View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

Available online at www.sciencedirect.com

ScienceDirect

www.sciencedirect.com/science/journal/16720229

GENOMICS PROTEOMICS & BIOINFORMATICS

Review

The Origin of Patterning Systems in Bilateria —Insights from the Hox and ParaHox Genes in Acoelomorpha

Eduardo Moreno[1], Jon Permanyer[1], and Pedro Martinez[1,2*]

[1]*Department of Genetics, University of Barcelona, 08028-Barcelona, Spain;*
[2]*Catalan Institute for Research and Advanced Studies (ICREA), 08010-Barcelona, Spain.*

Genomics Proteomics Bioinformatics 2011 Jun; 9(3): 65-76 DOI: 10.1016/S1672-0229(11)60010-7

Received: Oct 26, 2010; Accepted: Feb 24, 2011

Abstract

Hox and ParaHox genes constitute two families of developmental regulators that pattern the Anterior–Posterior body axis in all bilaterians. The members of these two groups of genes are usually arranged in genomic clusters and work in a coordinated fashion, both in space and in time. While the mechanistic aspects of their action are relatively well known, it is still unclear how these systems evolved. For instance, we still need a proper model of how the Hox and ParaHox clusters were assembled over time. This problem is due to the shortage of information on gene complements for many taxa (mainly basal metazoans) and the lack of a consensus phylogenetic model of animal relationships to which we can relate our new findings. Recently, several studies have shown that the Acoelomorpha most probably represent the first offshoot of the Bilateria. This finding has prompted us, and others, to study the Hox and ParaHox complements in these animals, as well as their activity during development. In this review, we analyze how the current knowledge of Hox and ParaHox genes in the Acoelomorpha is shaping our view of bilaterian evolution.

Key words: Hox–ParaHox, Acoelomorpha, Bilateria, cluster, transposon, Hox code

Introduction

The Hox and ParaHox group of regulatory genes encode for proteins involved in the regionalization of the anterior–posterior (AP) axis during the early embryonic development of bilateral animals (*1*). The protein products of these genes are characterized by the presence of a DNA binding domain known as the homeodomain fold (*2*).

The Hox–ParaHox genetic system has been widely studied in the Nephrozoa, the clade comprising the three main groups of bilateral animals: Deuterostomia, Ecdysozoa and Lophotrochozoa (*3*). Likewise, the phylogenetic sister group of the Bilateria, the diploblastic Cnidaria, has recently attracted the attention of scientists working on this issue (*4-10*). Such studies have shown deep differences between the simple and disorganized Hox–ParaHox genetic system in Cnidaria [where it is composed of only a few genes that are not involved in the establishment of the main body axis (*4*)] and the complex and well-organized system operating in Bilateria [where the number of Hox genes is higher and both gene families are, mostly, organized in conserved genomic clusters and involved in the patterning of the AP axis (*11-13*)]. These fundamental differences make it particularly difficult to

*Corresponding author.
E-mail: pedro.martinez@ub.edu

© 2011 Beijing Institute of Genomics.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

understand the evolution of the Hox–ParaHox system at the dawn of the Bilateria. In order to gain a more accurate picture of the evolution of this important genetic system, it will be essential to clarify the evolutionary transformations in the Hox–ParaHox system that occurred in the time span between the appearance of the cnidarian–bilaterian last common ancestor (C-BLCA) and of the last common ancestor of all bilaterian animals (LCBA).

Fortunately, it was recently discovered that there is a group of bilateral animals that branched before the divergence of the three above-mentioned superclades. As a consequence, they represent a key group in our understanding of the evolutionary history of the Hox–ParaHox system from the C-BLCA to the Nephrozoa last common ancestor (NLCA). This interesting group of animals is the Acoelomorpha, a group of flatworms that were traditionally considered members of the turbellarian platyhelminthes (*14*). Nowadays, however, the Acoelomorpha appear as the earliest offshoot of the Bilateria in most of the recent phylogenetic and phylogenomic analyses (*15, 16*) (**Figure 1A**), although this position is contested by others who have suggested that the Acoelomorpha are

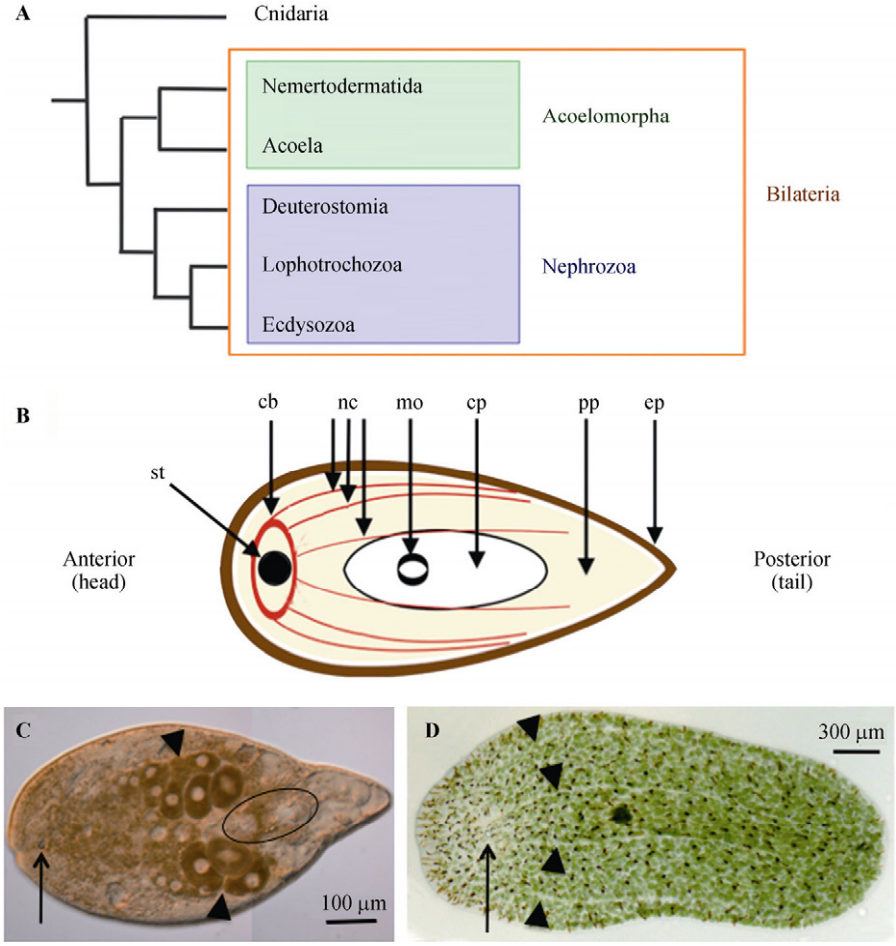

Figure 1 **A**. A consensus metazoan phylogenetic tree showing the relationships of all major clades. The Bilateria are subdivided into two major groups, the Acoelomorpha and the Nephrozoa (containing the superclades: Deuterostomia, Lophotrochozoa and Ecdysozoa). The Cnidaria is considered the sister group to the Bilateria. **B**. Diagrammatic view of the general morphology of Acoela from the dorsal side: st, statocyst (gravity receptor organ); cb, commissural brain; nc, nerve cords; mo, mouth opening; cp, central parenchyma or gut; pp, peripheral parenchyma; ep, epidermis. **C**. Adult specimen of *Isodiametra pulchra* under light microscope. The position of the statocyst is indicated by the arrow; arrowheads point to the developing eggs; the male and female copulatory organs are surrounded by the oval. **D**. Adult specimen of *Symsagittifera roscoffensis* under light microscope. The green cells correspond to the symbiotic algae *Tetraselmis colvolutae*. The position of the statocyst is indicated by the arrow; arrowheads point to the nerve cords.

the earliest branching clade of the Deuterostomia (*17*).

The clade Acoelomorpha is composed of two groups of small acoelomate flatworms: Acoela and Nemertodermatida. In general, these animals possess complex copulatory organs, a blind gut with a single ventral opening as a mouth, and a relatively simple nervous system with a true brain. However, protonephridia, segments and appendages are absent in these animals (for a general description of the morphology of the Acoela see **Figure 1B**). Special stem cells called neoblasts allow these animals to regenerate missing body parts (*18, 19*). Acoelomorpha are direct developers; a juvenile worm hatches out from the eggs, and becomes an adult after several days or weeks, depending on the species (**Figure 1C**). The Acoela comprises approximately 370 species, living in marine habitats from the tropics to the poles, with some of them establishing symbiotic relationships with unicellular algae (*20*) (see the green cells in **Figure 1D**). In contrast, Nemertodermatida is a clade composed of only six genera; most are free-living animals that live in marine coastal areas, though one studied species (*Meara stichopi*) lives as a parasite in the pharynx of sea cucumbers (*21*).

Whereas in nemertodermatids the gut is fully epithelial, most acoels, with the exception of some with truly cellular guts, lack a gut lumen, and the gut epithelium has been transformed into a syncytial tissue mass (*22*).

Even though the basal position of the Acoelomorpha at the base of the Bilateria has not been proved definitively, the change in the phylogenetic position of the Acoelomorpha, from the Platyhelminthes to an early diverging bilaterian group, has stirred a great interest in this group of animals. Obviously, and in order to clarify the origin of the major body axis, one of the first goals has been the characterization of the Hox–ParaHox genetic system in the Acoelomorpha (*23-27*). In this review we revisit all the available information on Hox–ParaHox genes in the Acoelomorpha and discuss the most plausible scenarios for the early evolution of this genetic system. For this aim we will also need to take into account all the information gathered over the last few years on the characterization of the cnidarian Hox–ParaHox counterparts.

The Hox and ParaHox Gene Complements in Acoelomorpha

The complement of Hox–ParaHox genes in the Acoelomorpha has been analyzed by different authors and in different species of acoels [*Symsagittifera roscoffensis* and *Paratomella rubra* (*23*); *S. roscoffensis* and *Isodiametra pulchra* (*24*); *Convolutriloba longifissura* (*26*); and *Convolutriloba retrogemma* (*27*)] and one nemertodermatid species [*Nemertoderma westbladi* (*28*)]. All the studies show that the acoels contain three Hox genes, belonging to the anterior (PG1), central (PG5) and posterior (PG9-10) classes, respectively. The presence of two posterior Hox genes in *P. rubra* (*23*) is probably the result of a lineage-specific gene duplication in the Paratomellidae. However, in the nemertodermatid *N. westbladi*, the complement of Hox genes is slightly different, with only one posterior and two (instead of one) central Hox genes. No anterior relatives were identified in the PCR screens that revealed those genes (*28*). It has been suggested that a cis-duplication in this lineage was responsible for the appearance of the second central gene (*28*).

One of the most striking findings derived from these analyses was the great difference in the number of genes existing between the traditionally assumed complex Hox complement proposed for the NLCA (with 7-8 genes) (*29*) and the small and simple Hox complement (composed of three genes) that was probably present in the last common ancestor of the Acoelomorpha (*24, 26*). The most obvious implication of these results is that there was a general increase in the number of Hox genes from the LCBA to the NLCA during the evolution of the bilateral animals. The three Hox genes present in acoels would be, in the new scenario, those present in the first bilaterian HOX cluster (*26*).

Using informatics approaches and motif-based reconstruction methodologies, Ogishima and Tanaka (*30*) reconstructed the putative sequences of the ancestral Hox genes present in the C-BLCA and the LCBA, which were precursors to the Hox genes found in present-day animals. Strikingly, the sequence from the three Hox genes of acoels (containing the motifs Na, Nb, Nbx and Ca) (*26*) perfectly matched those

predictions, providing further support for the acoel Hox genes as *bona fide* proxies for those present in the first bilaterian HOX cluster. Taking into account the fact that the different analyses performed in Cnidaria suggest that the C-BLCA possessed only genes belonging to the anterior and posterior classes, the Hox complement of the Acoelomorpha represents an intermediate stage in the evolution of the Hox group of genes between those present in the C-BLCA and those present in the rest of the extant bilaterian groups. It is interesting to note the appearance of the Bilateria at the same time as the appearance of a new class of Hox genes in this group of animals, *i.e.*, the central class, represented most probably by a PG5-like gene.

In summary, the set of three Hox genes (belonging to the anterior, central and posterior classes of paralogous groups, respectively) found in the five acoel species studied so far might represent the ancestral condition in Bilateria, corresponding to the minimal set needed for establishing the first ''Hox code'' in the evolution of animals, a code that provides positional information along the AP body axis. From these data, it has been suggested that one gene from each of the three major groups of Hox genes would represent the minimal set compatible with a bilaterian grade of structural organization (*24*).

With regards to the ParaHox genes, only Cdx orthologues have been found in the acoels *S. roscoffensis* and *C. longifissura* (*23, 31*). Thus, until sequences of whole genomes are available, the exact number of ParaHox genes will remain uncertain. It is important to note that in nemertodermatids the presence of an Xlox gene (central ParaHox) plus one Cdx orthologue (*28*) has been reported. Given that Xlox is involved in midgut patterning in bilaterians as well as in the epithelial digestive system in general (*12, 32*), it has been proposed that the transformation of a true epithelial gut into a syncytial digestive system in acoels (*33*) might be linked to the loss or modification of the Xlox gene in the acoel genomes (*26*). However, we understand that this is a simplistic scenario. In any case, and according to that hypothesis, the ancestral epithelial digestive system (present in nemertodermatids and cnidarians) would have been reduced to a syncytium in most of the acoel lineage, in concert with the loss of Xlox from the genome.

HOX Clusters in Basal Taxa Have Disintegrated

In the Bilateria, Hox genes are often organized in evolutionarily conserved clusters in the genome. However, the level of conservation of these clusters varies among groups. Vertebrates possess the most compact clusters, with all of their genes in the same transcriptional orientation (*34*). In amphioxus, the genes are also in the same transcriptional orientation, but the distance between the genes is greater and as a consequence the cluster is considerably longer (*35-37*). However, in some bilaterian groups, the original cluster has been atomized, with the different Hox genes being dispersed in the genome and some of them having been lost. Such losses might be related to the secondary adaptation of a lineage to a sessile lifestyle, *e.g.*, ascidians (*38*), or to parasitism, *e.g.*, nematodes (*39*) and Platyhelminthes (*40*). Interestingly, the Hox relatives are not linked in single genomic clusters in any cnidarian species studied to date (*6, 7*).

The lack of available sequenced genomes from any acoelomorph has prevented us from knowing the detailed arrangement of these genes in their genome. But recently, thanks to the generation of a genomic library from *S. roscoffensis* in BACs, Moreno and co-workers (*24*) have been able to localize the three identified Hox genes on metaphase chromosomes, using fluorescence *in situ* hybridization techniques. This study revealed that these Hox genes are all located on different chromosomes. Therefore, the authors showed that, at least in this acoel species, Hox genes do not form an organized cluster. It is highly plausible that an original HOX cluster present in the ancestor of all bilaterians has been atomized in this lineage and that the Hox genes are now scattered in the genome, as has occurred in other groups, for instance in the urochordate *Oikopleura dioica* (*41*) or in the studied cnidarians (*7*).

The stimuli that promoted the disintegration of the cluster in *S. roscoffensis* are difficult to explain, given that this organism neither has a sessile life style nor a parasitic one (which might explain the breaking of other HOX clusters). However, it is clear that there have been no strong constraints to maintain the cluster

intact over the long, independent evolution of this lineage (*e.g.*, the need for global regulators or the sharing of exons) (*34*). The mechanisms underlying the disintegration are also unclear, though it is worth highlighting the presence of transposon-related sequences adjacent to the central and posterior Hox genes in the *S. roscoffensis* genome (*24*). A recent analysis of the sequenced *S. roscoffensis* BACs revealed the presence of several copies of class I transposable elements (TEs) in the Hox gene vicinities (**Figure 2**). Each BAC harbors at least one truncated copy of an LTR retrotransposon. Order of the canonical structural features (aspartic protease, reverse transcriptase and integrase), phylogeny based on the reverse transcriptase domain and BLAST similarity clearly suggest that these copies belong to the Ty3-Gypsy or Bel-Pao clades. To which of them cannot be ascertained now, since both are closely related and the information gathered from the copies is partial. There is some evidence (*i.e.*, low BLAST similarity) that points to the presence of extra copies of repetitive elements on BACs related to central and posterior Hox genes. These copies are highly degenerated and truncated, thus even though they can still be recognized as ancient elements, the type cannot be determined with confidence. Homology to integrase and Pox_A32 domains suggests that these elements could also belong to the LTR superfamily, although the analysis of other genomic copies is essential. All these identified sequences could promote chromosomal rearrangements by the unequal recombination between these and other similar sequences present in different chromosomes (*42*). In fact, it has been noticed that TEs are also present in the neighborhood of Hox genes in the genomes of several of those organisms with fragmented HOX clusters (*43*).

In the groups in which the HOX clusters are intact, it is not only the Hox genes that maintain their relative positions within the cluster, but also some of the adjacent genes. For example, the homeobox gene Evx is usually located at the 5' end of the cluster, just behind the last posterior Hox gene (*44*). There are also two specific microRNAs associated with the HOX cluster. One is *mir-10*, located between *hox5* and *hox4* in vertebrates and arthropods, but absent in some nematodes, echinoderms and urochordates, probably a secondary effect of their cluster disruptions (*45, 46*). Another is *mir-196*, which is present upstream of the PG9 gene, though only in vertebrates. No invertebrate homologues of *mir-196* have been found to date (*47*).

In *S. roscoffensis*, three positive BACs (each containing a different Hox) have been completely sequenced (covering a total area of 300 kb) with the main aim of analyzing the possible existence of syntenic relationships between these and the corresponding regions of other animal genomes. This has allowed the prediction (and subsequent validation by PCR) of the presence of some putative transcriptional units in the vicinity of the three Hox genes. These open reading frames encode proteins without orthologues in the equivalent, Hox-containing genomic regions of other bilaterians or in the cnidarian *Nematostella*. In addition, neither Evx nor the microRNAs associated with HOX clusters seem to be present in the three BAC clones sequenced from *S. roscoffensis*, suggesting a complete lack of syntenic

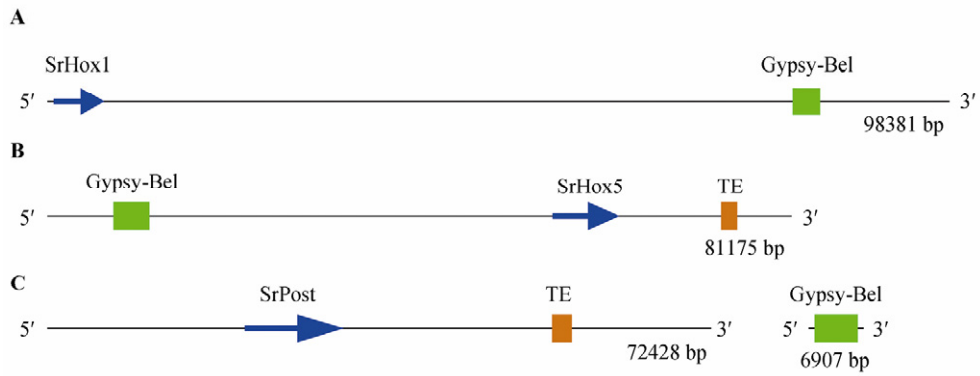

Figure 2 *S. roscoffensis* BACs containing the anterior (**A**), central (**B**) and posterior (**C**) Hox genes harbor at least one copy of a class I TE (green) that belongs to the Ty3-Gypsy or Bel-Pao clades. Traces of ancient copies of TEs are also present (orange).

relationships between these HOX "equivalent" regions, although the genomic region upstream from the anterior Hox has not yet been explored (*24*). This lack of conservation of synteny is a typical feature of dispersed and broken clusters (*38*). Since *S. roscoffensis* is, to date, the only acoel species from which we have genomic sequences in the regions containing Hox genes, we will have to wait until the sequencing of complete genomes from several acoel and nemertodermatid species is known before we can obtain a clear picture of the evolutionary process that led to the complete dispersal of the HOX cluster in, most probably, the whole Acoelomorpha lineage.

Evolution of Bilaterian Hox Functions

The expression patterns of the Hox and Cdx genes have been analyzed in different species of acoels during embryonic and postembryonic development, as well as in adulthood. Unfortunately, analysis of the expression patterns of these genes (and Xlox) in nemertodermatids has still not been carried out.

The embryonic expression of Hox genes has only been described in the acoel *C. longifissura* (*26*). In this species the anterior Hox is active in the animal hemisphere of the embryo, whereas the central and the posterior Hox are expressed at the site of gastrulation and in the vegetal hemisphere of the embryo. Though the spatial domains are partially nested (remnants of spatial colinearity), the temporal expressions of all three Hox genes, which are activated simultaneously after gastrulation, do not show any evidence of temporal collinearity. If the absence of a Hox cluster is confirmed for the whole acoel clade, this would suggest that the absence of temporal collinearity is, perhaps, associated with the processes of cluster disintegration, since temporal collinearity seems to depend tightly on the presence of intact HOX clusters (*48*).

The postembryonic expression of Hox genes has been analyzed in *C. longifissura, S. roscoffensis* and *I. pulchra* (*24-26*), though in the latter species only the posterior Hox expression has been determined. The anterior Hox is expressed in *S. roscoffensis* in an anterior domain that includes the brain and areas of peripheral parenchyma surrounding the statocyst. In *C. longifissura* this anterior expression domain begins behind the statocyst. While in *C. longifissura* a posterior domain is also detected covering a large part of the body length, in *S. roscoffensis* this posterior domain consists of two lateral bands of peripheral parenchyma. The central Hox is expressed in both species in two bilaterally paired domains, from the statocyst to the posterior third of the body. Finally, the posterior Hox is expressed in the peripheral parenchyma of the three mentioned species and in areas covering the terminal part of the body. The expression patterns of these three Hox genes in nested domains along the AP body axis of the juvenile worm suggest the presence of a Hox-based vectorial patterning system, or "Hox code" in the animal that provides positional information along the major body axis (as in other Bilateria). It is important to note that the presence of a putative "Hox code" functioning in *S. roscoffensis* and *C. longifissura* would represent the oldest (evolutionary) example known for any metazoan, indicating that the "code" was already present in the LCBA. In contrast, the expression of Hox genes is not nested in cnidarians, which suggests that the implementation of a Hox-collinearity system was a bilaterian innovation that evolved in the lineage leading to LCBA (*8*). The expression patterns of Hox genes in different cnidarians in fact suggest that Hox genes specified some specific structures along the oral–aboral axis in a radial C-BLCA, but never in a coordinated fashion. It was only later on in evolution that they were co-opted, as a group, for patterning the LCBA major (AP) axis. The system as such was then utilized in many developmental contexts, such as limbs and genital tract.

The Hox gene expression in acoels provides yet more evidence that the nested expression of the Hox genes does not require the presence of an intact cluster (*38, 39, 41, 45*). Since there is evidence showing that a cluster is needed only for coordinated temporal expression, it is clear that the temporal and spatial patterning of Hox systems is controlled, most probably, by distinct mechanisms (*49-51*).

An interesting aspect derived from the analysis of acoel Hox patterns is the possibility of suggesting models for the ancestral, bilaterian role of Hox genes. Two alternatives have been proposed. Hejnol and Martindale (*26*) suggest that since in the acoel *C.*

longifissura Hox genes might have a major role in the axial patterning of the nervous system, a neural role would be the ancestral one within bilaterians, subsequently being co-opted for the patterning of other tissues along the AP axis [an ancestral role also proposed by Deutsch and Le Guyader (*52*)]. In contrast to the Hejnol and Martindale proposal (*26*), the results from Moreno *et al* (*25*), which included the knockdown of this gene in *I. pulchra*, seem to indicate the possibility that posterior Hox genes were mainly devoted to the regulation of the postembryonic mesoderm and musculature. Similar mesodermal roles may have been ancestral for the other Hox. In any case, deciding between a neural-first or a mesoderm-first role for the bilaterian Hox-system will depend, eventually, on having more data from embryonic usage in other acoel species as well as improving the resolution of the acoel *in situ* patterns and having access to reference tissue markers.

In a phylogenetic context, however, the lack of knowledge on the function of Hox genes in cnidarians makes it difficult to trace any parallelisms between the uses of Hox genes in cnidarians and acoels (or bilaterians). At present it is difficult to make an educated guess about the evolution of Hox activities from the C-BLCA onwards. In fact, the only knockdown of Hox genes performed in Cnidaria was achieved by inhibiting the activity of the posterior-like Hox genes *Cnox-1* and *Cnox-3* in the species *Eleutheria dichotoma.* This experiment showed some phenotypic effects concentrated in one (the oral) pole of the medusa (*53*). Obviously this is not enough to understand the commonalities in the usage of Hox genes within the phylum Cnidaria.

The Role of ParaHox Genes in Early Bilaterians

The expression of only one of the two ParaHox genes reported from the Acoelomorpha has been analyzed, the acoel caudal orthologue Cdx. The expression patterns of the identified Cdx and Xlox orthologues in nemertodermatids are still unknown.

In the remaining Bilateria, the expression of Cdx in *Drosophila* (*54*), vertebrates (*55*), or amphioxus (*56*) is fundamentally restricted to the hindgut, anal structures and central nervous system (CNS). This functional conservation indicates that Cdx was probably involved in patterning the same regions in the LCBA (*57*).

In line with the findings in other bilaterians, in the acoels *C. longifissura* and *S. roscoffensis*, Cdx is expressed in the most posterior area of adult animals, a region that will form the male gonopore. According to Hejnol and Martindale (*31*), and because the metazoan mouth evolved first, it is more parsimonious to consider that the anal opening arose independently in different groups by co-opting hindgut genes in posterior domains at the ectodermal–endodermal boundary. This would explain why, in the absence of an anus, the expression of Cdx is restricted to the posterior area where the male gonopore opens.

In juveniles of *S. roscoffensis*, the pattern is clearly different from the one detected in adults, but very similar to the pattern described in juveniles of *C. longifissura* (*26, 31*). Strikingly, at this developmental stage Cdx is expressed within the nervous system. The domain of expression includes the commissural area located around the statocyst, extending further down the nerve cords and along the whole AP body axis (**Figure 3**). These results indicate that while Cdx may play a role in the differentiation of the nerve cords during post-embryonic development, it changes its role during adulthood to regulate the formation of the male gonoporal region.

Since the expression patterns of the Cdx relatives are highly variable among cnidarian species [*e.g.*, *Nematostella vectensis* (*6*), *E. dichotoma* (*8*) and *Clytia hemisphaerica* (*4*)], the ancestral role of this gene within the Cnidaria and whether the role changed (or not) at the origin of the Bilateria remain unclear.

Before closing this section, it is important to note that, in addition to the well-conserved posterior gut domain of Cdx, another ParaHox gene, Xlox, also has conserved roles in the Bilateria, always in domains anterior to Cdx. These commonalities of expression suggest that the ParaHox system (or at least the two posterior ParaHox genes) might have been in place, as was the case with the Hox, at the origin of the Bilateria. While the Hox system was deployed fundamentally in the mesoderm and ectoderm layers, the ParaHox system was devoted to regionalizing the gut and, most probably, parts of the CNS.

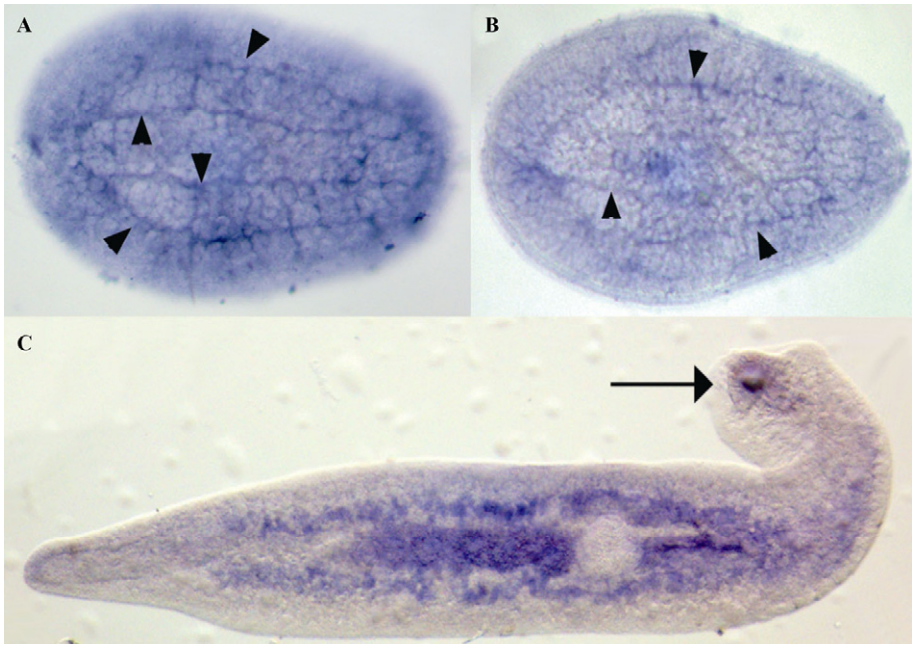

Figure 3 SrCdx expression patterns in juveniles (**A** and **B**) and adult specimens (**C**) of *S. roscoffensis*. In juveniles, *SrCdx* is mainly expressed in the nervous system. It labels thin nerve tracts running in parallel from the commissures around the statocyst along the AP body axis (arrowheads). In adults, expression is found around the male gonopore (arrows), in two rows of cells of peripheral parenchyma running along the AP axis, and in the central parenchyma. All pictures show a dorsal view, with the anterior to the left.

Evolution of the Hox–ParaHox Genetic System

The proposed models of Hox and ParaHox gene evolution in the Metazoa take into account the various phylogenetic studies published during recent years that have attempted to reconcile gene and lineage evolution (measured by other parameters). In general there are two main scenarios for the evolutionary history of the two families (*58*). On the one hand, the "multigene (or segmental) duplication" scenario assumes that the HOX and ParaHox clusters arose by segmental (whole cluster) duplication of a ProtoHox cluster, which later on split to originate the HOX and ParaHox clusters (*56, 59-64*). Under this scenario, the HOX and ParaHox are sister clusters. The model is supported by the affinities between the Hox and ParaHox genes, where the ParaHox genes Gsx, Xlox and Cdx are most closely related to Hox1-2, Hox3 and Hox9-15, respectively. The alternative is the "tandem duplication" scenario that assumes that the HOX and ParaHox clusters originated by a series of individual gene duplications originating in a ProtoHox gene cluster. These duplication events culminated in a final split between the two sets of genes, the Hox and ParaHox, which were moved to different genome regions by a translocation event. In this case the ParaHox genes would represent a group of detached genes from the ProtoHox cluster, rather than Hox sister cluster genes (*4, 6, 26*). Most current phylogenies are consistent with the latter hypothesis.

Taking into account the existing data on the Hox–ParaHox genes present in the Cnidaria and Bilateria, this scenario assumes the presence of anterior and protoposterior Hox genes, plus the three ParaHox genes in the C-BLCA (**Figure 4**). Since orthologies between posterior Hox genes in cnidarians and bilaterians are still not clear, it is fair to suppose that cnidarian and bilaterian posterior Hox genes were produced independently in both lineages from the proto-posterior Hox gene (PPHox) present in the C-BLCA. This gene duplicated in the C-BLCA, giving rise to two posterior Hox genes, one is inherited by the cnidarian lineage (CPHox), and the other is inherited by the bilaterian lineage. These two genes were the precursors of all the extant cnidarian and bilaterian (PG9/15) posterior Hox genes (*4*). In this

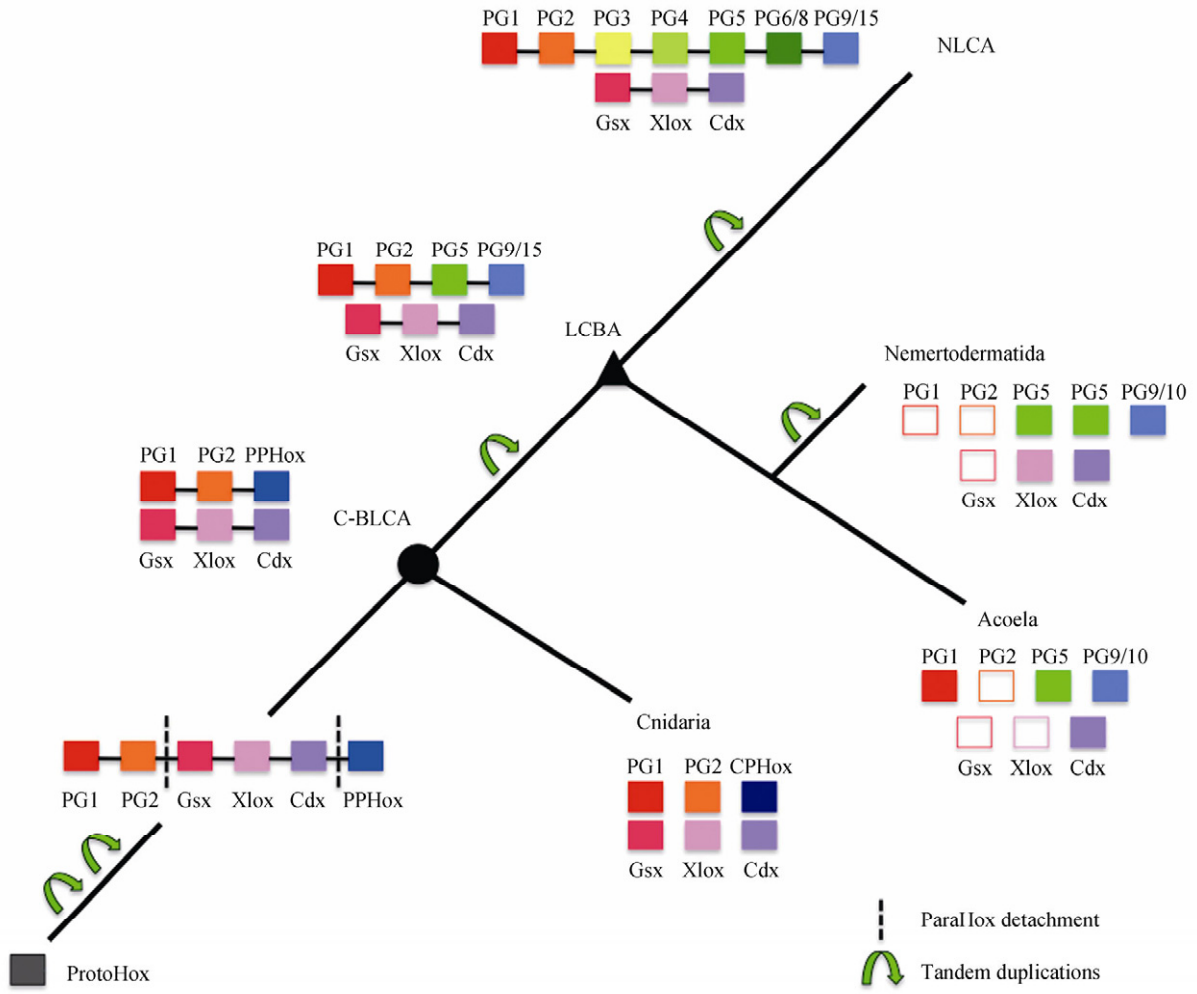

Figure 4 A model of the evolution of HOX and ParaHox clusters in metazoans from a single ProtoHox cluster after the divergence of Porifera. The model integrates data from recent phylogenetic studies in Cnidaria, and tries to combine the most parsimonious hypothesis for Hox–ParaHox evolution. A unique ANTP gene in the lineage leading to the C-BLCA duplicated several times in tandem, giving rise to an ancestral HOX–ParaHox gene cluster, which later split (broken line) and moved to different regions of the genome. Since orthologies between posterior Hox genes in cnidarians and bilaterians are still not clear, it is fair to suppose that cnidarian and bilaterian posterior Hox genes were produced independently in both lineages from the PPHox present in the C-BLCA. This gene duplicated in the C-BLCA, giving rise to two posterior Hox genes, one inherited by the CPHox and the other by the bilaterian lineage. These genes were the precursors of all the extant cnidarian and bilaterian (PG9/15) posterior Hox genes (*4*). In the lineage leading to the LCBA, another tandem duplication gave rise to the PG5 gene. From the LCBA, two sister-groups formed: one leading to present day acoelomorphs, and another giving rise to the NLCA. In the acoel lineage, the original cluster might have disintegrated (at least in *S. roscoffensis*) and some genes were lost (tentatively but not proven, the genes PG2, Gsx and Xlox), here represented by empty boxes. A similar process (yet to be proven) could have occurred in nemertodermatids, although the absence of anterior genes (empty boxes) is more probably the result of limited sampling. A further tandem duplication in the Nemertodermatida lineage originated a second PG5 gene. Finally, a series of tandem duplications, involving the central Hox class, gave rise to the extended HOX cluster present in the NLCA.

model, the origin of the bilateral animals would be related to two important events in the evolution of the Hox–ParaHox genetic system: on the one hand with the appearance of the central Hox class of genes (since genes homologous to central Hox genes have been found only in Bilateria and not in cnidarians; *i.e.*, the PG5 in acoelomorphs), and on the other hand with the establishment of a simple "Hox code" for patterning the AP axis at the LCBA. The huge variability in Hox expression patterns among cnidarian species

suggests that the HOX cluster, as such, was still not involved in patterning the oral–aboral body axis in the C-BLCA.

In the acoel lineage, the original clusters would have disintegrated, as the data from *S. roscoffensis* suggest, and some genes might have been lost (tentatively, but not proven, the genes PG2, Gsx and Xlox). A similar process could have occurred in nemertodermatids. These assertions should, in any case, be toned down since we are taking as a reference the status in *S. roscoffensis,* which is considered a derived species within acoels. A more solid conclusion could be reached after exploring the gene complement and genomic organization in other more basal acoel species. Then a general statement about Hox genes in all acoels can be made.

Later on during the evolution of the Bilateria, the basic complement of Hox genes present in the LCBA duplicated in tandem, giving rise to the PG3 relative. Moreover, the central paralogue gave rise to the five central PG genes (PG4-8) that are present in many deuterostomes and protostomes. In addition, in the deuterostome lineage, the posterior paralogue was duplicated many times *in cis*, giving rise to seven posterior PGs (PG9-15). On the other hand, the number of ParaHox genes has also changed in different lineages over evolutionary time. They are the result of specific losses in some groups [*e.g.*, Xlox in acoels (*24*), hagfish (*65*) and nematodes (*66*)], and selective expansions in others, the latter mainly linked to genome duplication events, but not always [*e.g.*, Xlox genes in the hemichordate *Ptychodera* (*67*)].

The results on the Hox complement in acoels suggest a gradual expansion of genes in the HOX cluster of the early bilaterians, starting with a Hox complement of four genes in the LCBA. The number of acoel Hox genes would represent an intermediate number in the evolution from the small complement present in the C-BLCA to the large complement in the NLCA. These data contradict the hypothesis that the sudden appearance of most bilaterian phyla during the Cambrian was causally related to the sudden expansion of the HOX cluster, and suggest an alternative scenario with the progressive incorporation of new characters (Hox and others) during the early evolution of bilateral animals (*68*).

However, we should not forget that the information about the Hox–ParaHox complement in the most basal bilaterian clades (acoels, nemertodermatids and xenoturbellids) is still very incomplete. The sequencing of the genomes of different species within these clades would be most desirable given the limited power that PCR screens have to detect the full complement of Hox genes in many organisms. Extensive genome sequencing will allow us to have a more precise indication of the complement of Hox genes in the different early bilaterian groups and thus we should be able to predict, with accuracy, the complement of Hox genes present in the LCBA. In addition, such genomic information should help us in the understanding of the syntenic relationships between the different Hox–ParaHox genome-containing areas, thus clarifying the history of the HOX clusters at the dawn of the Bilateria.

Acknowledgements

The work done in the laboratory of P. Martinez was supported by a grant from the Ministry of Science and Innovation of Spain. We would like to express our gratitude to Drs. Andreas Hejnol (Bergen, Norway), Mark Martindale (Hawaii, USA) and Jaume Baguñà (Barcelona, Spain) for sharing their knowledge on Hox genes over the years. We also thank the referees for helping us improving the manuscript.

References

1 McGinnis, W. and Krumlauf, R. 1992. Homeobox genes and axial patterning. *Cell* 68: 283-302.
2 Bürglin, T.R. 1996. Homeodomain proteins. In *Encyclopedia of Molecular Biology and Molecular Medicine* (ed. Meyers, R.A.), Vol.3, pp.55-76. VCH Verlagsgesellschaft mbH, Weinheim, Germany.
3 Adoutte, A., *et al*. 2000. The new animal phylogeny: reliability and implications. *Proc. Natl. Acad. Sci. USA* 97: 4453-4456.
4 Chiori, R., *et al*. 2009. Are Hox genes ancestrally involved in axial patterning? Evidence from the hydrozoan *Clytia hemisphaerica* (Cnidaria). *PLoS One* 4: e4231.
5 Quiquand, M., *et al*. 2009. More constraint on ParaHox than Hox gene families in early metazoan evolution. *Dev. Biol.* 328: 173-187.
6 Ryan, J.F., *et al*. 2007. Pre-bilaterian origins of the Hox

cluster and the Hox code: evidence from the sea anemone, *Nematostella vectensis*. *PLoS One* 2: e153.
7 Chourrout, D., *et al*. 2006. Minimal ProtoHox cluster inferred from bilaterian and cnidarian Hox complements. *Nature* 442: 684-687.
8 Kamm, K., *et al*. 2006. Axial patterning and diversification in the Cnidaria predate the Hox system. *Curr. Biol.* 16: 920-926.
9 Finnerty, J.R., *et al*. 2004. Origins of bilateral symmetry: Hox and Dpp expression in a sea anemone. *Science* 304: 1335-1337.
10 Gauchat, D., *et al*. 2000. Evolution of Antpclass genes and differential expression of Hydra Hox/paraHox genes in anterior patterning. *Proc. Natl. Acad. Sci. USA* 97: 4493-4498.
11 Gilbert, S.F. 2006. *Developmental Biology* (the eighth edition). Sinauer Associates, Sunderland, USA.
12 Arnone, M.I., *et al*. 2006. Genetic organization and embryonic expression of the ParaHox genes in the sea urchin *S. purpuratus*: insights into the relationship between clustering and colinearity. *Dev. Biol.* 300: 63-73.
13 Pearson, J.C., *et al*. 2005. Modulating Hox gene functions during animal body patterning. *Nat. Rev. Genet.* 6: 893-904.
14 Gegenbaur, C. 1859. Grundzüge der Vergleichenden Anatomie. W. Engelmann, Leipzig, Germany.
15 Ruiz-Trillo, I., *et al*. 1999. Acoel flatworms: earliest extant bilaterian Metazoans, not members of Platyhelminthes. *Science* 283: 1919-1923.
16 Hejnol, A., *et al*. 2009. Assessing the root of bilaterian animals with scalable phylogenomic methods. *Proc. Biol. Sci.* 276: 4261-4270.
17 Philippe, H., *et al.* 2007. Acoel flatworms are not platyhelminthes: evidence from phylogenomics. *PLoS One* 2: e717.
18 De Mulder, K., *et al*. 2009. Characterization of the stem cell system of the acoel *Isodiametra pulchra*. *BMC Dev. Biol.* 9: 69.
19 Karling, T.G. 1974. On the anatomy and affinities of the turbellarian orders. In *The Biology of the Turbellaria* (eds. Riser, N.W. and Morse, M.P.), pp.1-16. McGraw-Hill, New York, USA.
20 Keeble, F. 1910. *Plant-Animals: A Study in Symbiosis.* Cambridge University Press, London, UK.
21 Westblad, E. 1949. On *Meara stichopi* (Bock) Westblad, a new representative of *Turbellaria archoophora*. *Ark. Zool.* 1: 43-57.
22 Doe, D.A. 1981. Comparative ultrastructure of the pharynx simplex in Turbellaria. *Zoomorphology* 97: 133-193.
23 Cook, C.E., *et al*. 2004. The Hox gene complement of acoel flatworms, a basal bilaterian clade. *Evol. Dev.* 6: 154-163.
24 Moreno, E., *et al*. 2009. Tracking the origins of the bilaterian Hox patterning system: insights from the acoel flatworm *Symsagittifera roscoffensis*. *Evol. Dev.* 11: 574-581.
25 Moreno, E., *et al*. 2010. Inferring the ancestral function of the posterior Hox gene within the Bilateria: controlling the maintenance of reproductive structures, the musculature and the nervous system in the acoel flatworm *Isodiametra pulchra*. *Evol. Dev.* 12: 258-266.
26 Hejnol, A. and Martindale, M.Q. 2009. Coordinated spatial and temporal expression of Hox genes during embryogenesis in the acoel *Convolutriloba longifissura*. *BMC Biol.* 7: 65.
27 Sikes, J.M. and Bely, A.E. 2010. Making heads from tails: development of a reversed anterior-posterior axis during budding in an acoel. *Dev. Biol.* 338: 86-97.
28 Jiménez-Guri, E., *et al*. 2006. Hox and ParaHox genes in Nemertodermatida, a basal bilaterian clade. *Int. J. Dev. Biol.* 50: 675-679.
29 Balavoine, G., *et al*. 2002. Hox clusters and bilaterian phylogeny. *Mol. Phylogenet. Evol.* 24: 366-373.
30 Ogishima, S. and Tanaka, H. 2007. Missing link in the evolution of Hox clusters. *Gene* 387: 21-30.
31 Hejnol, A. and Martindale, M.Q. 2008. Acoel development indicates the independent evolution of the bilaterian mouth and anus. *Nature* 456: 382-386.
32 Wright, C.V., *et al*. 1988. Xlhbox 8: a novel *Xenopus* homeo protein restricted to a narrow band of endoderm. *Development* 104: 787-794.
33 Smith, J.P.S. and Tyler, S. 1985. The acoel turbellarians: kingpins of metazoan evolution or a specialized offshoot? In *The Origins and Relationships of Lower Invertebrates* (eds. Conway Morris, S., *et al.*), pp.123-142. Oxford University Press, Oxford, UK.
34 Duboule, D. 2007. The rise and fall of Hox gene clusters. *Development* 134: 2549-2560.
35 Garcia-Fernàndez, J. and Holland, P.W. 1996. Amphioxus Hox genes: insights into evolution and development. *Int. J. Dev. Biol.* Supp1: 71-72S.
36 Ferrier, D.E., *et al*. 2000. The amphioxus Hox cluster: deuterostome posterior flexibility and Hox14. *Evol. Dev.* 2: 284-293.
37 Holland, L.Z., *et al.* 2008. The amphioxus genome illuminates vertebrate origins and cephalochordate biology. *Genome Res.* 18: 1100-1111.
38 Ikuta, T., *et al*. 2004. *Ciona intestinalis* Hox gene cluster: its dispersed structure and residual colinear expression in development. *Proc. Natl. Acad. Sci. USA* 101: 15118-15123.
39 Aboobaker, A.A. and Blaxter, M.L. 2003. Hox gene loss during dynamic evolution of the nematode cluster. *Curr. Biol.* 13: 37-40.
40 Koziol, U., *et al*. 2009. Hox genes in the parasitic platyhelminthes *Mesocestoides corti*, *Echinococcus multilocularis*, and *Schistosoma mansoni*: evidence for a

reduced Hox complement. *Biochem. Genet.* 47: 100-116.
41 Seo, H.C., *et al*. 2004. Hox cluster disintegration with persistent anteroposterior order of expression in *Oikopleura dioica*. *Nature* 431: 67-71.
42 Lim, J.K. and Simmons, M.J. 1994. Gross chromosome rearrangements mediated by transposable elements in *Drosophila melanogaster*. *Bioessays* 16: 269-275.
43 Fried, C., *et al*. 2004. Exclusion of repetitive DNA elements from gnathostome Hox clusters. *J. Exp. Zool. B Mol. Dev.* 302: 165-173.
44 Minguillón, C. and Garcia-Fernández, J. 2003. Genesis and evolution of the Evx and Mox genes and the extended Hox and ParaHox gene clusters. *Genome Biol.* 4: R12.
45 Cameron, R.A., *et al*. 2006. Unusual gene order and organization of the sea urchin Hox cluster. *J. Exp. Zool. B Mol. Dev. Evol.* 306: 45-58.
46 Tanzer, A., *et al*. 2005. Evolution of microRNAs located within Hox gene clusters. *J. Exp. Zool. B Mol. Dev. Evol.* 304: 75-85.
47 Yekta, S., *et al*. 2004. MicroRNA-directed cleavage of *HOXB8* mRNA. *Science* 304: 594-596.
48 Monteiro, A.S. and Ferrier, D.E.K. 2006. Hox genes are not always colinear. *Int. J. Biol. Sci.* 2: 95-103.
49 Soshnikova, N. and Duboule, D. 2009. Epigenetic temporal control of mouse Hox genes *in vivo*. *Science* 324: 1320-1323.
50 Tarchini, B. and Duboule, D. 2006. Control of Hoxd genes' collinearity during early limb development. *Dev. Cell* 10: 93-103.
51 Tschopp, P., *et al*. 2009. Uncoupling time and space in the collinear regulation of Hox genes. *PLoS Genet.* 5: e1000398.
52 Deutsch, J. and Le Guyader, H. 1998. The neuronal zootype. An hypothesis. *C. R. Acad. Sci. III* 321: 713-719.
53 Jakob, W. and Schierwater, B. 2007. Changing hydrozoan bauplans by silencing Hox-like genes. *PLoS One* 2: e694.
54 Wu, L.H. and Lengyel, J.A. 1998. Role of caudal in hindgut specification and gastrulation suggests homology between *Drosophila* amnioproctodeal invagination and vertebrate blastopore. *Development* 125: 2433-2442.
55 James, R., *et al*. 1994. Structure of the murine homeobox gene cdx-2. Expression in embryonic and adult intestinal epithelium. *J. Biol. Chem.* 269: 15229-15237.
56 Brooke, N.M., *et al*. 1998. The ParaHox gene cluster is an evolutionary sister of the Hox gene cluster. *Nature* 392: 920-922.
57 de Rosa, R., *et al*. 2005. Caudal and even-skipped in the annelid *Platynereis dumerilii* and the ancestry of posterior growth. *Evol. Dev.* 7: 574-587.
58 Ferrier, D.E. 2010. Evolution of Hox complexes. *Adv. Exp. Med. Biol*. 689: 91-100.
59 Lewis, E.B. 1978. A gene complex controlling segmentation in *Drosophila*. *Nature* 276: 565-570.
60 Gehring, W.J., *et al*. 1994. Homeodomain proteins. *Annu. Rev. Biochem.* 63: 487-526.
61 Schubert, F.R., *et al*. 1993. The Antennapedia-type homeobox genes have evolved from three precursors separated early in metazoan evolution. *Proc. Natl. Acad. Sci. USA* 90: 143-147.
62 Zhang, J. and Nei, M. 1996. Evolution of Antennapedia-class homeobox genes. *Genetics* 142: 295-303.
63 García-Fernández, J. 2005. Hox, ParaHox, ProtoHox: facts and guesses. *Heredity* 94: 145-152.
64 García-Fernández, J. 2005. The genesis and evolution of homeobox gene clusters. *Nat. Rev. Genet.* 6: 881-892.
65 Furlong R.F., *et al*. 2007. A degenerate ParaHox gene cluster in a degenerate vertebrate. *Mol. Biol. Evol.* 24: 2681-2686.
66 Ruvkun, G. and Hobert, O. 1998. The taxonomy of developmental control in *Caenorhabditis elegans*. *Science* 282: 2033-2041.
67 Peterson, K.J. 2004. Isolation of Hox and Parahox genes in the hemichordate *Ptychodera flava* and the evolution of deuterostome Hox genes. *Mol. Phylogenet. Evol*. 31: 1208-1215.
68 Baguñà, J., *et al*. 2008. Back in time: a new systematic proposal for the Bilateria. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363: 1481-1491.