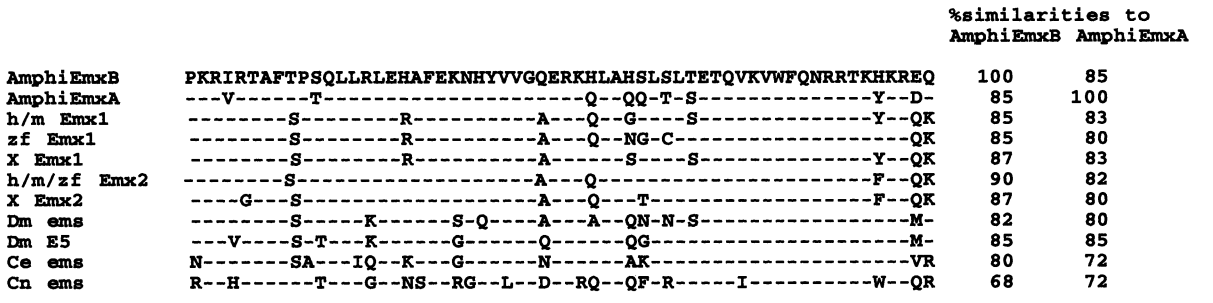

Fig. 4. Alignment of the homeodomain sequence of AmphiEmxA and B with vertebrate, fly, nematode and cnidarian homeodomain sequences. Dashes indicate identities to AmphiEmxB. The percentages of identity to the AmphiEmxB or AmphiEmxA homeodomains are shown. Species abbreviations are: Ce, *Caenorhabditis elegans*; Cn, cnidarian (*Hydractinia symbiolongicarpus*); Dm, *Drosophila melanogaster*; h, human; m, mouse; X, *Xenopus*; zf, zebrafish.

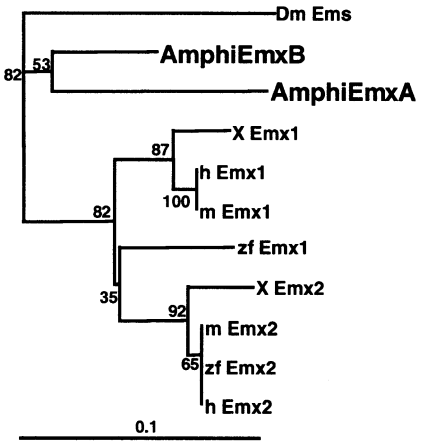

Fig. 5. Neighbour-joining phylogenetic tree with homeodomain sequences of amphioxus AmphiEmxB, AmphiEmxA, vertebrate Emx proteins, and *Drosophila* Ems. The numbers at the nodes are scores from 1000 bootstrap resamplings of the data. The abbreviations used are the same as in Fig. 4.

AmphiEmxA and B are 85% similar in the homeodomain, with only scattered similarities outside: two small domains can be recognized and aligned, the hexapeptide just upstream of the homeodomain, and the Emx peptide domain (Williams and Holland, 2000) at the amino-terminus of the protein, where they are 78.6% similar (underlined in Fig. 3).

A Neighbour-joining tree using 1000 bootstrap replicates, with AmphiEmxA and B and representative vertebrate Emx homeodomains (incomplete sequences have been omitted), and rooted with *D. melanogaster* Ems, shows that the amphioxus genes group together, and form a sister group to both vertebrate Emx1 and 2 classes (Fig. 5). This topology is consistent with vertebrate and amphioxus Emx genes having arisen by independent duplications in both lineages.

A closer look at the sequence similarities (Fig. 4) and branch lengths (Fig. 5) suggests that AmphiEmxB is slightly more related to vertebrate genes than AmphiEmxA, and that its rate of evolution is lower than that of AmphiEmxA. With no expression data available, it is difficult to speculate on which copy, if any, or both, have retained an ancestral role in patterning the most anterior part of the neural tube, a presumed ancestral function for this gene family in chordates. However, based on the above reasoning, we speculate is that AmphiEmxB may be the one retaining ancestral roles.

6. Further gene duplications in the amphioxus lineage

Table 2 summarizes published data on amphioxus-specific gene duplications. A further report of an extra actin gene from *B. belcheri* expressed in the notochord (Suzuki and Satoh, 2000) is not included as it is not clear whether it originated from a duplication of a cytoplasmic or a muscle actin gene. Where there is information about the expression of duplicated genes, the behaviour of the duplicates is diverse. For brachyury, duplicates may have indistinguishable patterns of expression, or the expression reported may be the summation of both genes (Holland et al., 1995). For HNF-3 and muscle actin genes the differences of expression are most probably quantitative (Shimeld, 1997; Kusakabe et al., 1997). *AmphiEvx A* and *B*, however, have entirely unrelated patterns of expression. It is noteworthy that a single RDH gene has been found in *B. lanceolatum*, whereas two

Table 2
List of reported gene duplications specific to the amphioxus lineage[a]

Genes	Type[b]	Species	Number[c]	Expression[d]	Reference
Brachyury	TF	*B. floridae*	2	Unknown	Holland et al., 1995
Myogenic bHLH	TF	*B. floridae*	2	Unknown	Araki et al., 1996
HNF-3	TF	*B. floridae*	2	Very similar*	Shimeld, 1997
Muscle actin	STR	*B. floridae*	2	Very similar*	Kusakabe et al., 1997
Cholinesterase	ENZ	*B. lanceolatum*	2	Unknown	Sutherland et al., 1997
Tyr K (Ephs)	SIG	*B. belcheri*	2	Unknown	Suga et al., 1999
Tyr K (src)	SIG	*B. belcheri*	2	Unknown	Suga et al., 1999
Hox 13/14	TF	*B. floridae*	2	Unknown	Ferrier et al., 2000
Calmodulin	SIG	*B. lanceolatum*	2	Unknown	Karabinos and Bhattacharya, 2000
Calmodulin-like	SIG	*B. lanceolatum*	2	Unknown	Karabinos and Bhattacharya, 2000
		B. floridae	3	Unknown	Karabinos and Bhattacharya, 2000
PTP (PTPR4)	SIG	*B. belcheri*	3	Unknown	Ono-Koyanagi et al., 2000
Evx	TF	*B. floridae*	2	Distinct	Ferrier et al., 2001
RDH	ENZ	*B. floridae*	2	Unknown	Dalfó et al., 2001
Emx	TF	*B. floridae*	2	Unknown	Present work

[a] bHLH, basic helix–loop–helix; Tyr K, tyrosine kinase; PTP, protein tyrosine phosphatases; RDH, retinol dehydrogenase; TF, transcription factor; STR, structural molecule; SIG, signalling molecule; ENZ, enzyme.
[b] Type of molecule encoded.
[c] Number of genes for the given family reported in amphioxus. In all cases, a single gene is expected in the last common ancestor of cephalochordates and vertebrates.
[d] When expression data are available, the relation between the duplicates is indicated. *Quantitative differences.

are present in *B. floridae* (Dalfó et al., 2001). The authors convincingly demonstrate that there is only a single copy in *B. lanceolatum* and that the two copies of *B. floridae* arose after an independent duplication. Taking into account that both species may have diverged as long as 190 Myr ago (Cañestro et al., 2002), it is perhaps not surprising to find species-specific gene duplications within these anciently separated *Branchiostoma* lineages.

7. Discussion

The rate of gene duplication and subsequent gene loss has been a matter of discussion for many years (see Meyer and Schartl, 1999 and references therein), and can now be addressed with the aid of complete genome sequences (e.g. Wolfe, 2001 and references therein). It has been proposed that gene duplication may occur at a high rate, over 0.01 per gene per million years, with gene loss after duplication being more common than maintaining a duplicate by functional divergence (Lynch and Conery, 2000). The origin of vertebrates may be an exception to the high rate of gene loss after gene duplication. It is clear that numerous gene families expanded by gene duplication on the vertebrate stem lineage. Although it is difficult to demonstrate experimentally, it is tempting to speculate that vertebrate innovations were possible due to the availability of extra genetic raw material after polyploidization or widespread gene duplications.

Amphioxus is in the privileged position of being morphologically not very derived, since the divergence from the vertebrate lineage. The invariance of amphioxus morphology is realized by consideration of the estimates of sequence divergence between three amphioxus species: *B. floridae* (West Atlantic Ocean), *B. lanceolatum* (East Atlantic and Mediterranean), and *B. belcheri* (Pacific Ocean). The estimates range from 110 to 190 Myr (Nohara et al., 2001; Cañestro et al., 2002), for species of the same genus, estimates that are much higher than the divergence between primate and rodent genes (different orders). This implies than the morphological uniformity among living lancelets is attributable to morphological stability rather than to recent speciation, and by extension, amphioxus morphology has not evolved much from the cephalochordate-vertebrate ancestor.

However, the amphioxus genome has been evolving separately from that of vertebrates for over 550 Myr, and is thus a remnant of the ancestor's plus the particular history of the amphioxus lineage, with gene duplications, functional divergence and gene losses. We have presented three examples of duplications: the probable expansion at the posterior end of the amphioxus Hox cluster, the tandem duplication of Evx, and duplication of an Emx class gene.

Amphioxus morphology may have remained astonishingly invariant since the origin of vertebrates, but the amphioxus genome has not escaped evolution. Gene duplication events are not so rare, as perhaps should be expected in any lineage with a long independent evolutionary history. Amphioxus has undergone duplication of developmentally important genes (Emx). Such duplications raise questions about the chordate and vertebrate ancestors; for example, Hox14 and how many Posterior Hox genes were there? Also duplication of genes can lead to very divergent members of their respective classes (*AmphiEvxB*), which may have unusual, atypical expression for the class, if indeed they are not recently formed pseudogenes. Gene duplications in amphioxus are not restricted to homeobox genes (see Table 2), and so when analysing an amphioxus gene to trace the history of a given gene family in chordates, special attention should be paid to the fact that amphioxus is not the ancestor of vertebrates, but 'only' the closest living relative to the vertebrate ancestor, an invaluable position to help understand the origin of vertebrates.

Acknowledgements

We are indebted to Peter Holland for discussions, and in whose laboratory some of the work reported here was conducted, Cristian Cañestro and Iñaki Ruiz for technical advice, and Linda and Nic Holland for help in amphioxus collection. We also acknowledge Ricard Albalat, Gemma Marfany, Seb Shimeld and Becky Furlong for discussions and comments on the manuscript. Research was supported by a grant from DGESIC PB98-1261-C02-02 (Ministerio de Educación y Cultura, Spain). D.E.K.F. was supported by a Marie Curie TMR postdoctoral fellowship, C.M. holds a CIRIT (Generalitat de Catalunya) predoctoral fellowship. Discussions and collaborations between Spanish and English laboratories have been facilitated by grants from Acciones Integradas of the British Council/Ministerio de Ciencia y Tecnología.

References

Amores, A., Force, A., Yan, Y.L., Joly, L., Amemiya, C., Fritz, A., Ho, R.K., Langeland, J., Prince, V., Wang, Y.L., Westerfield, M., Ekker, M., Postlethwait, J.H., 1998. Zebrafish *hox* clusters and vertebrate genome evolution. Science 282, 1711–1714.

Araki, I., Terazawa, K., Satoh, N., 1996. Duplication of an amphioxus myogenic bHLH gene is independent of vertebrate myogenic bHLH gene duplication. Gene 171, 231–236.

Cañestro, C., Albalat, R., Hjelmqvist, L., Godoy, L., Jörnvall, H., Gonzàlez-Duarte, R., 2002. Ascidian and amphioxus *Adh* genes reveal functional and molecular evolution of the ADH family expansion during vertebrate evolution. J. Mol. Evol. 54, 81–89.

Dalfó, D., Cañestro, C., Albalat, R., Gonzàlez-Duarte, R., 2001. Characterization of a microsomal retinol dehydrogenase gene from amphioxus: retinoid metabolism before vertebrates. Chem. Biol. Interact. 130–132, 359–370.

Di Gregorio, A., Spagnuolo, A., Ristoratore, F., Pischetola, M., Aniello, F., Branno, M., Cariello, L., Di Lauro, R., 1995. Cloning of ascidian homeobox genes provides evidence for a primordial chordate cluster. Gene 156, 253–257.

Ferrier, D.E.K., Holland, P.W.H., 2001. Ancient origin of the Hox gene cluster. Nat. Rev. Genet. 2, 33–38.

Ferrier, D.E.K., Minguillón, C., Holland, P.W.H., Garcia-Fernàndez, J., 2000. The amphioxus Hox cluster: deuterostome posterior flexibility and Hox14. Evol. Dev. 2, 284–293.

Ferrier, D.E.K., Minguillón, C., Cebrián, C., Garcia-Fernàndez, J., 2001. Amphioxus Evx genes: implications for the evolution of the Midbrain-Hindbrain Boundary and the chordate tailbud. Dev. Biol. 237, 270–281.

Garcia-Fernàndez, J., Holland, P.W.H., 1994. Archetypal organization of the amphioxus *Hox* gene cluster. Nature 370, 563–566.

Graham, A., Papalopulu, N., Krumlauf, R., 1989. The murine and *Drosophila* homeobox gene complexes have common features of organisation and expression. Cell 57, 367–378.

Holland, P.W.H., 1999. Gene duplication: past, present and future. Semin. Cell Dev. Biol. 10, 541–547.

Holland, P.W.H., Garcia-Fernandez, J., Williams, N.A., Sidow, A., 1994. Gene duplications and the origins of vertebrate development. Development Suppl., 125–133.

Holland, P.W.H., Koschorz, B., Holland, L.Z., Herrmann, B.G., 1995. Conservation of Brachyury (T) genes in amphioxus and vertebrates: developmental and evolutionary implications. Development 121, 4283–4291.

Karabinos, A., Bhattacharya, D., 2000. Molecular evolution of calmodulin and calmodulin-like genes in the cephalochordate *Branchiostoma*. J. Mol. Evol. 51, 141–148.

Kusakabe, R., Kusakabe, T., Satoh, N., Holland, N.D., Holland, L.Z., 1997. Differential gene espression and intracellular mRNA localization of amphioxus actin isoforms throughout development: Implications for conserved mechanisms of chordate development. Dev. Genes Evol. 207, 203–215.

Lundin, L.G., 1993. Evolution of the vertebrate genome as reflected in paralogous chromosomal regions in man and the house mouse. Genomics 16, 1–19.

Lynch, M., Conery, J.S., 2000. The evolutionary fate and consequences of duplicate genes. Science 290, 1151–1155.

Martinez, P., Rast, J.P., Arenas-Mena, C., Davidson, E.H., 1999. Organization of an echinoderm *Hox* gene cluster. Proc. Natl. Acad. Sci. USA 96, 1469–1474.

Méndez, A.T., Roig-López, J.L., Santiago, P., Santiago, C., GarcÍa-Arrarás, J.E., 2000. Identification of Hox gene sequences in the Sea Cucumber *Holothuria glaberrima* Selenka (Holothuroidea: Echinodermata). Mar. Biotech. 2, 231–240.

Meyer, A., Schartl, M., 1999. Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. Curr. Opin. Cell Biol. 11, 699–704.

Mito, T., Endo, K., 2000. PCR survey of Hox genes in the Crinoid and Ophiuroid: evidence for Anterior conservation and Posterior expansion in the echinoderm Hox gene cluster. Mol. Phyl. Evol. 14, 375–388.

Nohara, M., Nishida, M., Yamaguchi, T., Tomiyama, M., Zhang, P., Yokoyama, H., Manthacitra, V., Nishikawa, T., 2001. Mitochondrial DNA variation and genetic relationships of *Branchiostoma* species from the Pacific and Atlantic Oceans. Zool. Sci. 18, 453–454.

Ohno, S., 1970. Evolution by Gene Duplication, George Allen and Unwin, London.

Ono-Koyanagi, K., Suga, H., Katoh, K., Miyata, T., 2000. Protein Tyrosine Phosphatases from amphioxus, hagfish and ray: Divergence of tissue-specific isoform genes in the early evolution of vertebrates. J. Mol. Evol. 50, 302–311.

Panesse, M., Lupo, G., Kablar, B., Boncinelli, E., Barsacchi, G., Vignalli, R., 1998. The *Xenopus* Emx genes identify presumptive dorsal telencephalon and are induced by head organizer signals. Mech. Dev. 73, 73–83.

Raff, R.A., 2000. Evo-devo: the evolution of a new discipline. Nat. Rev. Genet. 1, 74–79.

Salazar-Ciudad, I., Newman, S.A., Sole, R.V., 2001. Phenotypic and dynamical transitions in model genetic networks. I. Emergence of patterns and genotype-phenotype relationships. Evol. Dev. 3, 84–94.

Shimeld, S.M., 1997. Characterisation of amphioxus HNF-3 genes: conserved expression in the notochord and floor plate. Dev. Biol. 183, 74–85.

Shimeld, S.M., Holland, P.W., 2000. Vertebrate innovations. Proc. Natl. Acad. Sci. USA 97, 4449–4452.

Suga, H., Hoshiyama, D., Kuraku, S., Katoh, K., Kubokawa, K., Miyata, T., 1999. Protein tyrosine kinase cDNAs from amphioxus, hagfish, and lamprey: isoform duplications around the divergence of cyclostomes and gnathostomes. J. Mol. Evol. 49, 601–608.

Sutherland, D., McClellan, J.S., Milner, D., Soong, W., Axon, N., Sanders, M., Hester, A., Kao, Y.H., Poczatek, T., Routt, S., Pezzementi, L., 1997. Two cholinesterase activities and genes are present in amphioxus. J. Exp. Zool. 277, 213–229.

Suzuki, M.M., Satoh, N., 2000. Genes expressed in the amphioxus notochord revealed by EST analysis. Dev. Biol. 224, 168–177.

Von Dassow, G., Munro, E., 1999. Modularity in animal development and evolution: elements of a conceptual framework for EvoDevo. J. Exp. Zool. 285, 307–325.

Wada, H., Garcia-Fernàndez, J., Holland, P.W., 1999. Colinear and segmental expression of amphioxus Hox genes. Dev. Biol. 213, 131–141.

Walldorf, U., Gehring, W.J., 1992. *Empty spiracles*, a gap gene containing a homeobox involved in *Drosophila* head development. EMBO J. 11, 2247–2259.

Williams, N.A., Holland, P.W., 2000. An Amphioxus Emx homeobox gene reveals duplication during vertebrate evolution. Mol. Biol. Evol. 17, 1520–1528.

Wolfe, K.H., 2001. Yesterday's polyploids and the mystery of diploidization. Nat. Rev. Genet. 2, 333–341.