The evolution and maintenance of Hox gene clusters in vertebrates and the teleost-specific genome duplication

SHIGEHIRO KURAKU and AXEL MEYER*
Lehrstuhl für Zoologie und Evolutionsbiologie, Department of Biology, University of Konstanz, Konstanz, Germany

ABSTRACT  Hox genes are known to specify spatial identities along the anterior-posterior axis during embryogenesis. In vertebrates and most other deuterostomes, they are arranged in sets of uninterrupted clusters on chromosomes, and are in most cases expressed in a “colinear” fashion, in which genes closer to the 3’-end of the Hox clusters are expressed earlier and more anteriorly and genes close to the 5’-end of the clusters later and more posteriorly. In this review, we summarize the current understanding of how Hox gene clusters have been modified from basal lineages of deuterostomes to diverse taxa of vertebrates. Our parsimony reconstruction of Hox cluster architecture at various stages of vertebrate evolution highlights that the variation in Hox cluster structures among jawed vertebrates is mostly due to secondary lineage-specific gene losses and an additional genome duplication that occurred in the actinopterygian stem lineage, the teleost-specific genome duplication (TSGD).

KEY WORDS: colinearity, two-round genome duplication, secondary gene loss

Introduction

Hox genes are transcription factors that serve crucial roles during development in particular in embryonic anterior-posterior (A-P) patterning. In vertebrates and most other deuterostomes, Hox genes are arranged in sets of uninterrupted clusters on chromosomes. They specify the positional identities along the A-P axis and are in most cases expressed in a “colinear” fashion, i.e., genes closer to the 3’-end of the Hox clusters are expressed earlier and more anteriorly and genes close to the 5’-end of the clusters later and more posteriorly (Duboule, 1994; Kessel and Gruss, 1991; Lewis, 1978; McGinnis and Krumlauf, 1992). By now, it is understood that the multiple Hox gene clusters in the genomes of vertebrates are the remnants of an ancestral single homeobox gene cluster that was generated by successive rounds of tandem duplications early during metazoan evolution (reviewed in Garcia-Fernandez, 2005a). At least one (more or less complete and uninterrupted) Hox gene cluster is present in the genomes of almost all extant animal phyla, except for poriferans (Garcia-Fernandez, 2005b; Kamm, et al., 2006). The so-called colinear relationship between their genomic arrangement and their temporal and spatial expression remains one of the most interesting aspects of Hox clusters. It has been suggested that there is a link between this special genomic architecture and the origin of morphological novelties, such as modifications of axial segmental elements seen in the carapace of turtles (Ohya et al., 2005), loss of limbs in snakes (Cohn and Tickle, 1999), and the acquisition of jaws in gnathostomes (Cohn, 2002; Takio et al., 2004). In this review, we aim to summarize briefly the standing variation in the structures of Hox gene cluster architectures among vertebrates and attempt to reconstruct their evolutionary history. In light of known phylogenetic relationships we discuss alternative evolutionary processes that might have led to the clustered chromosomal arrangement of Hox genes. We also briefly survey the potential evolutionary forces that kept Hox genes clustered.

Early deuterostome origins of the Hox cluster

Invertebrates typically, but not always, possess a single uninterrupted cluster of Hox genes while vertebrates have at least four such clusters (Fig. 1). There are no reports so far of invertebrates with more than one Hox cluster. In some invertebrate lineages,
their Hox cluster was secondarily broken - as seen, for example, in the fly (Drosophila melanogaster; Von Allmen et al., 1996) and the nematode (Caenorhabditis elegans; Van Auken et al., 2000). This Hox cluster breakage is thought to have been caused by lineage-specific events that interrupted and dislocated an ancestrally intact Hox cluster (Akam, 1989; Aboobaker and Blaxter, 2003; Negre et al., 2003). Future sequencing efforts will determine in how many animal phyla the Hox cluster is intact and how often during evolution it disassembled and partly relocated onto different chromosomes.

During the evolution of chordates from deuterostome ancestors, the genome was duplicated most likely twice consecutively in its entirety (Ohno, 1970; Lundin, 1993; Sidow, 1996; Dehal and Boore, 2005; McLysaght et al., 2002; Kasahara, 2007). The evolution of Hox gene repertoires and their genomic structures in deuterostomes need to be reconstructed based on correctly inferred phylogenetic relationships between deuterostome phyla and major lineages within them (Fig. 1). Formerly, the hemichordata and echinodermata were placed basally, and the cephalochordata was regarded as the sister group of vertebrates (‘Euchordata’ hypothesis; Rowe, 2004; Schaeffer, 1987; Wada and Satoh, 1994). This relationship was, however, recently revised through the analysis of large amounts of sequence information for a larvacean tunicate (Oikopleura dioica; Delsuc et al., 2006). These new data suggested a closer affinity of urochordates to vertebrates by excluding cephalochordata and clustered them with the echinodermata at the base of the deuterostomes (‘Olfactore’ hypothesis, after the name of clade combining urochordates and vertebrates; Jefferies, 1991). The Olfactore hypothesis has been supported by subsequent analyses that included the newly introduced phylum Xenoturbellida (Bourlat et al., 2006) and genome-wide sequence data of the Florida lancelet (Branchiostoma floridae) (Putnam et al., 2008), both of which placed cephalochordata at the basal lineage of chordates.

Among all deuterostome invertebrates, the genomic analyses of B. floridae provided the first report of what was interpreted to be the ancestral condition of vertebrate Hox clusters (Garcia-Fernandez and Holland, 1994). Together with Hox14 (Ferrier et al., 2000) and Hox15found more recently (Holland et al., 2008), the amphioxus Hox genes constitute an uninterrupted array of 15 Hox genes (AmphiHox1-15) and one Evx gene (Fig. 1; Amemiya et al., 2008; Holland et al., 2008). By contrast, in the sea urchin Strongylocentrotus purpuratus, only eleven Hox genes plus one Evx gene are aligned within 700 kb in a single genomic contig (Cameron et al., 2006). Notably, the gene order and orientation of transcription are quite scrambled in this echinoderm lineage (Fig. 1). For the phylum hemichordata, so far only cDNA sequences are available for Saccoglossus kowaleskii and Phycodiera flava (Lowe et al., 2003; Peterson, 2004), and the genomic structure of their Hox genes remains still unknown (Fig. 1). Within urochordates, the ascidian species Ciona intestinalis exhibits a broken Hox cluster with scrambled gene order and transcriptional orientation (Ikuta and Saiga, 2005), while the Hox cluster was disassembled secondarily in another urochordate, a larvacean O. dioica (Seo et al., 2004; Fig. 1). This is interpreted as an aberration, and might be related to an altered upstream machinery, where major components of retinoic acid (RA) signaling, retinoaldehyde dehydrogenase (Aldh1a1), cytochrome P450 family-26 (Cyp26), and the retinoic acid receptor (RAR), that normally specifies the spatial identities along the anterior-posterior axis, seems to be absent from its genome (Canestro and Postlethwait, 2007; Canestro et al., 2006). Strangely, however, the spatial expression patterns of Hox genes in this species still exhibit “colinearity”, in spite of the atomized cluster organization and a lack of what heretofore was seen as crucial signaling cascade (Seo et al., 2004; reviewed in Duboule, 2007). This surprising example poses the question of which kind of constraints keep multiple Hox genes clustered (reviewed in Duboule, 2007; also see Hoegg and Meyer, 2007 and Siegel, et al., 2007, for examples in other conserved gene clusters).

In spite of the observed variation of gene repertoires and their genomic organization among deuterostome invertebrates (Monteiro and Ferrier, 2006), the apparent similarity between the amphioxus Hox cluster and the gnathostome clusters seems to support the notion that it represents the ancestral structure of the vertebrate and also chordate Hox cluster. It should be noted again, however, that this inference strongly depends on the

Fig. 1. Evolution of Hox clusters in deuterostomes. The relationship between urochordates, cephalochordates, and vertebrates is depicted as a trichotomy, based on a conservative view between the conflicting reports of Delsuc et al. (2006) and Bourlat et al. (2006). The relationships between hagfishes, lampreys, and gnathostomes are based on references included in Kuraku and Kuratani (2006). The information for hagfish and lamprey is based on previous analyses (Force et al., 2002; Irvine et al., 2002; Studier et al., 2004). See Fig. 4 for details of the hypothetical gnathostome ancestor. Direction of pentagonal boxes indicates transcriptional orientation, where this information is available. For hemichordates, Hox cDNAs isolated both from Saccoglossus kowaleskii and Phycodiera flava are shown as in Monteiro and Ferrier (2006).
for the Euchordata, the loss of Hox14 gene occurred in the urochordate lineage plus one Evx gene that were all transcribed in the same direction (Fig. 1).

How many Hox genes made up the ancestral vertebrate Hox cluster?

A related question is how many Hox genes made up the ancestral vertebrate Hox cluster. The answer to this question hinges on some difficult issues that are specific to Hox genes. Molecular phylogenetic analyses of Hox genes usually can only provide phylogenetic trees with limited confidence (Malaga-Trillo and Meyer, 2001; Meyer, 1998). This is mainly due to the conserved nature of the homedomain that is only 60 amino acids in length. Regions outside of the homeobox cannot be aligned reliably across large evolutionary distances and therefore are not available for phylogenetic inferences. This is particularly problematic for the posterior abdominal genes (Hox9-13), since it is often difficult to determine paralogy group relationships, (posterior flexibility; Ferrier et al., 2000). This is also reflected in the uncertain nomenclature for posterior Hox genes, for example in hemichordate and echinoderms (Fig. 1).

The situation is further complicated by the recent discovery of the Hox14 genes located between the Hox13 and Evx gene in vertebrate Hox clusters. This paralogy group was first reported from rather basal vertebrate lineage – from horn shark and the coelacanth (Powers and Amemiya, 2004a;). This has been followed by the discovery in the elephant fish (Venkatesh et al., 2007) and lamprey (Kuraku et al., 2008). Interestingly, the molecular phylogenetic tree including available posterior genes did not suggest orthology of the Hox14 genes between amphioxus and vertebrates, whereas the position of the intron (‘split homeobox’) and gene location (between Hox13 and Evx) supported a single origin of Hox14 in the common ancestor of cephalochordates and vertebrates (Powers and Amemiya, 2004a; also see Ferrier, 2004; Garcia-Fernandez, 2005b, for reviews). Especially, when a single origin of Hox14 shared by both amphioxus and vertebrates is assumed, the Euchordata hypothesis described above seemed to fit this scenario better (Fig. 2A).

However, under the Olfactore hypothesis that is currently supported by large-scale molecular phylogenetic data, the scenario is not so clear-cut. If the Hox14 of amphioxus and vertebrates are orthologous to each other, a loss of Hox14 should be assumed in the urochordate lineages, because none of species in this group has been shown to possess a Hox14 homolog (Fig. 2B). As mentioned above, in fact, molecular phylogenetic analysis has never supported this orthology (Powers and Amemiya, 2004a; Kuraku et al., 2008; Holland et al., 2008). Moreover, it has been shown that an intron in the homeodomain does not serve as an unambiguous marker of Hox14 – a lamprey Hox13 homolog (LjHox13β) also has this feature (Kuraku et al., 2008). This accumulating evidence favors paralogy of Hox14 between amphioxus and vertebrates. So far, molecular phylogenetic trees have strongly suggested that amphioxus Hox13 and Hox14 duplicated in the cephalochordate lineage (Fig. 2C and 2D; Powers and Amemiya, 2004a; Kuraku et al., 2008). For the origin of vertebrate Hox14, however, two scenarios, shown in Fig. 2C and 2D, have been supported with relatively high confidence (Kuraku et al., 2008). In fact, the origin of Hox14 at the Olfactore ancestor (Fig. 2D), which compels a secondary loss of Hox14 in the urochordate lineage, was favored over other scenarios that are more parsimonious (Kuraku et al., 2008).

In vertebrates, Hox14 is retained by horn shark and elephant fish, whereas the coelacanth possesses Hox14 (Powers and Amemiya, 2004a; Venkatesh et al., 2007). Therefore, it is likely that Hox14 was gained before the quadruplication (1R/2R genome duplications) of Hox clusters (Powers and Amemiya, 2004a; also see Ferrier, 2004; Garcia-Fernandez, 2005b; Hoegg and Meyer 2005; Powers and Amemiya, 2004b for reviews). Based on a parsimony reconstruction of the Hox cluster evolution, it is likely that the Hox14 paralogy group is the only one that might have undergone two independent gene losses during two subsequent rounds of whole genome duplications (Fig. 3). To date, Hox14 genes have been found neither in any tetrapod nor any teleost genome in spite of relatively abundant sequence information for

A single origin of Hox14 + ‘Euchordata’ phylogeny

A single origin of Hox14 + ‘Olfactore’ phylogeny

Independent origins of Hox14 + ‘Olfactore’ phylogeny

Independent origins of Hox14 + ‘Olfactore’ phylogeny

Fig. 2. The phylogenetic timing of the origin(s) of Hox14 genes. Possible scenarios for gain(s) of Hox14 were inferred. Under the assumption of the ‘Euchordata’ tree, if Hox14 of amphioxus and vertebrates shared a common ancestry, the presence of Hox 14 can be interpreted as a synapomorphy for the Euchordata (A). Under the ‘Olfactore’ tree, and if Hox14 of amphibious and vertebrates are orthologous, it can be assumed that a secondary loss of Hox14 gene occurred in the urochordate lineage (B). If Hox14 of amphioxus and vertebrates are paralogous, the vertebrate Hox14 originated after the divergence from the urochordate lineage (C), or before (that is, at the common ancestor of vertebrates and urochordates) (D). In (D), the absence of the Hox14 gene in the urochordates is regarded as a result of a secondary gene loss in the urochordate lineage.
these lineages. It is suggested that less functional constraint, suggested by loss of expression in axial elements where the Hox-code is normally functioning, permitted the secondary losses of Hox14 genes in these lineages (Kuraku et al., 2008). Data from other taxa, such as hagfishes, non-teleost actinopterygians, non-tetrapod sarcopterygians will provide further information about the history of gains and losses of Hox14.

The example of Hox14 potentially cautions that gene order in Hox clusters does not necessarily imply orthology. Early in the evolution of bilateral body plans, the Hox cluster was generated as the result of successive tandem duplications (reviewed in Garcia-Fernandez, 2005a). Moreover, secondary shuffling of gene order in a Hox gene cluster is frequently observed in invertebrate deuterostomes (Fig. 1; Monteiro and Ferrior, 2006). Thus, relative positions of genes within a cluster itself, which have been the basis of gene annotation especially in studies of Hox genes, cannot alone serve as sole and unequivocal criterion for postulating homology relationships among Hox genes and their assignment to paralogy groups. This circumstance combined with the difficulty of reliable phylogeny reconstruction renders orthology statements tenuous. This makes it difficult to reconstruct unequivocally the evolutionary history of posterior Hox genes - at least between some deuterostome lineages (Fig. 1).

Agnathans: how many Hox clusters?

Agnathans are jawless fishes that branched off the chordate stem lineage early during vertebrate evolution. Extant agnathans are grouped into hagfishes (Myxiniformes) and lampreys (Petromyzoniformes), and the phylogenetic relationships of these two lineages with gnathostomes (jawed vertebrates) remained a controversial issue for over a century. However, recent molecular phylogenetic analysis using a large number of genes suggested that hagfishes and lampreys form a monophyletic group, the Cyclostomata (Fig. 1; Blair and Hedges, 2005; Delarbre et al., 2002; Delsuc et al., 2006; Furlong and Holland, 2002; Kuraku et al., 1999; Mallatt and Sullivan, 1998; Stock and Whitt, 1992; Takezaki et al., 2003; also see Kuraku and Kuratani, 2006; reviewed in Meyer and Zardoya, 2003).

Although several attempts have been made to determine the Hox repertoires and their genomic organization of lamprey by cDNA isolation (Pendleton et al., 1993; Sharman and Holland, 1998; Takio et al., 2007; Takio et al., 2004) and genomic sequencing (Force et al., 2002; Irvine et al., 2002), not all Hox genes have been discovered so far and their genomic organization remains uncertain. In addition to a high GC-content in protein coding regions (Kuraku and Kuratani, 2006), available cDNA sequences are usually incomplete and often derived from multiple closely related species. These factors prevented a precise categorization of available sequences into paralogy groups. Targeted genomic sequencing of Hox clusters in a sea lamprey Petromyzon marinus succeeded in the identification of regions containing multiple Hox genes, but they did not encompass entire Hox clusters (Force et al., 2002; Irvine et al., 2002). A phylogenetic analysis using available sequences suggested that lamprey has at least three or four Hox clusters (Force et al., 2002; Irvine et al., 2002; Fig. 1), which might, at least partly, be the result of cluster duplications specific to the cyclostome or lamprey lineage (Fried et al., 2003). In contrast, a PCR survey of genomic sequences in the Pacific hagfish Eptatretus stoutii detected at most nine genes in a single paralogy group (Fig. 1), suggesting again the possibility of independent duplication(s) of Hox clusters in the cyclostome or hagfish lineage (Stadler et al., 2004). Since hagfishes and lampreys diverged apparently relatively early (more than 400 million years ago) during cyclostome evolution—(Blair and Hedges, 2005; Fig. 3. Processes of Hox cluster quadruplication based on currently available data. A possible scenario of quadruplication of Hox clusters was reconstructed based on parsimony criteria. For the gene repertoires of the hypothetical gnathostome ancestor (bottom), see Fig. 4. The cluster structure at the pre-1R/2R state (top) represents that of a hypothetical ancestor just before the cluster quadruplication that had already gained Hox14. It is notable that Hox14 is the only paralogy group of Hox genes that might have undergone two gene losses in the course of cluster quadruplication. The pattern of Hox cluster relationships we assume here is that the A and B and the C and D clusters are “sister clusters”, that is, [[A,B],[C,D]] (see Meyer, 1998). This assumption is confirmed by our parsimony reconstruction of patterns of cluster duplications using the hypothetical gnathostome ancestor: each of three possible scenarios of cluster duplications, that is, [[A,B],[C,D]], [[A,C],[B,D]], and [[A,D],[B,C]], requires nine, eleven, and ten steps of gene losses, respectively.
Kuraku and Kuratani, 2006), it would therefore not be surprising if these two major groups of extant jawless fishes would turn out to possess different organizations of Hox clusters.

The incomplete knowledge of the organization of Hox clusters in cyclostomes has implications for the understanding of the phylogenetic timing of the genome expansion event, the so-called two-round (2R) whole genome duplications (Ohno, 1970; Lundin, 1993; Sidow 1996; Dehal and Boore, 2005; McLysaght et al., 2002). In these two rounds of whole genome duplications, a single gene or gene array of a pre-vertebrate ancestral genomes, would have been duplicated into multiple (up to four) copies in derived vertebrates. If the quadruplication of Hox clusters was caused by this event (Fig. 3), it is expected that the size of other gene families as well should provide clues on how many Hox clusters cyclostomes should possess. The same line of reasoning was used in the initial investigations of the teleost-specific genome duplication (TSGD or 3R), which will be discussed later (Vandepoele et al., 2004; Hoegg et al., 2004; Meyer and Schartl, 1999; Taylor et al., 2001b, Wittbrodt et al., 1998). However, attempts to analyze the phylogenetic position of cyclostome genes have so far failed to provide a reliable timing of genome duplications in early vertebrate evolution (e.g., Escriva et al., 2002; also see Kuratani et al., 2002 for review), probably due to the lack of strong phylogenetic signal that would be required for resolving multiple events (two rounds of genome duplications and the cyclostome/gnathostome divergence) that occurred within a short period of time (Horton et al., 2003). In order to infer the timing of this genome expansion event reliably, larger numbers of gene families would need to be characterized in the future (see Kuraku et al., 2009 for a recent study on this issue). The ongoing genome sequencing project for Petromyzon marinus (URL: http://genome.wustl.edu/genome.cgi?GENOME=Petromyzon%20marinus) is expected to provide additional information about the timing of the 2R genome duplications and the evolution of the genomic organization of Hox clusters in different lineages of deuterostomes.

Apart from genomic structure of Hox clusters, recent studies of lamprey Hox genes provided an overview about the colinearity in gene expression (Takio et al., 2007; Takio et al., 2004). Based on the expression patterns of paralogy group 2-8, the lamprey exhibits spatial colinearity in the central nervous system and pharyngeal arches, whereas no significant temporal colinearity could be detected. This observation will have to be investigated by more solid genomic linkage data and the collection of expression patterns of more gene repertoires. There are no reports at all about expression patterns of Hox genes in hagfish, so far. The recent success of hagfish embryology is expected to change this in the near future (Ota et al., 2007).

Below, we discuss the evolutionary processes during the evolution of Hox clusters in gnathostomes by reconstructing ancestral cluster structures. We follow the methodology of parsimony reconstructions as was previously done (Hoegg and Meyer, 2005; Malaga-Trillo and Meyer, 2001; Meyer, 1998).

**Hox clusters in chondrichthyanas**

Cartilaginous fishes (chondrichthyes) diverged first among extant gnathostome lineages from the gnathostome stem lineage (Nelson, 1994; Janvier, 1996; also see Kikugawa et al., 2004 for recent molecular phylogenetic evidence). They are divided into elasmobranchii (sharks and rays) and holocephali (chimaeras; Fig. 4). For elasmobranchii, data of the horn shark Heterodontus francisci are available (Kim et al., 2000, Prohaska et al., 2004), and those of the elephant fish Callorhinus mili are available for holocephali (Venkatesh et al., 2007). One of the interesting genomic aspects of this group is the retention of Hox14 genes (Powers and Amemiya, 2004a; Venkatesh et al., 2007). In the horn shark, HoxA14 was found to be a pseudogene. It can safely be presumed that this gene was present still in the chondrichthyan ancestor (Fig. 4). It is also notable that in the holocephalian lineage as much as five genes were secondarily lost mainly in the Hox D cluster. Based on updated information for horn shark presented by Powers and Amemiya (2004b), in which the Hox C cluster is not fully reported, the structure of HoxA, B and D clusters of the horn shark is identical to that of the genomic reconstructions of the chondrichthyan and gnathostome ancestors (Fig. 4), except for the pseudogenization of HoxA14 gene. This possible retention of the ancestral gnathostome state by elasmobranchs needs to be confirmed by additional studies on the Hox C cluster and on other representative species in these lineages.

**The actinopterygian lineage: Pre-3R and Post-3R modes**

Almost 10 years ago extranumeral Hox clusters were first reported in teleosts (Amores et al., 1998; Prince et al., 1998; also see Hoegg and Meyer, 2005; Meyer, 1998; Meyer and Malaga-Trillo, 1999; Wittbrodt et al., 1998 for reviews). We now know that the Hox cluster duplicates arose by a whole genome duplication that occurred early in the evolution of teleost fishes (VandePoule et al., 2004; Christopheis and Venkatesh, 2004; Crow et al., 2006). This was previously also shown by the analyses of non-Hox gene families (Hoegg et al., 2004; Taylor et al., 2003; Taylor et al., 2001a). This event was formerly termed the “fish-specific genome duplication” (FSGD; reviewed in Meyer and Van de Peer, 2005). However, this event has been proven to be shared by teleost fish lineages only and excludes more basal actinopterygian fish lineages (Hoegg et al., 2004), therefore the term “teleost-specific genome duplication (TSGD)” more accurately describes this important evolutionary event.

In recent years, information on Hox gene cluster architecture became available for more teleost species. In the order Cypriniformes, the data of blunt snout bream Megalobrama amblycephala was published (Zou et al., 2007) and it closely resembles those of the zebrafish Danio rerio (data not shown in Fig. 1). For example, these two cypriniform species share the loss of the HoxDb cluster. Interestingly, all available data for the more derived Neoteleostei, namely, two pufferfishes (Tetraodon nigroviridis· Jaillon et al., 2004; Takitugu rubripes· Aparicio et al., 2002), and stickleback (Gasterostes aculeatus· Hoegg et al., 2007), medaka (Oryzias latipes· Kasahara et al., 2007; Kurosawa et al., 2006), and a haplochromine cichlid fish (Astastolatilapia burtoni· Hoegg et al., 2007; also see Thomas-Chollier and Ledent, 2008), showed that they shared the loss of the HoxDb cluster, as described previously (Hoegg and Meyer, 2005). In addition to sporadic small-scale gene losses that occurred recently in each lineage (e.g., loss of the hoxB7a gene in the medaka lineage), massive gene losses seem to have occurred just after the origin of Neoteleostei (Fig. 4).

The data for bichirs are only available for the most basal...
The actinopterygian lineage that represents the pre-3R (TSGD) state (Chiu et al., 2004; Ledje et al., 2002). For Polypterus palmas, only cDNA isolation was performed (Ledje et al., 2002), while the genomic sequences of the entire Hox A cluster is determined for Polypterus senegalus (Chiu et al., 2004). Only very few gene losses seem to have occurred around the actinopterygii/sarcopterygii split, though this might be partly due to insufficient information for non-teleost actinopterygians. Although the structure of the Hox B, C, and D clusters remain to be determined for this lineage, it seems probable that bichirs have retained a similar set of Hox gene repertoires to that of the hypothetical ancestor of the Actinopterygii, Euteleostomi, and Sarcopterygii. Currently, there is no information available yet about Hox genes in Acipenseriformes (sturgeons), Amiiformes (bow fins), and Semionotiformes (gars). Knowledge about the genomic situation of Hox genes in these basal fish lineages would provide valuable information about the transition from the pre-3R state (before the TSGD and the evolutionary diversification of teleosts) to the post-3R genomic architecture.

**Structural stability in the sarcopterygian lineage**

For the sarcopterygian lineage including tetrapods, there are some reports for early-branching taxa. In the Australian lungfish *Neoceratodus forsteri* and the Indonesian coelacanth *Latimeria menadoensis*, only fragmental data have been reported so far based on cDNA isolation (Koh et al., 2003; Longhurst and Joss, 1999). Recently, an overview of coelacanth Hox clusters was reported (Powers and Amemiya, 2004b), and its genome was found to contain highly similar Hox gene repertoires to those of tetrapods (Fig. 4). In the tetrapod lineage, four Hox genes including *HoxA14*, that is still present in the coelacanth, have been lost.
secondarily (Fig. 4). Especially, it is curious to see whether the lungfishes possess Hox14 genes in their genomes. In the lineages leading to western clawed frog (Xenopus tropicalis), two genes (HoxB13 and HoxD12) are thought to have been lost (Fig. 4), based on an in silico survey by Hoegg and Meyer (2005). For the chicken (Gallus gallus), a recent report on genomic annotation of Hox clusters describes that the assembly is not complete, and thus entire or partial coding regions of some Hox genes (HoxB2, HoxC4, HoxC5, HoxC13, HoxD1, HoxD8, HoxD9 and HoxD10) are still missing (Richardson et al., 2007). Among these genes, entire coding regions of two genes (HoxC4 and HoxC5) have not been found by our survey in available genomic and cDNA sequences (Fig. 4). These potential losses have to be confirmed with more complete genomic assembly in the future.

It remains an open question if phenotypic evolution and genomic evolution march to the beat of the same drummer. Interestingly, the hypothetical ancestral organization of tetraptods reconstructed from currently available information is identical to that still found today in the human genome (Fig. 4). In general it would appear that during almost 400 million years of tetrapod evolution only very few gene losses seem to have occurred. It appears as if the diversification of body plans in land vertebrates and a possible macroevolutionary trend towards increased complexity during the evolution of vertebrates was not accompanied by any (at least obvious major) changes in their Hox cluster architecture. Based on our preliminary analysis, however, the grey short-tailed opossum Monodelphis domestica might be an exception. In an in silico survey of its still somewhat preliminary genomic sequences (version MonDom5; URL, http://www.broad.mit.edu/mammals/opossum/; Mikelsen et al., 2007), only a single Hox gene (HoxC6) has been found so far in the HoxC cluster. But, the Hox A, B, and D clusters are almost identical to the situation in the human genome. We also investigated the genomic sequences of the duck-billed platypus Ornithorhynchus anatinus (Ensembl database: URL, http://www.ensembl.org/Ornithorhynchus_anatinus/), but it would be premature to derive any conclusions based on this incomplete data set.

It is possible, however, that the TSGD is not only temporally correlated but even causally linked to the significantly accelerated rates of diversification/speciation and increased levels of phenotypic complexity in teleost fish (e.g. Wittbrodt et al., 1998; Meyer and Scharl, 1999; for more detailed discussion). Future work on a possible relationship of Hox cluster architecture and phenotypic diversification and increased complexity should also attempt to quantify changes in regulatory regions as well as protein coding regions and numbers of Hox gene repertoires.

Conclusions

In contrast to invertebrates, all vertebrates have four or more Hox gene clusters. Vertebrate Hox clusters are peculiar in that the linkages of genes in clusters are never broken and that all genes in a cluster are transcribed in the same direction (Duboule, 2007). A possible exception is the opossum HoxC cluster, but more data are needed to confirm this hypothesis. Also for agnathans, more reliable information is eagerly awaited since it might shed light on the phylogenetic timing of the 1R and 2R genome duplication events during chordate evolution. For the gnathostomes our reconstruction of the genomic organizations of the Hox clusters during vertebrate evolution suggests that there are two major types of Hox cluster architectures: (1) the four-cluster type and (2) the post-3R teleost type. The former is further divided into three themes: (1A) an ancestral gnathostome type (still seen in the horn shark A, B, and D clusters); (1B) non-teleost bony fish type (as found in the coelacanth and probably in the bichir); and (1C) the tetrapod type (e.g. human). To date, there is no solid genomic data about the organization of Hox gene clusters in some crucial lineages, such as hagfishes, lampreys and reptiles, and marsupials. These data are needed urgently for a more complete understanding of the patterns and processes of Hox cluster evolution in deuterostomes.

Knowledge of expression patterns of Hox genes might help in the understanding of the evolutionary history, mechanisms and constraints that shaped Hox cluster evolution. Expression information would also aid in the identification of potentially homologous morphological structures among species that belong to phenotypically extremely diverged lineages of vertebrates.

Acknowledgements

This study was financially supported by the Deutsche Forschungsgemeinschaft (DFG). We thank Michael Richardson, Shigeru Kuratani, Joost Woltering, Byarra Venkatesh, Simone Hoegg, and Jenny Graves for helpful comments and discussion.

References


Further Related Reading, published previously in the *Int. J. Dev. Biol.*

See our Special Issue *Evolution & Development* edited by Jaume Baguñà and Jordi García-Fernández at:

See our Special Issue *Fertilization*, in honor of David L. Garbers and edited by Paul M. Wassarman and Victor D. Vacquier at:
http://www.ijdb.ehu.es/web/contents.php?vol=52&issue=5-6

*From the American to the European amphioxus: towards experimental Evo-Devo at the origin of chordates*
Jordi Garcia-Fernàndez, Senda Jiménez-Delgado, Juan Pascual-Anaya, Ignacio Maeso, Manuel Irímiia, Carolina Minguillón, Elia Benito-Gutiérrez, Josep Gardenyes, Stéphanie Bertrand and Salvatore D’Aniello
*Int. J. Dev. Biol.* in press doi: 10.1387/ijdb.072436jg

Peter Holland, *homeobox genes and the developmental basis of animal diversity*
Sebastian M. Shimeld
*Int. J. Dev. Biol.* (2008) 52: 3-7

*Cloning and pattern of expression of the shiro-uo vasa gene during embryogenesis and its roles in PGC development*
Akimitsu Miyake, Taiju Saito, Tohru Kashiwagi, Daisuke Ando, Akitsugu Yamamoto, Tohru Suzuki, Norio Nakatsuji and Takako Nakatsuji
*Int. J. Dev. Biol.* (2006) 50: 619-625

*Geometry and mechanics of teleost gastrulation and the formation of primary embryonic axes*
Elena M. Cherdantseva and Vladimir G. Cherdantsev

*Hox and ParaHox genes in Nemertodermatida, a basal bilaterian clade*
Eva Jiménez-Guri, Jordi Paps, Jordi García-Fernández and Emili Saló
*Int. J. Dev. Biol.* (2006) 50: 675-679

*Pulling forces acting on Hox gene clusters cause expression collinearity*
Spyros Papageorgiou
*Int. J. Dev. Biol.* (2006) 50: 301-308

*Evolution of the Hox/ParaHox gene clusters.*
David E K Ferrier and Carolina Minguillón
*Int. J. Dev. Biol.* (2003) 47: 605-611

*Colinearity and non-colinearity in the expression of Hox genes in developing chick skin.*
Alasdair I Reid and Stephen J Gaunt

*Novel interactions between vertebrate Hox genes.*
M H Hooiveld, R Morgan, P in der Rieden, E Houtzager, M Pannese, K Damen, E Boncinelli and A J Durston
*Int. J. Dev. Biol.* (1999) 43: 665-674

*Estimation of Hox gene cluster number in lampreys.*
A C Sharman and P W Holland
*Int. J. Dev. Biol.* (1998) 42: 617-620