Hox, ANTP and Homeobox Gene Evolution in Metazoa: Insights from Cnidaria

Von der Naturwissenschaftlichen Fakultät
der Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades eines
Doktors der Naturwissenschaften
Dr. rer. nat.
genehmigte Dissertation
von

Dipl.-Biol. Kai Kamm
geboren am 02.01.1971 in Mayen

2007

Referent: Prof. Dr. Bernd Schierwater
Institut für Tierökologie und Zellbiologie
Stiftung Tierärztliche Hochschule Hannover

Korreferent: Prof. Dr. Hans-Jörg Jacobsen
Institut für Pflanzengenetik
Gottfried Wilhelm Leibniz Universität Hannover

Tag der Promotion: 19.07.2007

Die größten Schwierigkeiten liegen da, wo wir sie nicht suchen.

Johann Wolfgang von Goethe,
Wilhelm Meisters Wanderjahre

Aber das ahnungsvolle, Jahre währende Suchen im Dunkeln mit seiner gespannten Sehnsucht, seiner Abwechslung von Zuversicht und Ermattung und seinem endlichen Durchbrechen zur Klarheit, das kennt nur, wer es selber erlebt hat.

Albert Einstein,
1933

„Wenn jemand sucht“, sagte Siddhartha, „dann geschieht es leicht, daß sein Auge nur noch das Ding sieht, das er sucht, daß er nichts zu finden, nichts in sich einzulassen vermag, weil er nur immer an das Gesuchte denkt, weil er ein Ziel hat, weil er vom Ziel besessen ist. Suchen heißt: ein Ziel haben. Finden aber heißt: frei sein, offen stehen, kein Ziel haben.“

Hermann Hesse,
Siddhartha

Contents

Zusammenfassung

Homeobox-Gene kodieren für Transkriptions-Faktoren, die die Genexpression während der Entwicklung und der Zellproliferation in Metazoen steuern. Die größte Gruppe der Homeobox-Gene ist die ANTP-Klasse, die aus den Hox-, „extended Hox“ und den NK-artigen Genen besteht. Von besonderem Interesse sind die Hox-Gene, weil sie wahrscheinlich eine zentrale Rolle bei der Entstehung und Diversifizierung von Bauplänen im Tierreich gespielt haben. Während der Embryonal-Entwicklung der Bilateria determinieren sie die Ausbildung von Körperstrukturen und Geweben entlang der Anterior-Posterior-Achse. Ihr kennzeichnendes Merkmal ist, dass sie in Clustern organisiert sind und ihre Reihenfolge auf dem Chromosom mit ihrem Expressionsmuster entlang der Körperachse übereinstimmt.

Die Evolution von Entwicklungs-Mechanismen scheint eng mit der Evolution der Homeobox-Gene verbunden zu sein. Ein besseres Verständnis vom Ursprung des Hox-Systems, der ANTP-Klasse und der Homeobox-Gene, kann uns deswegen wichtige Einblicke geben über die genetischen Mechanismen, die für die Radiation der Tierstämme verantwortlich sind. Während viel über die Homeobox-Gene und das Hox-System höherer Tiere bekannt ist, haben wir sehr wenig Informationen von basalen Taxa. Der Stamm der Nesseltiere ist deshalb wichtig für die Aufklärung der Evolution dieser Gene, weil der Bauplan der Cnidaria eine Schlüsselstellung einnimmt in Richtung zunehmend komplexerer Formen bei den Bilateria. In der vorliegenden Arbeit wurde versucht: (I) aufzuklären, ob Cnidaria bereits ein echtes Hox-System besitzen, (II) die Komplexität der ANTP-Klasse innerhalb der Cnidaria zu ermitteln, und (III) aus den Ergebnissen Schlussfolgerungen für die Evolution des Hox-Systems, der ANTP-Klasse und der Homeobox-Gene im Allgemeinen zu ziehen.

Die Ergebnisse zeigen, dass ein echtes Hox-System wahrscheinlich erst nach der Trennung der Cnidaria und Bilateria entstanden ist und deswegen eine Synapomorphie der Letzteren ist. Obwohl die Cnidaria mehrere Gene aufweisen, die mit Hox-Genen verwandt sind, besitzen sie kein vergleichbares System hinsichtlich Sequenz-Homologie, genomischer Organisation und konservierter Funktion. Einzelne Gene zeigen eine hohe Übereinstimmung mit echten Hox-Genen höherer Tiere, was zwar auf einen gemeinsamen Ursprung dieser Gene, aber auch auf eine Trennung beider Linien deutet, bevor ein ausgereiftes Hox-System entstanden ist. Dagegen zeigte das nicht-Hox ANTP-Gen-Repertoire der Cnidaria eine erstaunliche Diversität, die mit der höherer Tiere vergleichbar ist. Die Diversifizierung der ANTP-Gen-Familien scheint also bereits abgeschlossen gewesen zu sein, bevor ein Hox-System entstanden ist. Die Analyse von Homeobox-Genen, die in *Eleutheria* und *Nematostella* benachbart liegen, unterstützt zudem die Hypothese, dass die Mehrheit der ANTP-

Gen-Familien, und darüber hinaus auch die Vorläufer der meisten Homeobox-Klassen, einer gemeinsamen ancestralen genomischen Region entsprungen sind.

Schlüsselwörter: Hox, Homeobox, ANTP, Cnidaria

Abstract

Homeobox genes code for important transcription factors that regulate gene expression during development and cell proliferation in metazoan animals. By far the most diverse group of homeobox genes is the ANTP class which comprises the Hox, the “extended Hox” and NK-like genes. Of particular interest are the Hox genes because they presumably have played a crucial role in the diversification of metazoan bauplan patterns. Hox genes are responsible for patterning most or all tissues along the anterior-posterior axis of bilaterian animals and their defining characteristic is that they are typically organised in clusters in which gene order directly reflects expression domains along the body axis.

The evolution of developmental mechanisms seems intimately connected with the evolution of homeobox genes. Understanding the origin of the Hox system, the ANTP class, and homeobox genes in general, is therefore of particular interest for understanding the genetic mechanisms that may have been deployed for the radiation of metazoan phyla. Whilst much has been known about homeobox genes and the Hox system in higher animals, we lack important information from more basal taxa. The phylum Cnidaria is therefore most valuable for trying to elucidate aspects of the evolution of these genes because it represents a key transition regarding animal complexity towards the Bilateria. In the present work it was tried to: (I) clarify if Cnidaria already possess an equivalent of the bilaterian Hox system, (II) assess the complexity of the ANTP class in Cnidaria, and (III) evaluate resulting implications for the evolution of the Hox system, the ANTP class of homeobox genes, and the homeobox genes in general.

The results indicate that a true Hox system postdates the split between Cnidaria and Bilateria and is therefore a synapomorphy of the latter. Although some genes related to Hox classes of higher animals are present in Cnidaria they do not possess an equivalent of the Hox system in terms of sequence identity, genomic organisation or conserved function. Nevertheless, single Cnidarian genes show high affinities to true Hox classes of higher animals, indicating common origin of particular genes but possible secession before a complete Hox system evolved. Conversely, the non-Hox ANTP homeobox gene complement in Cnidaria was found to be as complex as in Bilateria, indicating that the diversification of the ANTP gene families took place before a Hox cluster emerged. Moreover, analyses of homeobox genes that were found to be linked in *Eleutheria* and *Nematostella* provided support for the hypothesis that most ANTP gene families, and probably also the predecessors of all metazoan homeobox classes, have once emerged from a common ancestor region.

Keywords: Hox, Homeobox, ANTP, Cnidaria

1. Introduction

1.1 The Hox system

Introduction to the Hox system

The famous “Pierre de Rosette” or “Rosetta Stone” was discovered in 1799 during Napoleon’s Egyptian Campaign. It yielded a most valuable secret: The same text is engraved in Greek and in Demotic and Egyptian hieroglyphs (Figure 1). Before that discovery, scientists had not been able to unravel the underlying plan or code of the hieroglyphic figures. By comparing letters and figures of the Rosetta texts, the French linguist Champollion was able to deduce the basic grammar and meanings of the first hieroglyphs which subsequently led to the decoding of the ancient scriptures.

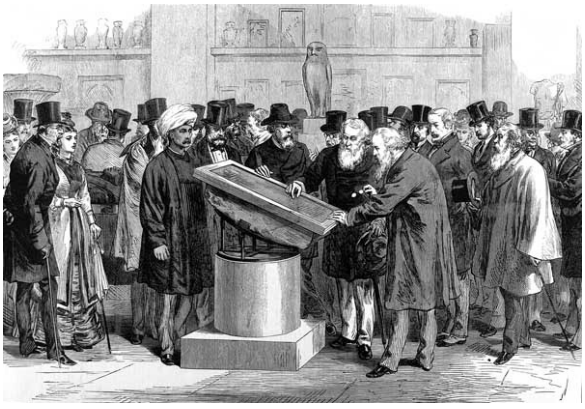


Figure 1: Experts inspecting the Rosetta Stone during the International Congress of Orientalists of 1874. Illustrated London News.

The discovery of the Hox system [1] in metazoan animals has often been compared with the Rosetta Stone, because likewise it seemed to provide a blueprint for unravelling the genetic basis of metazoan bauplan diversity. The products of the Hox genes are transcription factors which provide positional information along the anterior-posterior axis during development [2]; and they perform this task in a striking similar fashion across bilaterians. The presence or absence of expression of a particular (set of) Hox gene(s) defines body regions in Bilateria. This is called the Hox code, and similar to the code underlying the Rosetta hieroglyphs, the same or a similar code underlies the development of all bilaterians studied so far. It seems as if evolution had deployed a “platform strategy” for animal design.

Hox genes belong to the homeobox genes which are a large and diverse family of transcription factors characterised by the presence of a conserved 180bp sequence encoding a DNA binding motif - the homeodomain [3]. All homeobox genes play important roles in metazoan development and cell proliferation, making them important tools for studying the evolution of genomes, bauplans and developmental patterns.

One of the defining characteristics of Hox genes in particular is that they are typically organised in clusters in which genomic organisation directly reflects domains of expression along the anterior-posterior body axis (Figure 2): genes at one end of the cluster pattern the anterior end of the embryo, those genes at the opposite end pattern the posterior end. This is referred to as spatial collinearity [4]. In vertebrates, which show very tight clusters, there is also a temporal aspect: the genes are turned on successively from one end to the other, reflect-

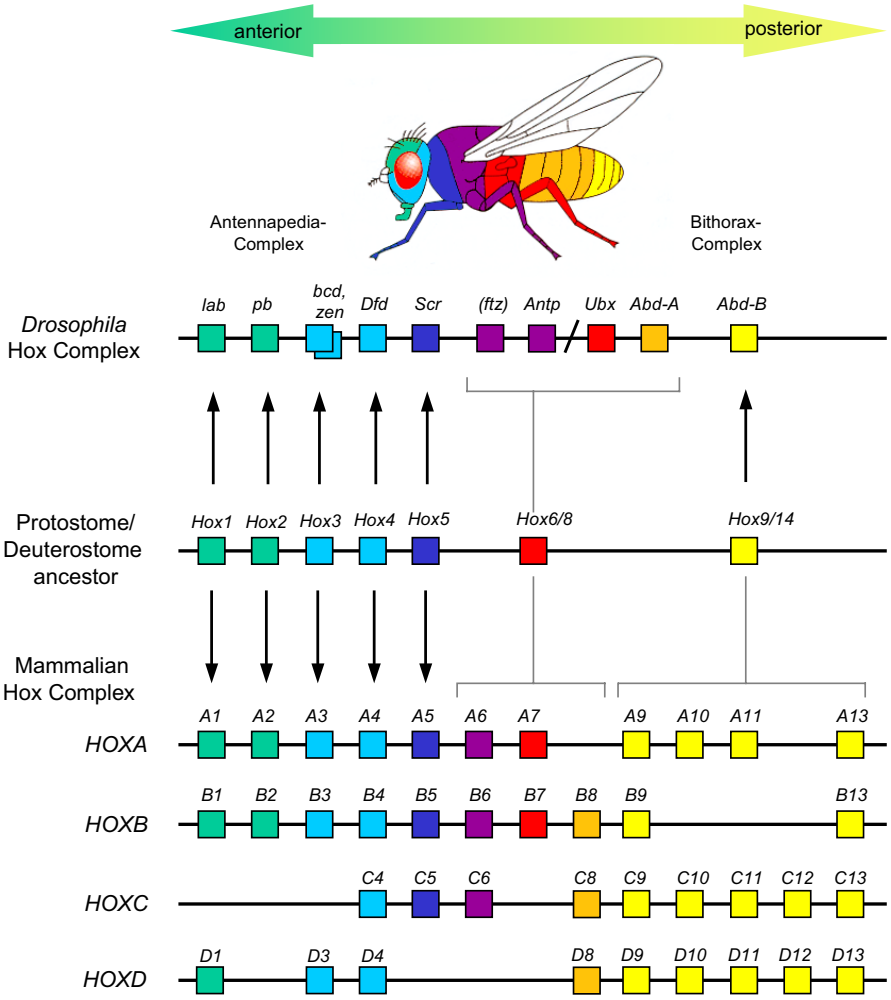
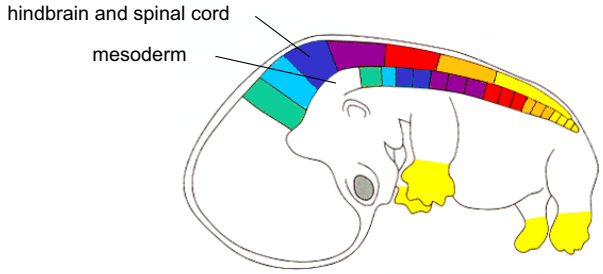


Figure 2: The Hox system in bilaterian animals: The *Drosophila* antennapedia and bithorax complex in chromosomal order above. The corresponding genes of the four mammalian complexes below. A possible cluster of the last common ancestor of protostomes and deuterostomes in the middle. Expression domains of particular genes are colour coded and simplified (details of expression patterns depend on developmental stage and the expression domains of many genes overlap with those of more anterior or posterior genes). The *Drosophila* Hox complex is split into the antennapedia and bithorax complexes, both lying on chromosome 3. In the lineage of mammals (and most vertebrates) the Hox complex was duplicated twice yielding four Hox clusters (modified from [5, 6]).



ing enhancer sharing and common regulatory mechanisms - this mode of expression seems to be crucial for slowly developing animals like vertebrates [4].

Amongst chordates, a correlation exists between increasing Hox complement and extent of morphological variation. Like other invertebrates, *Amphioxus* has a single Hox cluster whilst most vertebrates have four [7]. In the bony fish, the morphologically most diverse vertebrate group, an extra round of duplication followed by losses has resulted in the seven clusters of pufferfish [8] and zebrafish [9] (if we exclude tetraploid bony fishes with up to fourteen [10]). In arthropods, differential Hox gene use underlies much of the extensive variation on the same basic

body plan [11]. These observations, together with their central role in axial patterning, have led to the assumption that much of the variation within bilaterian phyla can be attributed to different numbers of Hox clusters or differential use of the Hox system. Hence, the Hox cluster seems to be one of the key inventions that have driven the radiation of bauplan patterns.

Since the discovery that animals related as far as arthropods and mammals share the same regulatory mechanisms for their developing body axes, the question has been: how far can we trace back the origin of this system? The find that even diploblasts, such as cnidarians, possess genes related to bilaterian Hox genes [12-14] has galvanised the evo-devo community because cnidarians are amongst the most simple organised metazoans; concurrently they represent a key transition regarding animal complexity towards the Bilateria. Subsequently, several efforts have been made to unite all metazoan animals by the possession of the Hox system [15, 16]. In this view the Hox cluster would be a synapomorphy defining the Metazoa.

The paradigm of an “ancient Hox system” has been challenged and controversially discussed by many authors [6, 16-19], notably because of the uncertain nature of cnidarian Hox-like genes: the phylogenetic relation of cnidarian Hox-like genes to Hox classes of higher animals, as well as their presumable function in axial patterning is still speculative and almost nothing is known about their genomic organization. However, this clarification is crucial for deciding if these genes represent Hox genes *sensu stricto*; or rather independent offspring of ancient genes that were also the predecessors of bilaterian true Hox classes. Nevertheless, hypotheses about the origin of the Hox system have been mainly deduced from the vast evidence available for Bilateria, regardless of the scarce information about cnidarian Hox-like genes.

Hypothesis about the origin of the Hox system: The ParaHox hypothesis

The genes of the bilaterian Hox system are commonly subdivided into the anterior, central and posterior Hox classes. Anterior genes (Hox 1-5) are sometimes referred to as the “head genes”, central genes (Hox 6-8) as the “trunk genes” and posterior genes (Hox 9-13/14) as the “tail genes”. However, classifications based on more elaborate phylogenetic analyses [16, 20] divide bilaterian Hox genes into the four groups anterior (Hox 1-2), group 3 (Hox 3), central (Hox 4-8) and posterior (Hox 9-13/14). In either case, all these genes are typically arrayed in bilaterian Hox clusters and it is assumed that they have originated by successive duplications from a single ancestor gene [6, 19]

Another group of Hox related genes has puzzled scientists for many years: The genes belonging to the Gsx, Xlox and Cdx classes show a close relationship to the Hox genes but do not reside within Hox clusters. In 1998 Brook et al. [21] discovered that these genes build a sin-

gle cluster in the cephalochordate *Amphioxus* (*Amphioxus* possesses only a single Hox cluster and single orthologs of Gsx, Xlox and Cdx). Furthermore, each of the three genes seems to be more closer related to the anterior, group 3 and posterior Hox classes, respectively, than to each other. Accordingly, it seemed reasonable to deduce that this so called “ParaHox” cluster is the evolutionary sister (the “paralog”) of the Hox cluster; and both must have originated from an ancestral (Proto)Hox cluster, consisting of three to four genes, that duplicated to give rise to distinct Hox and ParaHox clusters (Figure 3). If correct, this means that each metazoan group possessing distinct orthologs of Hox and ParaHox genes must also have possessed a true Hox system in its evolutionary history [16].

Surprisingly, the Hox-like genes present in diploblasts include a Gsx ortholog (Diplox-2) in Cnidaria and Placozoa [20, 22, 23], and some authors even claim the presence of a Cdx ortholog in Cnidaria [16]. Furthermore, there is some agreement that cnidarians possess genes related to bilaterian anterior and posterior Hox classes, while no study has ever revealed genes related to group 3 Hox, its “ParaHox counterpart” Xlox, or central Hox classes [6].

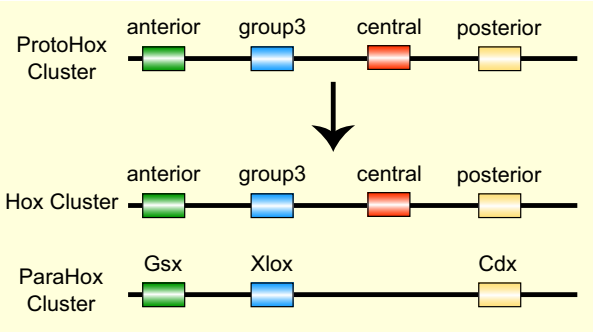


Figure 3: Origin of Hox and ParaHox clusters according to the ParaHox theory. An ancestral Proto cluster, harbouring the predecessors of the four Hox classes, duplicated to distinct Hox and ParaHox clusters. Central genes were either lost in the ParaHox cluster, or emerged after the duplication event in the Hox cluster only. This assumption is mainly based on the observation that, for example, the Gsx gene seems more closely related to the anterior Hox than to Xlox or Cdx. These sister relationships are, however, not unambiguously supported in phylogenetic analyses.

The consensus view over the last years has thus been that the three missing gene classes have either escaped the surveys or that cnidarians simply lost them; but the claim that cnidarians have distinct Hox and ParaHox genes has been used as a proof that the common cnidarian/bilaterian ancestor already must have had a simple but true Hox cluster [16]. A crucial weak point in such assumptions is that this is rather indirect proof of an ancient Hox system and that the fate of the missing genes is highly speculative. Moreover, the assignment of a sister relationship of Gsx, Xlox and Cdx to the anterior, group 3 and posterior Hox classes, respectively, is far from confident [6]. More information about cnidarian Hox-like genes is urgently needed. A closer look at the Hox system in Bilateria might help to see what we are actually looking for.

A closer view: The Hox system of the Bilateria

Within vertebrates and cephalochordates we typically observe tight and uninterrupted clusters - from the single 500Kb cluster in *Amphioxus* to the 100-150kb clusters of vertebrates [7, 24] - and spatiotemporal expression of Hox genes

along the body axis. In urochordates there seems to be a trend for disintegration of the cluster. However, urochordate Hox genes show high sequence identity to all four Hox classes and despite the cluster having fragmented their expression patterns show a persistent anterior-posterior orientation, reminiscent of an ancestor with an uninterrupted cluster [25-27].

Another derived deuterostome Hox cluster can be seen in the echinoderm *Strongylocentrotus* where the single cluster spans 600Kb with a highly derived gene order - the central genes lie in reverse order at the end of the cluster - suggesting several re-arrangements. But again we observe collinear expression of central and posterior Hox genes in the somatocoel of the developing larvae - in a pentameral animal that has emerged from a bilateral ancestor [28-30].

In protostomes we find not more than one cluster and it seems that the clusters in different phyla have a greater tendency to disintegrate. The intergenic distances of insect Hox clusters are generally bigger (as compared to vertebrates) and, in the case of the drosophilids, the cluster is split into two pieces with different breakpoints in different species [31]. That this is a derived condition within insects can be seen in *Schistocerca gregaria* and *Anopheles gambiae*, both of which have a single and uninterrupted cluster [32, 33]. Nematodes again show a high level of divergence from the primitive bilaterian Hox condition. *Caenorhabditis elegans* has lost group 3 genes, in contrast to other nematodes, and its cluster is organized in three pairs of Hox genes spanning a distance of 5Mb on one chromosome [34-36].

The Hox systems of lophotrochozoan taxa yet have to be examined more detailed, but it is already evident that all four Hox classes are present. Some of the genes were shown to be linked in the nemertine *Lineus* and the polychaete *Nereis*, whereas collinear expression of particular genes has been observed in molluscs, suggesting that clustering and collinear expression of Hox genes is the primitive condition in Lophotrochozoa [37-40].

The presumably most primitive Hox system can be seen in the acoelomorph flatworms. This group, previously placed into the platyhelminthes, is considered to be the most basal bilaterian clade by some authors [41]. Their Hox complement consists of only four genes representing the four Hox classes [41, 42], though nothing is known so far about the genomic organization of these. However, it is compelling that this represents the primitive Hox complement of the last common bilaterian ancestor and that acoelomorphs diverged from other Bilateria before the more elaborate clusters, consisting of 8-10 genes in protostomes and up to 14 genes in deuterostomes, emerged.

The highly conserved spatial collinear expression of Hox genes in animals as distantly related as *Drosophila* and vertebrates suggests that this is the primitive condition in Bilateria; and from the data available today it is evident that a true (but

simple) Hox cluster was present in the last common bilaterian ancestor. Secondary disruption of bilaterian Hox clusters seems to be correlated with a very rapid embryogenesis in particular lineages [4]. In this view the temporal collinearity is the constraining force that keeps the cluster intact. The switch to a rapid developmental mode may simply not allow the successive activation of Hox genes, which in turn eliminates the need for maintaining a tight and uninterrupted cluster. However, throughout extant Bilateria we have vast data of more or less tight and uninterrupted clusters; for the other cases we have compelling reasons to regard this as a derived condition because either (Figure 4):

1. other members of a particular clade have uninterrupted clusters
2. clear orthologs of all four Hox classes are present
3. expression domains of Hox genes obey (at least in part) spatial collinearity

Elucidating the origin of the Hox system

The situation in Bilateria seems quite clear: the UrBilaterian ancestor must have possessed a true Hox system, regardless of individual gene losses and cluster disintegration in some extant species. Conversely, the situation in more basal metazoans remains unclear. Cnidarians possess genes related to bilaterian Hox and ParaHox genes which are here referred to as Hox-like genes. One must be cautious, however, in assigning a Hox system to a particular group just because some genes show affinities to Hox class genes. Similarly, for example, if we compared morphological features, like the tetrapod limbs and the lobe fins of the coelacanth. The coelacanth lobe fins are homologous to the tetrapod limbs but obviously cannot be regarded as such. Coelacanths split off the lineage leading to tetrapods before the typical five digit tetrapod limbs evolved. Therefore, we also have to distinguish between what we consider a true

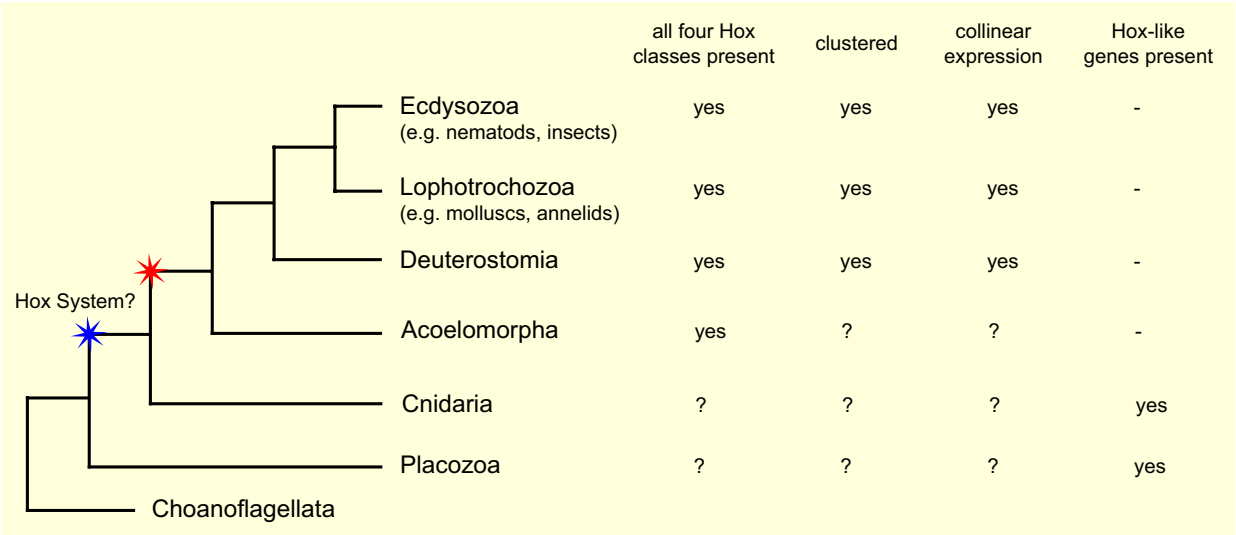


Figure 4: Simplified phylogenetic tree of the major metazoan clades rooted with Choanoflagellata, adapted from [41, 43, 44]. At the base of Bilateria we have evidence that a canonical Hox system was (at least primitively) present. The data available from more basal taxa do not yet allow to decide whether a true Hox system originated even earlier.

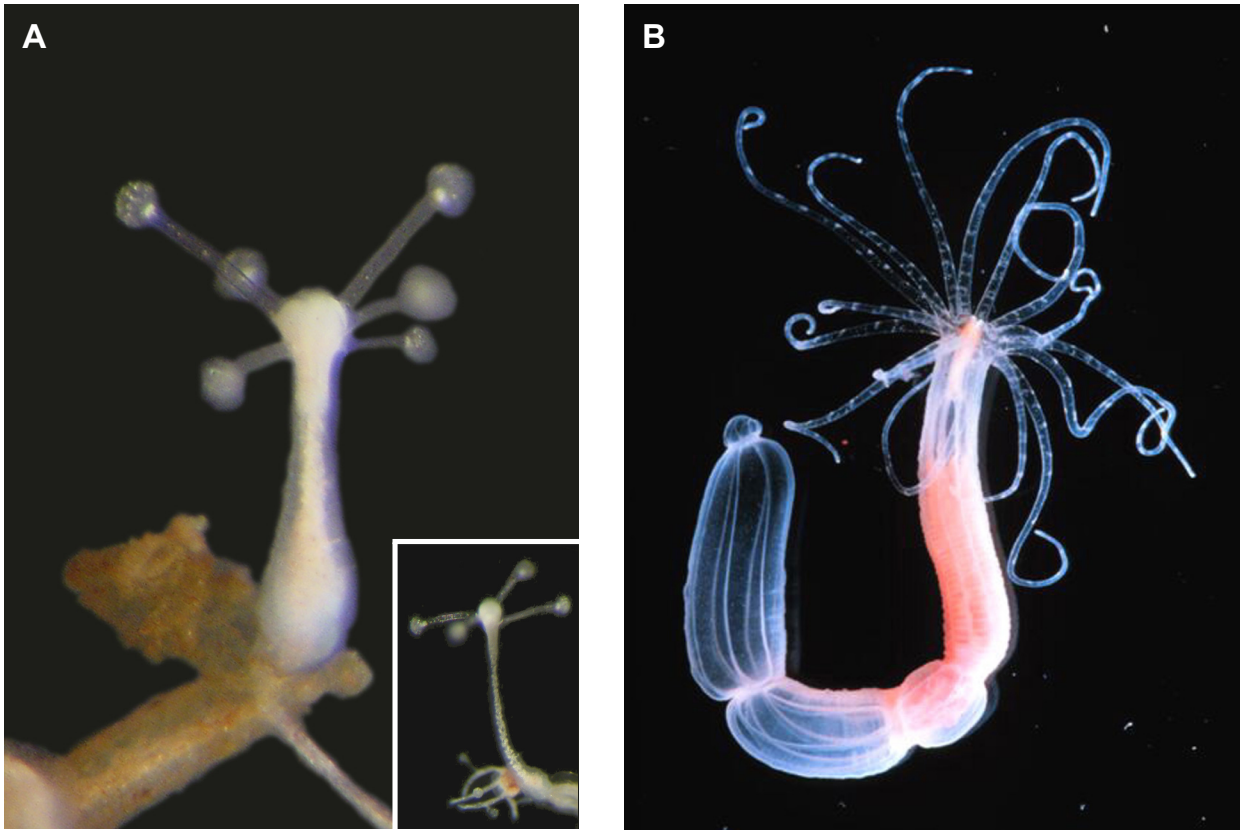


Figure 5: A Polyp of the hydrozoan *Eleutheria dichotoma*. Inserted picture: Polyp budding a medusa at its basis **B** The starlet sea anemone *Nematostella vectensis* (picture credit for B: Institut für Zoologie und Limnologie, Leopold-Franzens-Universität Innsbruck)

Hox system and genes that are merely related to members of this system, reflecting common origin of particular genes but possible secession before a mature system evolved. For this purpose it is necessary to define a true canonical Hox system as a full set of linked and interacting homeobox genes that are directly related to the Hox classes (anterior, group 3, central and posterior) of *Drosophila* and mammals and that, through their combined actions, are primarily responsible for patterning most or all tissues along the anterior-posterior body axis.

In an attempt to clarify the evolutionary origins of the Hox system, the Hox-like genes in two representative cnidarians, *Eleutheria dichotoma* (Hydrozoa, Figure 5A) and *Nematostella vectensis* (Anthozoa, Figure 5B), were thus characterized in terms of sequence relationships, genomic organization and expression patterns. *Eleutheria* is a typical cnidarian in having both polyp and medusa life cycle stages. It was the first cnidarian representative in which Hox-like genes were identified [12]. Genomic organisation of *Eleutheria* Hox-like genes was characterised by means of genomic fosmid libraries and expression patterns by RNA in situ hybridization (*Eleutheria* expression data provided by Dr. Wolfgang Jakob). *Nematostella*, on the other hand, represents the basal cnidarian class (Anthozoa) and several Hox-like genes have been already identified in this cnidarian [45]. The now fully sequenced, though unassembled, *Nematostella*

genome was screened for Hox-like genes and their genomic organisation was determined by assembling genomic contigs from the trace data available at the NCBI data bases. Phylogenetic analyses were conducted to infer relationships of *Eleutheria* and *Nematostella* Hox-like genes and bilaterian Hox/ParaHox classes and expression patterns for *Eleutheria* were compared with those from the literature [45, 46].

1.2 A broader view: Evolution of the ANTP class

Hox/ParaHox genes themselves belong to the ANTP (super)class of the homeobox genes [20, 47, 48] which also comprises the “extended Hox” and the NK-like (NKL) genes and seems to be restricted to the Metazoa [49, 50]. Besides the Hox/ParaHox genes, also many other members of the ANTP class are involved in important developmental programs. For example, members of the Emx gene family participate in vertebrate forebrain and midbrain development [50] and many NK genes play a role in the specification of muscle cell lineages. A well studied example are the NK 2.5 orthologs in *Drosophila* and mouse which play a crucial role in cardiogenesis in both species [48].

Linkage analyses in bilaterian genomes indicate that clustering was once a characteristic not only of Hox but of most ANTP homeobox genes (reviewed in [19]). The coordinated expression of Hox genes during development might be the major constraining force that keeps this particular cluster intact whilst such constrains seem to be much lower for the non-Hox ANTP genes - which subsequently led to their dispersal in the genomes of extant species [4, 19, 51]. Nevertheless, remnants of clusters still exist in many bilaterian animals [51-53] and sequence relationships among the ANTP genes, combined with comparative gene mapping in the genomes of *Drosophila*, human, mouse and *Amphioxus*, suggest that the ancestors of all ANTP gene families have once emerged by several cis-duplications from an ancestral ANTP mega-array [52-54]. Hence, the evolution of the Hox genes can only be fully understood in the broader context of the evolution of all ANTP genes - and vice versa.

Earlier studies revealed that cnidarians have clear orthologs to some non-Hox ANTP genes such as even-skipped, Emx, Hex, Not or Dlx [20], whilst the phylum Placozoa, which is most likely even more basal [44, 55, 56], features a very low diversity of these genes [57]. At least some steps of the expansion of the ANTP class must have taken place in the cnidarian bilaterian ancestor but only an almost complete assessment of cnidarian non-Hox ANTP genes can provide insight into the origin and evolution of these genes.

Recent analyses indicate that the last common ancestor of cnidarians and bilaterians already had a very complex genome which has been retained in basal cnidarians like *Nematostella* and *Acropora* [58-60]. For this rea-

son, the full, though unassembled, *Nematostella* genome was screened for non-Hox ANTP genes and phylogenetic relationships to their bilaterian counterparts were inferred. By taking advantage of existing linkage information from the literature and genome data bases a possible ANTP gene complement of the last common ancestor of Cnidaria and Bilateria was deduced.

1.3 Ancient linkage of distantly related homeobox genes: The case of POU homeobox genes

The homeobox genes present in Metazoa can be mainly assigned to the ANTP, PRD, POU, LIM, CUT, prospero, TALE and SIX classes, most of which seem to be restricted to Metazoa [47, 48, 50]. Homeobox genes of the atypical TALE-homeobox class, however, are also found in plants and fungi [61], while a homeobox gene possibly related to the metazoan LIM-homeobox class is present in the slime mold *Dictyostelium* [47].

By far the most diverse group of homeobox genes is the ANTP class which has been already introduced above [19]. Another important group of homeobox genes is the POU class which was first identified in the mammalian transcription factors **P**it-1, **O**ct-1 and Oct-2 and in the nematode **U**nc-86 factor (hence the name **POU**) sharing a novel domain N-terminal to the well known homeodomain [62, 63]. The bipartite POU domain thus consists of the N-terminal POU specific domain (POU_S / ~75aa) and the POU homeodomain (POU_H / 60aa) which are connected by a highly variable linker - and both sub-domains are involved in recognizing and binding of target DNA sequences [48, 62, 63]. POU genes play important roles in many developmental systems, notably in the nervous system [64, 65], and they have been detected in virtually all metazoan taxa, including Cnidaria and sponges [66-69]. Based on sequence similarity of the complete POU domain six subclasses are commonly recognized [62].

Phylogenetic analyses based on sequence identity (Figure 6) suggest that the ANTP and PRD classes are closer related to each other than to the remaining classes, whereas, for example, LIM and POU are more diverged [47, 48, 50]. A possible scenario has been proposed [47] in which a LIM-like ancestor, possibly related to

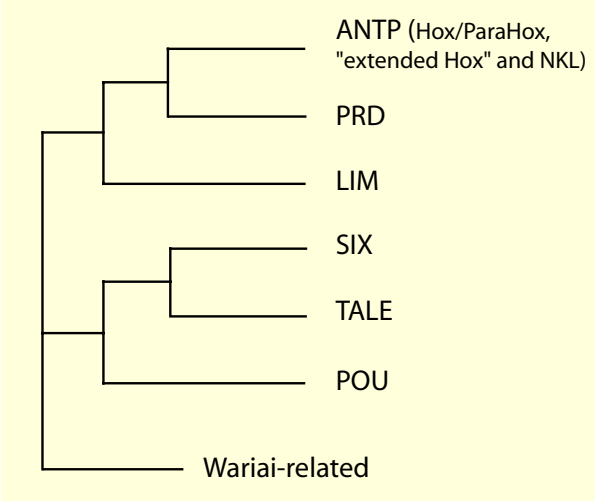


Figure 6: Simplified phylogenetic relationships between major metazoan homeobox classes deduced from analyses by Galliot et al., 1999 [47]. Note that the term “class” is not strictly used throughout the literature. For example, whereas all the ANTP genes are regarded as one class, there are also the different “Hox classes”. Therefore, the term ANTP superclass has also been suggested. Likewise, the POU class of homeobox genes consists of several (sub)classes (e.g. POU class 6 genes).

the *Dictyostelium* Wariai HD, gave rise to the ancestors of LIM and PRD classes [47]. Close to the base of PRD origin the ANTP class could have originated whilst the remaining classes, for example the SIX and POU genes, may have emerged even earlier in this scenario.

If similar genes are linked in the genome, co-evolution or common origin by duplication seems likely. This is evident for the clustered Hox genes in Bilateria which are thought to have originated by successive duplications from a single ancestor gene [6, 19]. However, if there were no functional constrains for maintaining of linkage - like coordinated expression - then transduplication events, gene loss and chromosomal breakage are likely to occur which might disperse genes that once have co-evolved. Hence, ancient linkage of genes is not easy to assess in extant species, especially for genes that are less closely related and have diverged for a longer time - like Hox and NK genes, or even Hox and POU genes.

Nevertheless, comparative gene mapping has shown that most likely the ANTP class genes have emerged and co-evolved once on a single ancestral genomic region - the proposed ANTP mega-array - before they were separated [51-54]. Similarly, analyses of the human and mouse genomes also suggest that most of the remaining metazoan homeobox classes possibly derived from a second common ancestor region - which was named the contraHox superparalogon [70].

The question remains whether there has been a link between the two proposed ancestor regions of metazoan homeobox genes - one containing most of the ANTP class genes, the other containing the PRD, LIM, POU, SIX, CUT, TALE and prospero classes. We might expect such a link because the PRD and ANTP classes seem much closer related than, for example, PRD and POU genes [47]. However, the founding event of the ANTP class might as well have been a transduplication that would have translocated a putative founder gene to a different chromosomal location, thereby separating the evolution of the ANTP class from that of the remaining classes.

The analyses of the *Eleutheria* genomic fosmids in the present study revealed that the anterior Hox-like gene Cnox-5 is linked to a putative POU gene. A closer look at this linkage might contribute to our understanding of metazoan homeobox gene diversification. Hence, the linked POU gene was further investigated by means of 3' and 5' RACE, phylogenetic sequence analyses and by comparing linkage patterns of putative orthologs in other metazoan genomes.

1.4 Specific aims of this study

In the present study it was attempted to elucidate some crucial aspects about the evolution of metazoan homeobox genes by analysing cnidarian genomes. The Hox-like genes of two cnidarian representatives were analysed to resolve whether Cnidaria already possess an

equivalent of the bilaterian Hox system. Genomic fosmid libraries for *Eleutheria dichotoma* were constructed and screened for clones harbouring the Hox-like genes Cnox-1 to Cnox-5 [12, 13]. The isolated genomic fosmids were sequenced by sub-cloning into plasmid vectors and the resulting contigs analysed for neighbouring genes. Additionally, the unassembled genome of *Nematostella vectensis*, deposited in the trace archives at NCBI, was screened for unknown Hox-like genes. Large genomic contigs were then constructed for the newly identified Hox-like genes and the ones already described in the literature [45]. As with *Eleutheria*, the genomic sequences were analysed for neighbouring genes. Phylogenetic analyses were conducted to infer unambiguous relationships between cnidarian Hox-like genes and bilaterian Hox/ParaHox classes. To assess conservation of function, specified expression patterns of cnidarian Hox-like genes from the literature were compared with those determined for *Eleutheria* (provided by Dr. Wolfgang Jakob) and with the expression patterns characteristic of true Hox genes.

To assess the complexity of the remaining ANTP gene complement in Cnidaria, the unassembled *Nematostella* genome was also screened for non-Hox ANTP class genes and phylogenetic analyses were conducted to identify clear orthologs to bilaterian ANTP class genes. To infer the possible composition of a hypothetical ANTP mega-array of the last common ancestor of Cnidaria and Bilateria, existing linkage of ANTP genes in metazoan genomes was deduced from the literature [24, 51-54] and by screening existing genome assemblies deposited at NCBI.

The linkage of a Hox-like gene to a POU homeobox gene in *Eleutheria* was identified in the course of this work. To assess the significance of this linkage for the evolution of metazoan homeobox genes the complete coding sequence of the POU gene was determined and phylogenetic analyses were conducted to assess orthology to bilaterian POU genes. Furthermore, the linkage patterns of bilaterian orthologs were inferred by analysing existing genome assemblies from NCBI, and compared with the linkage found in *Eleutheria*.

2. Summary of Results and Discussion

Sequence identity, genomic organisation and expression patterns of cnidarian Hox-like genes suggest that Cnidaria predate a mature Hox system

(Kamm et al., 2006, and references therein)

When the Gsx/Diplox-2 type genes are excluded (which are clearly distinct), a total of four Hox-like genes is present in *Eleutheria* and five genes of this type have been previously identified in *Nematostella*. In the present analyses three novel Hox-like genes have been identified in the unassembled *Nematostella* genome. Phylogenetic analyses of the *Eleutheria* and *Nematostella* Hox-like genes with representatives of bilaterian Hox and ParaHox classes suggest common origin for a type of cnidarian genes and anterior Hox genes. The analyses also confirmed the high similarity between some cnidarian genes, including *Eleutheria* Cnox-3 and Cnox-4, and bilaterian posterior gene types. These cnidarian gene types are here referred to as anterior Hox-like and posterior Hox/Cdx-like genes. Although these relationships are well supported, cnidarians clearly lack orthologs of group 3 and central gene types, while the remaining cnidarian genes cannot be assigned to any of the bilaterian Hox/ParaHox classes (i.e. anterior, group 3, central, posterior). Whereas alternative orthology relationships and gene loss have sometimes been suggested for the case of group 3 and central genes, a simpler interpretation is that the common ancestor of Cnidaria and Bilateria had genes which later gave rise to the anterior and posterior Hox genes, but that intermediate (group 3, central) Hox genes - and hence Hox clusters - postdate the split between cnidarians and bilaterians. The remaining cnidarian Hox-like genes rather seem to be genuine cnidarian genes, representing an independent increase in the complexity of regulatory genes in this lineage.

The genomic organisation of the *Eleutheria* and *Nematostella* Hox-like genes also did not show any resemblance of the linkage characteristic for true bilaterian Hox genes. Most Cnidarian Hox-like genes are flanked by unrelated genes in the genomes of *Eleutheria* and *Nematostella*. In the latter, however, two cases of linkage of Hox-like genes were identified. The first case is the linkage of three paralogous genes which also cannot be assigned to true Hox classes of higher animals, and thus most likely represent a genuine cnidarian Hox-like class which underwent a cis-duplication in the *Nematostella* lineage. The second case is the linkage of Anthox1A and Anthox9. It is not clear whether these two are paralogs because Anthox9 is highly derived and might even be a pseudo-gene (although clearly related to Hox genes it does not have the ANTP characteristic Leu residue at homeobox position 16). Nevertheless, neither Anthox1A nor Anthox9 are orthologs of true Hox genes. Hence, both cases of linkage cannot be compared with

the linkage of different Hox classes in bilaterian Hox clusters. In a true Hox cluster no non-Hox genes lie within the cluster; although in some cases (paralogous) Hox-like genes are linked in the Cnidaria, no evidence was found for organization characteristic of true Hox genes. Moreover, as in the coral *Acropora*, an even-skipped gene (a member of the "extended Hox") was found to be tightly linked to an anterior Hox-like gene in *Nematostella,* similar to the tight linkage of even-skipped orthologs to the Hox clusters in vertebrates. This tight linkage is in contrast with the mapping data of the other Hox-like genes and implies that one might reasonable expect to have found other linked genes in the range of the assembled contigs if an equivalent of a Hox cluster was present.

Even bilaterians that exhibit a highly derived cluster show spatial collinearity and similar AP-restricted domains of Hox gene expression. Therefore the expression patterns of the *Eleutheria* Hox-like genes were determined and compared with those of other cnidarians from the literature. The results revealed that the expression of cnidarian Hox-like genes cannot be compared with the axially restricted expression patterns of true Hox genes; even the expression of orthologous genes differs dramatically across Cnidaria and is thus inconsistent with the conserved function of true Hox genes. A striking example is the expression of the *Eleutheria* anterior Hox-like gene Cnox-5 and its orthologs in *Nematostella* and *Podocoryne*. Whereas this ortholog is expressed in the aboral end of the planula in *Eleutheria* and *Podocoryne*, the corresponding *Nematostella* gene is expressed at the opposite end at this stage.

In summary, no equivalent of the bilaterian Hox system was found in Cnidaria in terms of sequence identity, genomic organisation and function. The cnidarian/bilaterian ancestor presumably possessed genes related to bilaterian anterior Hox and posterior Hox/Cdx classes, whilst a mature Hox system - and the unique cnidarian Hox-like genes - most likely postdate the split between Cnidaria and Bilateria. Although the Hox system seems to be responsible for the enormous morphological diversity in Bilateria, it clearly is not mandatory for axial patterning throughout the Metazoa. An alternative - but less parsimonious - scenario is that in those cnidarians examined to date, an ancestral cluster of Hox genes has fragmented, and both individual sequences as well as expression patterns diverged beyond recognition.

Diversity and complexity of the non-Hox ANTP genes in Cnidaria is comparable to the Bilateria

(Kamm and Schierwater, 2006, and references therein)

Hox genes are members of the ANTP class and their origin and evolution is intimately connected with that of the entire class, because we have evidence from Bilateria that most of

the ANTP class genes have emerged and co-evolved on the same ancestral genomic region. To assess to what extent cnidarians share the expansion of the ANTP class genes with Bilateria, the unassembled genome of *Nematostella vectensis* was screened for non-Hox ANTP class genes and phylogenetic analyses were conducted to infer clear orthologies to bilaterian non-Hox ANTP class genes.

The results show that *Nematostella* harbours unambiguous orthologs to almost all non-Hox ANTP gene families that are present in Bilateria, comprising 19 gene families of the "extended Hox" and NKL genes. Furthermore, *Nematostella* possesses genes that are missing in some bilaterian lineages, like the rough gene and NK7 which are absent in vertebrates. The only exceptions are the engrailed and Tlx genes, for which no orthologs were found. Vax genes are possibly also absent in *Nematostella*; two genes were identified that grouped in between the Emx and Vax gene family and it was not possible to decide to which of either these belong. The reasons for the failure to detect these two (or three) gene types could be: gene loss, highly derived sequences, or these genes are an invention of the Bilateria. At this point, however, favouring any of the alternatives seems to be premature.

Many of the identified gene families contain paralogs and thus represent independent duplications of homeobox gene loci. This is consistent with a previous analysis of three duplicated homeobox gene loci in the coral *Acropora* and also with the identified linkage of paralogous Hox-like genes in *Nematostella* (see above). A good example are the Mox gene family, which contains four paralogs, and a subfamily of NK2 genes for which five paralogs could be identified in *Nematostella*. The independent duplication of homeobox genes thus seems to be a general phenomenon in Cnidaria.

Remnants of linkage of ANTP genes in bilaterian genomes suggest that at least the predecessors of most ANTP gene families were ancestrally clustered. The linkage analyses of Hox-like genes in *Nematostella* revealed that the anterior Hox-like gene Anthox6 is linked to an even-skipped ortholog. The same linkage has been previously identified in the coral *Acropora*. Because even-skipped orthologs are linked to vertebrate Hox clusters, this linkage has been interpreted as evidence for a Hox cluster in the common ancestor of Cnidaria and Bilateria. In light of the new data about cnidarian Hox-like genes from this study, however, this linkage most likely reflects ancient linkage of ANTP class members which predates the Hox system. Hence, Cnidaria seem to have split off the lineage leading to Bilateria after most of the ANTP gene families emerged, but before having a mature Hox system.

The linkage of *Eleutheria* POU6 to the anterior Hox-like gene Cnox-5 suggests the ancestral linkage of most metazoan homeobox gene classes

(Kamm and Schierwater, submitted, and references therein)

Linkage of similar genes can be a sign for common origin and co-evolution, especially if such linkage is shared between distantly related taxa on the phylum level. The linkage analyses of the *Eleutheria* Hox-like genes showed that most are flanked by non-Hox genes which contradicts a Hox cluster in Cnidaria. The anterior Hox-like gene Cnox-5, however, was found to be closely linked to a putative POU homeobox gene. In order to assess the significance of this linkage for the evolution of homeobox genes, the coding sequence of the putative *Eleutheria* POU gene was determined and phylogenetic analyses with the complete POU domain were conducted to infer relationships to bilaterian POU genes. Furthermore, bilaterian genome assemblies deposited at NCBI were screened for linkage of orthologous POU genes.

Phylogenetic analyses strongly suggest that the *Eleutheria* POU gene is an ortholog to the bilaterian POU class 6 genes, although it appears to be somewhat derived as compared with representatives from Bilateria. The analyses also suggest the POU class 6 genes as the sister clade to all the remaining POU classes.

Analyses of bilaterian genomes revealed that vertebrates also have retained linkage of POU class 6 genes to their Hox clusters. For example, orthologs of POU class 6 genes map to the same chromosomes as the HoxC and/or HoxA clusters in human, mouse and chicken. In teleost fish these orthologs map to the HoxCa cluster which is one of the duplicated HoxC clusters in the teleost lineage.

Although Cnidaria seem to predate a Hox cluster, the anterior Hox-like genes in Cnidaria, for example the *Eleutheria* Cnox-5 gene, likely share the same ancestor with bilaterian anterior Hox genes. Therefore the linkage of POU class 6 genes to Hox or anterior Hox-like genes can be regarded as ancestral linkage of distantly related homeobox genes.

Metazoan ANTP genes seem to have originated from the same genomic region. Similarly it has been deduced by analyses of the human and mouse genomes that most of the remaining metazoan homeobox classes - notably PRD, LIM, POU, CUT, prospero, TALE and SIX classes - possibly also derived from a common ancestor region. The conserved linkage of POU class 6 genes and Hox or anterior Hox-like genes in two distantly related phyla now suggests that there has been a link between these two ancestor regions. Hence, most likely the ancestors of most metazoan homeobox classes have a common origin.

3. Conclusions and further prospects

3.1 Implications for the evolution of metazoan homeobox genes

In the present work it was possible to clarify some crucial aspects of metazoan homeobox gene evolution by analyses of two cnidarian genomes. The results show that Cnidaria share important traits of homeobox gene evolution with Bilateria, while others postdate the cnidarian/bilaterian split, and even others are unique to Cnidaria. Hence, the homeobox gene complement in Cnidaria concurrently features ancient, complex and derived characters.

The complexity and diversity of non-Hox ANTP homeobox genes predates the split between Cnidaria and Bilateria. After this split the diversity of these genes has been independently enlarged in Cnidaria by several duplications. The functional significance of the unique cnidarian duplicates for the evolution of developmental mechanisms in Cnidaria, however, remains to be analysed in depth in the future by determining their expression patterns and other functional studies.

Two cases of ancient linkage of homeobox genes were found to be retained in Cnidaria. The linkage of a Hox-like gene to an even-skipped ortholog in *Nematostella* is consistent with the linkage patterns of ANTP gene members in bilaterian genomes and supports the view that many ANTP gene members have emerged by several cis-duplications from the same genomic region [19]. Likewise, the linkage of POU class 6 genes to Hox clusters or Hox-like genes, which can be traced from Vertebrates to Cnidaria, together with the linkage patterns of other homeobox gene classes in bilaterian genomes [51-54, 70], provides evidence that at least the predecessors of most metazoan homeobox gene classes have emerged from a common ancestor region. Ongoing and future genome projects of basal metazoans will show if other examples of ancient linkage are still present in extant species.

Cnidaria seem to miss one important step in homeobox gene evolution - which is the evolution of the Hox system. However, the evolutionary success of the Cnidaria - exemplified by about 9,000 extant species featuring a spectacular range of shapes and forms [71] - shows that a Hox system is not mandatory for axial patterning and elaboration of morphological diversity. Nevertheless, the true Hox system most likely has facilitated the even greater extent of morphological diversification in Bilateria - which most likely was jointly responsible for their successful radiation.

3.2 Implications for the origin of the Hox system

Although the evidence favours that Cnidaria predate the Hox system, we are left with several question on how it might have originated. The consensus view about Hox cluster origin has been the ParaHox hypothesis over the last years

[16, 21] - in *bona fide* that Cnidaria conformed to it. Now that we have conflicting evidence from Cnidaria we should carefully reconsider Pros and Cons. There are several weak points in the ParaHox hypothesis, though it fits to the situation in Bilateria. In light of the new evidence from Cnidaria, and also if we extend our view to other diploblasts, we find a situation that is hardly explainable with this theory: Despite enormous efforts no Hox or Hox-like genes have been isolated from sponges and ctenophores; only genes related to non-Hox ANTP families seem to be present [67, 72-75]. In the phylum Placozoa only the Gsx ortholog Diplox-2 (Trox-2) could be identified, in addition to four non-Hox ANTP class genes [23, 57]. The ANTP gene complement of these three phyla is thus in accordance with the hypothesis that the non-Hox ANTP gene families evolved before a true Hox cluster developed; and the presence of a Gsx/Diplox-2 type gene as the only Hox-like gene in Placozoa further contradicts the ParaHox hypothesis.

Cnidarians possess genes related to bilaterian anterior Hox and posterior Hox/Cdx classes and orthologs of Gsx [this work]. Therefore, according to the ParaHox hypothesis, a Hox cluster should have been present in the cnidarian/bilaterian ancestor [16]. However, the present study suggests that Cnidaria lack any equivalent of a Hox system in terms of genomic organisation and function. In terms of sequence identity they clearly lack group 3 Hox genes, its presumable ParaHox counterpart Xlox and the central Hox genes. Hence, the ParaHox theory requires the independent loss of these genes in the Cnidaria. Moreover, many cnidarian Hox-like genes cannot be assigned to any of the Bilaterian Hox classes and the most simple explanation is that these genes represent an independent increase in the complexity of regulatory genes in Cnidaria which has nothing to do with an ancient Hox system.

One could argue that the difficulty in assigning strict Hox classes to many cnidarian Hox-like genes could be a result of the very old divergence from the bilaterian lineage; and there are also examples of derived Hox genes (with no homeotic function) within clusters of Bilateria, like the zen and fushi-tarazu genes in arthropods. But contrary to cnidarians, these genes reside in an otherwise “normal” Hox context and it is now clear that their homologs in more basal arthropods are true Hox genes [76]. Even if we consider loss or divergence, the situation here differs from that in Bilateria because: 1. Contrary to Bilateria we have no compelling evidence that cnidarians ever had the missing genes. 2. It is not reasonable to assume that cnidarians have lost these important genes but otherwise have maintained and independently extended their Hox-like and non-Hox ANTP gene repertoire. 3. Cnidarian Diplox-2, anterior Hox-like and posterior Hox/Cdx-like genes, as well as the almost complete non-Hox ANTP gene families, show high sequence identity to their bilaterian counterparts [20, this work] and it is not convinc-

ing that just the missing genes should have diverged to an extent beyond recognition. In other words, the assumption of loss or divergence is much less parsimonious than the assumption that cnidarians have never possessed these genes. Moreover, whereas orthologies for Gsx and related diploblast genes (Diplox-2) are well supported, there has always been weak support for orthology of cnidarian Hox-like genes to Cdx [20, 22]. Cnidarians rather have genes that seem generally related to posterior gene types, like the posterior Hox/Cdx-like genes Cnox-4 and Cnox-3 of the hydrozoan *Eleutheria dichotoma* [this work].

A modified ParaHox hypothesis

How can we resolve this puzzle? We have convincing evidence that cnidarians possess anterior Hox-like and posterior Hox/Cdx-like genes. In Bilateria, the two outermost Hox genes (i.e. the most anterior and posterior) show highest divergence from the Hox consensus. It is therefore assumed that the bilaterian cluster may have evolved from the outside in, via a series of unequal crossing-over events [77]. This conclusion is consistent with the possibility that cnidarian anterior and posterior gene types may be derived from these ancestral two outer genes, in which case it is most appropriate to view them as independently derived from the predecessors of anterior and posterior Hox genes, rather than as having any direct relationship with Hox classes of higher animals.

Furthermore, the Diplox-2 genes in Cnidaria and Placozoa seem to be true orthologs of the bilaterian Gsx genes [20, 22, 57, this work]. The fact that this is the only Hox-like gene in the basal phylum Placozoa also deserves further consideration: We may speculate if a Gsx/Diplox-2 type gene was the founder of all Hox/ParaHox classes, hence the putative ProtoHox. If we compare the function of bilaterian and cnidarian Hox/ParaHox(-like) genes, the Gsx/Diplox-2 genes seem to be the only one with a conserved function. Across Bilateria Gsx orthologs are involved in neurogenesis [22, 78 and references therein]. With data for the coral *Acropora* [78] and the hydrozoan *Hydra* [Miljkovic-Licina & Galliot et al., personal communication] we have now good reason to assume that this is also the case in Cnidaria (there could be, however, additional functions for Gsx/Diplox-2 in cnidarians, [c.f. 20, 22, 46, 79, 80]).

Moreover, if Gsx was a direct descendant of an anterior ProtoHox, as postulated in the ParaHox hypothesis, then we would expect the Gsx genes to be the sister clade to all anterior Hox genes. Instead, in most analyses there is either weak support for a sister relationship of Gsx and anterior Hox genes or they group as a sister clade to most or all Hox/ParaHox classes [20-22, 81, 82, this work]. Gsx/Diplox-2 type genes are clearly distinct from other Hox/ParaHox genes. It could be that this divergence is due to an ancient duplication of an archetypical Gsx/Diplox-2 gene founding the predecessors of anterior

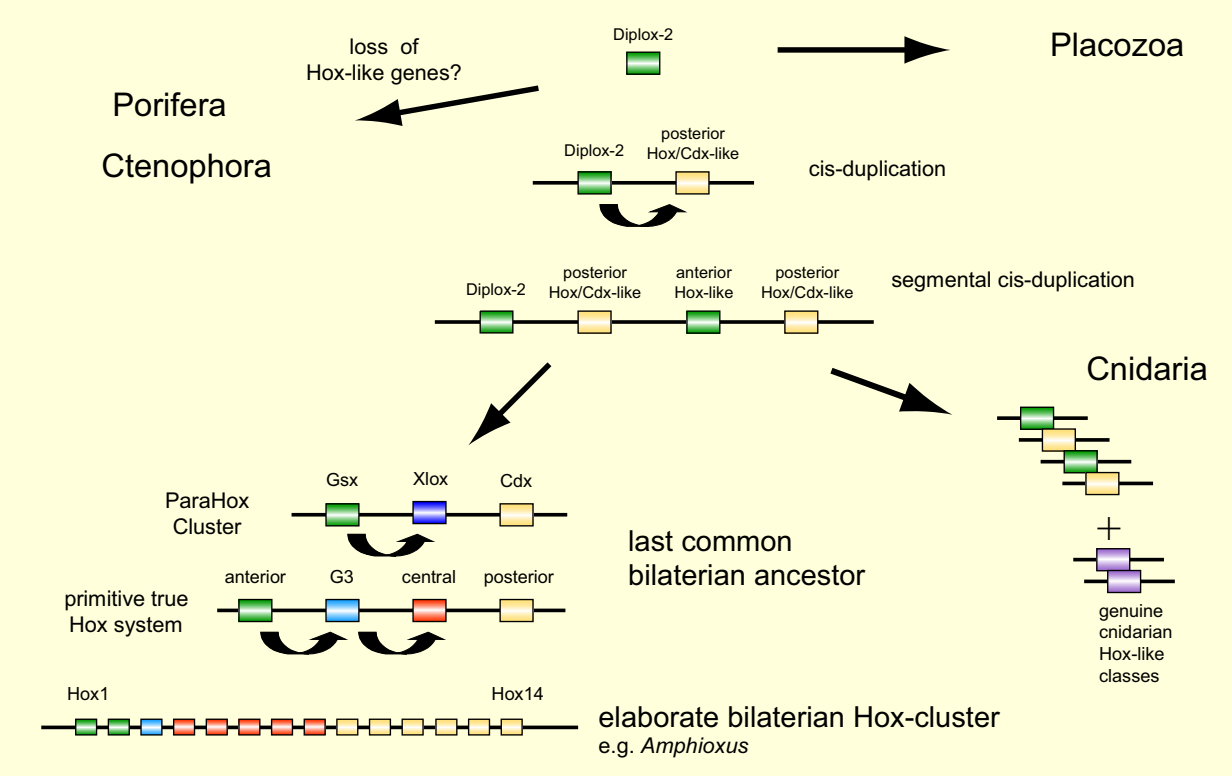


Figure 7: Possible scenario for the evolution of the Hox system in Bilateria and the Hox-like gene complement in Cnidaria. Starting from a Diplox-2/Gsx gene a duplication could have yielded Diplox-2/Gsx and a predecessor of posterior Hox/Cdx genes. The whole segment then underwent a segmental cis-duplication. In Bilateria one segment founded the ParaHox cluster while the other founded the Hox cluster. In Cnidaria most of the genes were dispersed and independent duplications increased the complexity of their Hox-like gene complement. In this scenario Hox 3 and Xlox would have arisen independently from related genes which may account for their similarity.

and posterior Hox/ParaHox classes. The copies of this event could have evolved more rapidly because they had no need for proper function. Such relaxed constraints on the evolution of duplicated genes also seem to have played a major role during duplication of the vertebrate Hox clusters and the subsequent recruitment of the new genes for other functions [83].

A possible scenario (Figure 7) therefore is that a Gsx/Diplox-2 type gene was the first metazoan Hox-like gene which probably evolved in the context of other genes of the ANTP class [19, 53] and other less related homeobox genes (see above). The ancestral function of Diplox-2 could have been neurogenesis, though we can only speculate what this means for Placozoa since they lack true nerve cells. However, the Diplox-2 gene Trox-2 is expressed in the marginal cells of *Trichoplax adhaeren*s [23] and these cells also appear to express the neurotransmitter RFamide [84].

The Diplox-2 could have duplicated to give rise to a posterior Hox/Cdx type gene. The next evolutionary event could have been a cis-duplication of the whole segment which gave rise on one hand to Diplox-2 plus a posterior Hox/Cdx type gene and, on the other hand, to the predecessors of anterior and posterior Hox genes. After the segmental cis-duplication the Diplox-2 gene and the posterior Hox/Cdx gene were pos-

sibly transposed away while the predecessors of anterior and posterior Hox still resided in the context of other ANTP genes, accounting for the still present linkage of even-skipped orthologs to anterior Hox-like genes in Cnidaria [85, this work] and to the Hox clusters in vertebrates [53]. In the bilaterian lineage the predecessors of Hox and ParaHox genes might have retained linkage for some time. In the urochordate *Oikopleura* we still see linkage of Cdx to an anterior Hox gene [26]. This, however, is the only known example and might thus be mere chance.

Nevertheless, at some point the duplicated genes must have undergone different fates in the respective lineages. In Cnidaria the genes were dispersed and only the linkage of anterior Hox-like genes to even-skipped was maintained in some species, or to a POU class 6 gene in others. Furthermore, independent duplications of Hox-like genes founded genuine cnidarian Hox-like classes. In Bilateria the segment containing the Diplox-2 and the posterior Hox/Cdx became the ParaHox cluster, while the predecessors of anterior and posterior Hox genes became the founder of the Hox cluster. A similar two gene origin of bilaterian Hox and ParaHox clusters (Figure 8C), starting from one anterior and one posterior type gene, has been proposed recently to accommodate conflicting views about the nature of cnidarian Hox-like genes [6, 19].

This scenario, however, requires that the group 3 Hox genes and the ParaHox Xlox have arisen independently in the bilaterian Hox and ParaHox clusters and that their resemblance

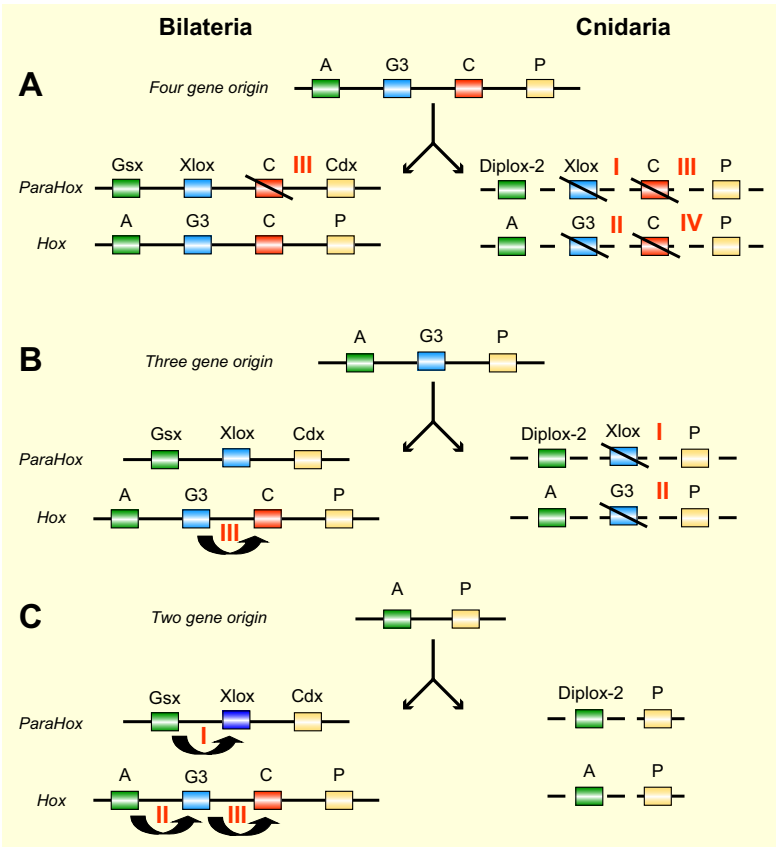


Figure 8: Proposed alternative steps that are necessary to accommodate current knowledge about the Hox/ParaHox(-like) gene complements in Metazoa (modified from [19]): **A** The original ParaHox theory, starting from a four gene model in the cnidarian/bilaterian ancestor, requires four steps to explain the Hox(-like) and ParaHox(-like) complements in extant Metazoa. **B** A three gene model, lacking the central genes, requires only three steps. **C** Likewise, a two gene model also requires only three steps. However, as outlined in the text, the assumption of several gene losses is less parsimonious than the assumption that cnidarians have never had the missing genes. Therefore, a two gene model, starting from one anterior and one posterior type gene seems much more likely. The founder of these two genes may have been a Diplox-2/Gsx type gene. The first appearance and further elaboration of the true Hox system in the bilaterian lineage would also fit to major transitions such as the origin of Bilateria and the Cambrian Explosion [19].

is simply convergence. In this case both could have originated from the independent duplication of paralogous genes which may account for their similarity. Altogether, this scenario seems much more likely than the assumption of several gene losses in Cnidaria for which we have no evidence at all. Figure 8 compares the alternative hypotheses for Hox cluster origin and the underlying assumptions that are necessary to accommodate the situation found in extant Metazoa [19].

The conserved function of Diplox-2 and their bilaterian orthologs Gsx seems to be neurogenesis and across Bilateria Hox genes are involved in patterning of the nervous system [86-96]. This might have been their ancestral function because Hox expression seems predominantly ectodermal in Bilateria [19], and patterning of the whole ectoderm in chordates is possibly generally neurally related [97]. We may therefore hypothesize that the duplicated genes enabled regionalisation of the nervous system in Bilateria and subsequently were co-opted for axial patterning.

Although the above scenario accommodates for the situation found in extant Metazoa, the precise course of Hox and Hox-like gene evolution is still speculative. It is unclear if we will ever find some of the missing links of this scenario. Most promising is the sequencing of other cnidarian genomes, or those of even more basal Metazoa like Placozoa or Porifera; and although it seems clear, that the role of cnidarian Hox-like genes is not comparable to that of true bilaterian Hox genes, we still need to assess precisely what their function is in the determination of the cnidarian bauplan.

4. References

1. Lewis, E.B. (1978). A gene complex controlling segmentation in *Drosophila*. Nature 276, 565-570.
2. McGinnis, W., and Krumlauf, R. (1992). Homeobox genes and axial patterning. Cell 68, 283-302.
3. Gehring, W.J. (1985). The homeo box: a key to the understanding of development? Cell 40, 3-5.
4. Ferrier, D.E., and Minguillon, C. (2003). Evolution of the Hox/ParaHox gene clusters. Int J Dev Biol 47, 605-611.
5. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). Molecular Biology of the Cell, 4th Edition (Garland Science).
6. Garcia-Fernandez, J. (2005). Hox, ParaHox, ProtoHox: facts and guesses. Heredity 94, 145-152.
7. Garcia-Fernandez, J., and Holland, P.W. (1994). Archetypal organization of the amphioxus Hox gene cluster. Nature 370, 563-566.
8. Amores, A., Suzuki, T., Yan, Y.L., Pomeroy, J., Singer, A., Amemiya, C., and Postlethwait, J.H. (2004). Developmental roles of pufferfish Hox clusters and genome evolution in ray-fin fish. Genome Res 14, 1-10.
9. Prince, V. (2002). The Hox Paradox: More complex(es) than imagined. Dev Biol 249, 1-15.
10. Moghadam, H.K., Ferguson, M.M., and Danzmann, R.G. (2005). Evidence for Hox gene duplication in rainbow trout (*Oncorhynchus mykiss*): a tetraploid model species. J Mol Evol 61, 804-818.
11. Gellon, G., and McGinnis, W. (1998). Shaping animal body plans in development and evolution by modulation of Hox expression patterns. Bioessays 20, 116-125.
12. Schierwater, B., Murtha, M., Dick, M., Ruddle, F.H., and Buss, L.W. (1991). Homeoboxes in cnidarians. J Exp Zool 260, 413-416.
13. Kuhn, K., Streit, B., and Schierwater, B. (1996). Homeobox genes in the cnidarian *Eleutheria dichotoma*: evolutionary implications for the origin of Antennapedia-class (HOM/Hox) genes. Mol Phylogenet Evol 6, 30-38.
14. Kuhn, K., Streit, B., and Schierwater, B. (1999). Isolation of Hox genes from the scyphozoan *Cassiopeia xamachana*: implications for the early evolution of Hox genes. J Exp Zool 285, 63-75.
15. Slack, J.M., Holland, P.W., and Graham, C.F. (1993). The zootype and the phylotypic stage. Nature 361, 490-492.
16. Ferrier, D.E., and Holland, P.W. (2001). Ancient origin of the Hox gene cluster. Nat Rev Genet 2, 33-38.
17. Schierwater, B., Dellaporta, S., and DeSalle, R. (2002). Is the evolution of Cnox-2 Hox/ParaHox genes “multicolored” and “polygenealogical?” Mol Phylogenet Evol 24, 374-378.
18. Schierwater, B., and Desalle, R. (2001). Current problems with the zootype and the early evolution of Hox genes. J Exp Zool 291, 169-174.
19. Garcia-Fernandez, J. (2005). The genesis and evolution of homeobox gene clusters. Nat Rev Genet 6, 881-892.
20. Gauchat, D., Mazet, F., Berney, C., Schummer, M., Kreger, S., Pawlowski, J., and Galliot, B. (2000). Evolution of Antp-class genes and differential expression of *Hydra* Hox/paraHox genes in anterior patterning. Proc Natl Acad Sci U S A 97, 4493-4498.
21. Brooke, N.M., Garcia-Fernandez, J., and Holland, P.W. (1998). The ParaHox gene cluster is

an evolutionary sister of the Hox gene cluster. Nature 392, 920-922.

22. Finnerty, J.R., Paulson, D., Burton, P., Pang, K., and Martindale, M.Q. (2003). Early evolution of a homeobox gene: the parahox gene Gsx in the Cnidaria and the Bilateria. Evol Dev 5, 331-345.
23. Jakob, W., Sagasser, S., Dellaporta, S., Holland, P., Kuhn, K., and Schierwater, B. (2004). The Trox-2 Hox/ParaHox gene of *Trichoplax* (Placozoa) marks an epithelial boundary. Dev Genes Evol 214, 170-175.
24. Minguillon, C., Gardenyes, J., Serra, E., Castro, L.F., Hill-Force, A., Holland, P.W., Amemiya, C.T., and Garcia-Fernandez, J. (2005). No more than 14: the end of the amphioxus Hox cluster. Int J Biol Sci 1, 19-23.
25. Spagnuolo, A., Ristoratore, F., Di Gregorio, A., Aniello, F., Branno, M., and Di Lauro, R. (2003). Unusual number and genomic organization of Hox genes in the tunicate *Ciona intestinalis*. Gene 309, 71-79.
26. Seo, H.C., Edvardsen, R.B., Maeland, A.D., Bjordal, M., Jensen, M.F., Hansen, A., Flaat, M., Weissenbach, J., Lehrach, H., Wincker, P., Reinhardt, R., and Chourrout, D. (2004). Hox cluster disintegration with persistent anteroposterior order of expression in *Oikopleura* dioica. Nature 431, 67-71.
27. Ikuta, T., Yoshida, N., Satoh, N., and Saiga, H. (2004). *Ciona intestinalis* Hox gene cluster: Its dispersed structure and residual colinear expression in development. Proc Natl Acad Sci U S A 101, 15118-15123.
28. Martinez, P., Rast, J.P., Arenas-Mena, C., and Davidson, E.H. (1999). Organization of an echinoderm Hox gene cluster. Proc Natl Acad Sci U S A 96, 1469-1474.
29. Arenas-Mena, C., Cameron, A.R., and Davidson, E.H. (2000). Spatial expression of Hox cluster genes in the ontogeny of a sea urchin. Development 127, 4631-4643.
30. Cameron, R.A., Rowen, L., Nesbitt, R., Bloom, S., Rast, J.P., Berney, K., Arenas-Mena, C., Martinez, P., Lucas, S., Richardson, P.M., Davidson, E.H., Peterson, K.J., and Hood, L. (2006). Unusual gene order and organization of the sea urchin hox cluster. J Exp Zoolog B Mol Dev Evol 306, 45-58.
31. Von Allmen, G., Hogga, I., Spierer, A., Karch, F., Bender, W., Gyurkovics, H., and Lewis, E. (1996). Splits in fruitfly Hox gene complexes. Nature 380, 116.
32. Devenport, M.P., Blass, C., and Eggleston, P. (2000). Characterization of the Hox gene cluster in the malaria vector mosquito, *Anopheles gambiae*. Evol Dev 2, 326-339.
33. Ferrier, D.E., and Akam, M. (1996). Organization of the Hox gene cluster in the grasshopper, *Schistocerca gregaria*. Proc Natl Acad Sci U S A 93, 13024-13029.
34. Aboobaker, A., and Blaxter, M. (2003). Hox gene evolution in nematodes: novelty conserved. Curr Opin Genet Dev 13, 593-598.
35. Aboobaker, A.A., and Blaxter, M.L. (2003). Hox Gene Loss during Dynamic Evolution of the Nematode Cluster. Curr Biol 13, 37-40.
36. Van Auken, K., Weaver, D.C., Edgar, L.G., and Wood, W.B. (2000). *Caenorhabditis elegans* embryonic axial patterning requires two recently discovered posterior-group Hox genes. Proc Natl Acad Sci U S A 97, 4499-4503.
37. Lee, P.N., Callaerts, P., De Couet, H.G., and Martindale, M.Q. (2003). Cephalopod Hox genes and the origin of morphological novelties. Nature 424, 1061-1065.
38. Hinman, V.F., O'Brien, E.K., Richards, G.S., and

Degnan, B.M. (2003). Expression of anterior Hox genes during larval development of the gastropod *Haliotis asinina*. Evol Dev 5, 508-521.

39. Andreeva, T.F., Kuk, C., Korchagina, N.M., C'Ikc'm, M., and Dondya, A.K. (2001). [Cloning and analysis of structural organization of Hox genes in the Polychaete *Nereis virens*]. Ontogenez 32, 225-233.
40. Kmita-Cunisse, M., Loosli, F., Bierne, J., and Gehring, W.J. (1998). Homeobox genes in the ribbonworm *Lineus sanguineus*: evolutionary implications. Proc Natl Acad Sci U S A 95, 3030-3035.
41. Baguna, J., and Riutort, M. (2004). The dawn of bilaterian animals: the case of acoelomorph flatworms. Bioessays 26, 1046-1057.
42. Cook, C.E., Jimenez, E., Akam, M., and Salo, E. (2004). The Hox gene complement of acoel flatworms, a basal bilaterian clade. Evol Dev 6, 154-163.
43. Aguinaldo, A.M., Turbeville, J.M., Linford, L.S., Rivera, M.C., Garey, J.R., Raff, R.A., and Lake, J.A. (1997). Evidence for a clade of nematodes, arthropods and other moulting animals. Nature 387, 489-493.
44. Dellaporta, S.L., Xu, A., Sagasser, S., Jakob, W., Moreno, M.A., Buss, L.W., and Schierwater, B. (2006). Mitochondrial genome of *Trichoplax adhaerens* supports placozoa as the basal lower metazoan phylum. Proc Natl Acad Sci U S A 103, 8751-8756.
45. Finnerty, J.R., Pang, K., Burton, P., Paulson, D., and Martindale, M.Q. (2004). Origins of bilateral symmetry: Hox and dpp expression in a sea anemone. Science 304, 1335-1337.
46. Yanze, N., Spring, J., Schmidli, C., and Schmid, V. (2001). Conservation of Hox/ParaHox-related genes in the early development of a cnidarian. Dev Biol 236, 89-98.
47. Galliot, B., de Vargas, C., and Miller, D. (1999). Evolution of homeobox genes: Q50 Paired-like genes founded the Paired class. Dev Genes Evol 209, 186-197.
48. Banerjee-Basu, S., and Baxevanis, A.D. (2001). Molecular evolution of the homeodomain family of transcription factors. Nucleic Acids Res 29, 3258-3269.
49. Holland, P.W. (2001). Beyond the Hox: how widespread is homeobox gene clustering? J Anat 199, 13-23.
50. Holland, P.W., and Takahashi, T. (2005). The evolution of homeobox genes: Implications for the study of brain development. Brain Res Bull 66, 484-490.
51. Luke, G.N., Castro, L.F., McLay, K., Bird, C., Coulson, A., and Holland, P.W. (2003). Dispersal of NK homeobox gene clusters in amphioxus and humans. Proc Natl Acad Sci U S A 100, 5292-5295.
52. Coulier, F., Popovici, C., Villet, R., and Birnbaum, D. (2000). MetaHox gene clusters. J Exp Zool 288, 345-351.
53. Pollard, S.L., and Holland, P.W. (2000). Evidence for 14 homeobox gene clusters in human genome ancestry. Curr Biol 10, 1059-1062.
54. Castro, L.F., and Holland, P.W. (2003). Chromosomal mapping of ANTP class homeobox genes in amphioxus: piecing together ancestral genomes. Evol Dev 5, 459-465.
55. Ender, A., and Schierwater, B. (2003). Placozoa are not derived cnidarians: evidence from molecular morphology. Mol Biol Evol 20, 130-134.
56. Miller, D.J., and Ball, E.E. (2005). Animal evolution: the enigmatic phylum placozoa revisited. Curr Biol 15, R26-28.
57. Monteiro, A.S., Schierwater, B., Dellaporta, S.L.,

and Holland, P.W. (2006). A low diversity of ANTP class homeobox genes in Placozoa. Evol Dev 8, 174-182.

58. Kortschak, R.D., Samuel, G., Saint, R., and Miller, D.J. (2003). EST analysis of the cnidarian *Acropora millepora* reveals extensive gene loss and rapid sequence divergence in the model invertebrates. Curr Biol 13, 2190-2195.
59. Miller, D.J., Ball, E.E., and Technau, U. (2005). Cnidarians and ancestral genetic complexity in the animal kingdom. Trends Genet 21, 536-539.
60. Technau, U., Rudd, S., Maxwell, P., Gordon, P.M., Saina, M., Grasso, L.C., Hayward, D.C., Sensen, C.W., Saint, R., Holstein, T.W., Ball, E.E., and Miller, D.J. (2005). Maintenance of ancestral complexity and non-metazoan genes in two basal cnidarians. Trends Genet 21, 633-639.
61. Burglin, T.R. (1997). Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals. Nucleic Acids Res 25, 4173-4180.
62. Ryan, A.K., and Rosenfeld, M.G. (1997). POU domain family values: flexibility, partnerships, and developmental codes. Genes Dev 11, 1207-1225.
63. Phillips, K., and Luisi, B. (2000). The virtuoso of versatility: POU proteins that flex to fit. J Mol Biol 302, 1023-1039.
64. Latchman, D.S. (1999). POU family transcription factors in the nervous system. J Cell Physiol 179, 126-133.
65. Zhou, H., Yoshioka, T., and Nathans, J. (1996). Retina-derived POU-domain factor-1: a complex POU-domain gene implicated in the development of retinal ganglion and amacrine cells. J Neurosci 16, 2261-2274.
66. Seimiya, M., Watanabe, Y., and Kurosawa, Y. (1997). Identification of POU-class homeobox genes in a freshwater sponge and the specific expression of these genes during differentiation. Eur J Biochem 243, 27-31.
67. Larroux, C., Fahey, B., Liubicich, D., Hinman, V.F., Gauthier, M., Gongora, M., Green, K., Worheide, G., Leys, S.P., and Degnan, B.M. (2006). Developmental expression of transcription factor genes in a demosponge: insights into the origin of metazoan multicellularity. Evol Dev 8, 150-173.
68. Shah, D., Aurora, D., Lance, R., and Stuart, G.W. (2000). POU genes in metazoans: homologs in sea anemones, snails, and earthworms. DNA Seq 11, 457-461.
69. Ryan, J.F., Burton, P.M., Mazza, M.E., Kwong, G.K., Mullikin, J.C., and Finnerty, J.R. (2006). The cnidarian-bilaterian ancestor possessed at least 56 homeoboxes. Evidence from the starlet sea anemone, *Nematostella vectensis*. Genome Biol 7, R64.
70. Popovici, C., Leveugle, M., Birnbaum, D., and Coulier, F. (2001). Homeobox gene clusters and the human paralogy map. FEBS Lett 491, 237-242.
71. Brusca, R.C., and Brusca, G.J. (1990). Invertebrates (Sinauer Associates, Sunderland).
72. Coutinho, C.C., Fonseca, R.N., Mansure, J.J., and Borojevic, R. (2003). Early steps in the evolution of multicellularity: deep structural and functional homologies among homeobox genes in sponges and higher metazoans. Mech Dev 120, 429-440.
73. Hill, A., Tetrault, J., and Hill, M. (2004). Isolation and expression analysis of a poriferan Antp-class Bar-/Bsh-like homeobox gene. Dev Genes Evol 214, 515-523.

74. Martinelli, C., and Spring, J. (2005). T-box and homeobox genes from the ctenophore *Pleurobrachia pileus*: comparison of Brachyury, Tbx2/3 and Tlx in basal metazoans and bilaterians. FEBS Lett 579, 5024-5028.
75. Lee, S.E., Gates, R.D., and Jacobs, D.K. (2003). Gene fishing: the use of a simple protocol to isolate multiple homeodomain classes from diverse invertebrate taxa. J Mol Evol 56, 509-516.
76. Hughes, C.L., and Kaufman, T.C. (2002). Hox genes and the evolution of the arthropod body plan. Evol Dev 4, 459-499.
77. Gehring, W.J., Affolter, M., and Burglin, T. (1994). Homeodomain proteins. Annu Rev Biochem 63, 487-526.
78. Hayward, D.C., Catmull, J., Reece-Hoyes, J.S., Berghammer, H., Dodd, H., Hann, S.J., Miller, D.J., and Ball, E.E. (2001). Gene structure and larval expression of cnox-2Am from the coral *Acropora millepora*. Dev Genes Evol 211, 10-19.
79. Shenk, M.A., Bode, H.R., and Steele, R.E. (1993). Expression of Cnox-2, a HOM/HOX homeobox gene in hydra, is correlated with axial pattern formation. Development 117, 657-667.
80. Cartwright, P., Bowsher, J., and Buss, L.W. (1999). Expression of a Hox gene, Cnox-2, and the division of labor in a colonial hydroid. Proc Natl Acad Sci U S A 96, 2183-2186.
81. Minguillon, C., and Garcia-Fernandez, J. (2003). Genesis and evolution of the Evx and Mox genes and the extended Hox and ParaHox gene clusters. Genome Biol 4, R12.
82. Finnerty, J.R., and Martindale, M.Q. (1999). Ancient origins of axial patterning genes: Hox genes and ParaHox genes in the Cnidaria. Evol Dev 1, 16-23.
83. Wagner, G.P., Amemiya, C., and Ruddle, F. (2003). Hox cluster duplications and the opportunity for evolutionary novelties. Proc Natl Acad Sci U S A 100, 14603-14606.
84. Schuchert, P. (1993). *Trichoplax adhaerens* (Phylum Placozoa) has cells that react with antibodies against the neuropeptide RFamide. Acta Zoologica 74, 115-117.
85. Miller, D.J., and Miles, A. (1993). Homeobox genes and the zootype. Nature 365, 215-216.
86. Hirth, F., and Reichert, H. (1999). Conserved genetic programs in insect and mammalian brain development. Bioessays 21, 677-684.
87. Hirth, F., Hartmann, B., and Reichert, H. (1998). Homeotic gene action in embryonic brain development of *Drosophila*. Development 125, 1579-1589.
88. Hirth, F., Loop, T., Egger, B., Miller, D.F., Kaufman, T.C., and Reichert, H. (2001). Functional equivalence of Hox gene products in the specification of the tritocerebrum during embryonic brain development of Drosophila. Development 128, 4781-4788.
89. Sprecher, S.G., Muller, M., Kammermeier, L., Miller, D.F., Kaufman, T.C., Reichert, H., and Hirth, F. (2004). Hox gene cross-regulatory interactions in the embryonic brain of *Drosophila*. Mech Dev 121, 527-536.
90. Wada, H., Garcia-Fernandez, J., and Holland, P.W. (1999). Colinear and segmental expression of amphioxus Hox genes. Dev Biol 213, 131-141.
91. Dasen, J.S., Liu, J.P., and Jessell, T.M. (2003). Motor neuron columnar fate imposed by sequential phases of Hox-c activity. Nature 425, 926-933.
92. Dasen, J.S., Tice, B.C., Brenner-Morton, S., and Jessell, T.M. (2005). A Hox regulatory network establishes motor neuron pool identity and tar-

get-muscle connectivity. Cell 123, 477-491.

93. Wilkinson, D.G. (1993). Molecular mechanisms of segmental patterning in the vertebrate hindbrain and neural crest. Bioessays 15, 499-505.
94. Wilkinson, D.G. (1993). Molecular mechanisms of segmental patterning in the vertebrate hindbrain. Perspect Dev Neurobiol 1, 117-125.
95. Kourakis, M.J., Master, V.A., Lokhorst, D.K., Nardelli-Haefliger, D., Wedeen, C.J., Martindale, M.Q., and Shankland, M. (1997). Conserved anterior boundaries of Hox gene expression in the central nervous system of the leech Helobdella. Dev Biol 190, 284-300.
96. Papillon, D., Perez, Y., Fasano, L., Le Parco, Y., and Caubit, X. (2005). Restricted expression of a median Hox gene in the central nervous system of chaetognaths. Dev Genes Evol 215, 369-373.
97. Holland, L.Z. (2005). Non-neural ectoderm is really neural: evolution of developmental patterning mechanisms in the non-neural ectoderm of chordates and the problem of sensory cell homologies. J Exp Zoolog B Mol Dev Evol 304, 304-323.

Acknowledgements

This section is to pay tribute to those people who have contributed to this work by sharing resources, knowledge, ideas, sympathy, passion or simply money. Without them this work would not have been possible.

First I would like to thank my supervisor Professor Dr. Bernd Schierwater for providing me with this outstanding topic, with facilities and with excellent assistance throughout the course of my thesis. Several collaborations of our institute have been indispensable for the conduction of my work: I like to thank Professor Dr. Steven Dellaporta, Yale University, for the opportunity to visit his lab and learn how to make genomic libraries. Likewise, I thank Professor Dr. Peter Holland, Oxford University, for his help and hospitality (and also his team) during my stay in Oxford and for several excellent suggestions and discussions. Professor Dr. David J. Miller, James Cook University, was always an important source of expertise on the phylogenetics of homeobox genes - thanks a lot. I also like to thank Professor Dr. Tosso Leeb (formerly Institut für Tierzucht, TiHo-Hannover) who was always open for help and suggestions regarding the sub-cloning of the *Eleutheria* fosmids.

I am also deeply grateful for three years of financial support granted by the Evangelische Studienstiftung e.V. Villigst and for a three months scholarship from the DAAD that enabled my visit to Yale University.

The help, the hospitality and the spirit of the entire ITZ team at the TiHo-Hannover, its former as well as its present members, was essential for me and my work. I am grateful especially to Werner Schroth, Sven Sagasser and Jessica Rach for all kinds of support. Many thanks also to Wolfgang Jakob who provided *Eleutheria* expression data for the Current Biology paper.

Last, but not least, I want to thank my parents who have always supported me.

Curriculum Vitae

Kai Kamm
Deisterstraße 58
30449 Hannover
Germany

Date, Place of Birth	02.01.1971, Mayen, Germany
Family Status	unmarried
Education	
1983 - 1990	Graduation from Gymnasium Großburgwedel with Allgemeine Hochschulreife
University Education	
Oct. 1992 - Feb. 1999	Study of Biology at the University of Hannover. Graduation with University Diploma
Further Education	
Sep. 2000 - Sep. 2001	Training as trade-journal/online-editor at the WBS Training AG in Düsseldorf
Experience	
Oct. 1990 - Dec. 1991	Civil Service in a „Factory for Handicapped“, Pestalozzi-Foundation Großburgwedel
1992 - 1999	Diversity of jobs before and during university education. (e.g. working with handicapped, at the Hannover-Fair, in university institutes etc.)
Sep. 1999 - Jun. 2000	Scientific Assistant in the Institute of Physiology, TiHo-Hannover
May - Jul. 2001	Practical training in the editorial office of PC Praxis
Oct. 2001 - May 2002	Freelancing Journalist (e.g. contributions to Journalismus.com, Agency Crossrelations)
since Jul. 2002	Scientific Assistant (PhD student) at the Institute of Ecology and Evolution, TiHo-Hannover
Studies Abroad and Scholarships	
Sep. 1995 - Apr. 1996	Enrolment in the Biology Masters Program at Northeastern University, Boston, USA. Courses at Northeastern University and at the Discovery Bay Marine Lab on Jamaica. Supported by a DAAD scholarship (IAS).
Oct. - Dec. 2002	Work in the lab of Professor Dr. Steven Dellaporta, Yale University, USA. Supported by a DAAD scholarship.
Nov. 2003 - Oct. 2006	Scholarship of the Evangelische Studienstiftung e.V. Villigst.
May 2005	Work in the lab of Professor Dr. Peter Holland, Oxford University, UK.

List of Scientific Publications

Kamm, K., Hoppe, S., Breves, G., Schroder, B., and Schemann, M. (2004). Effects of the probiotic yeast *Saccharomyces boulardii* on the neurochemistry of myenteric neurones in pig jejunum. Neurogastroenterol Motil 16, 53-60.

Kamm, K., Schierwater, B., Jakob, W., Dellaporta, S.L., and Miller, D.J. (2006). Axial patterning and diversification in the cnidaria predate the Hox system. Curr Biol 16, 920-926.

Kamm, K., and Schierwater, B. (2006). Ancient complexity of the non-Hox ANTP gene complement in the anthozoan *Nematostella vectensis*. Implications for the evolution of the ANTP superclass. J Exp Zoolog B Mol Dev Evol 306, 589-596.

Kamm, K., and Schierwater, B. (2007). Ancient Linkage of a POU Class 6 and an Anterior Hox-Like Gene in Cnidaria: Implications for the Evolution of Homeobox Genes. J Exp Zoolog B Mol Dev Evol.

Sagasser, S., Dellaporta, S., Kamm, K., and Schierwater, B. (in preparation). The placozoan Hox-like gene Trox-2 as a candidate for ProtoHox function. Cell.

Dellaporta, S., Rokhsar, D., DeSalle, R., Holland, P., Buss, L., Kamm, K., and Schierwater, B. (in preparation). The draft genome of *Trichoplax adhaerens*. Nature.

Schierwater, B., Kamm, K., and Miller, D.J. (in preparation). The evolution of the Hox system. Review. Bioessays.

Addendum

The publications upon which this cumulative dissertation is based:

Kamm, K., Schierwater, B., Jakob, W., Dellaporta, S.L., and Miller, D.J. (2006). Axial patterning and diversification in the cnidaria predate the Hox system. Curr Biol 16, 920-926.

Kamm, K., and Schierwater, B. (2006). Ancient complexity of the non-Hox ANTP gene complement in the anthozoan *Nematostella vectensis*. Implications for the evolution of the ANTP superclass. J Exp Zoolog B Mol Dev Evol 306, 589-596.

Kamm, K., and Schierwater, B. (submitted). Ancient Linkage of a POU Class 6 and an Anterior Hox-Like Gene in Cnidaria: Implications for the Evolution of Homeobox Genes. J Exp Zoolog B Mol Dev Evol.

Axial Patterning and Diversification in the Cnidaria Predate the Hox System

Kai Kamm[1], Bernd Schierwater[1,2]*, Wolfgang Jakob[1], Stephen L. Dellaporta[2], and David J. Miller[3,4]*

[1]ITZ, Ecology and Evolution, Tierärztliche Hochschule Hannover, D-30559 Hannover, Germany
[2]Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, Connecticut 06520, USA
[3]Comparative Genomics Centre &
[4]ARC Centre of Excellence for Coral Reef Studies, James Cook University, Townsville QLD 4811, Australia

*Correspondence: bernd.schierwater@ecolevol.de (B.S.); david.miller@jcu.edu.au (D.J.M.)

This is the author's version of a work published by Elsevier. The article has been originally published in:

Current Biology 16, 920–926, May 9, 2006, DOI 10.1016/j.cub.2006.03.036

The original article can be found online at: www.current-biology.com

Axial Patterning and Diversification in the Cnidaria Predate the Hox System

Kai Kamm[1], Bernd Schierwater[1,2]*, Wolfgang Jakob[1], Stephen L. Dellaporta[2], and David J. Miller[3,4]*

[1]ITZ, Ecology and Evolution, Tierärztliche Hochschule Hannover, D-30559 Hannover, Germany
[2]Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, Connecticut 06520, USA
[3]Comparative Genomics Centre &
[4]ARC Centre of Excellence for Coral Reef Studies, James Cook University, Townsville QLD 4811, Australia

Summary

Across the animal kingdom, Hox genes are organized in clusters whose genomic organization reflects their central roles in patterning along the anterior/posterior (A/P) axis [1–7]. While a cluster of Hox genes was present in the bilaterian common ancestor, the origins of this system remain unclear (cf. [8]). With new data for two representatives of the closest extant phylum to the Bilateria, the sea anemone *Nematostella* and the hydromedusa *Eleutheria*, we argue here that the Cnidaria predate the evolution of the Hox system. Although Hox-like genes are present in a range of cnidarians, many of these are paralogs and in neither *Nematostella* nor *Eleutheria* is an equivalent of the Hox cluster present. With the exception of independently duplicated genes, the cnidarian genes are unlinked and in several cases are flanked by non-Hox genes. Furthermore, the cnidarian genes are expressed in patterns that are inconsistent with the Hox paradigm. We conclude that the Cnidaria/Bilateria split occurred before a definitive Hox system developed. The spectacular variety in morphological and developmental characteristics shown by extant cnidarians demonstrates that there is no obligate link between the Hox system and morphological diversity in the animal kingdom and that a canonical Hox system is not mandatory for axial patterning.

Results and Discussion

The Hox cluster has been the Rosetta Stone of comparative developmental biology, but its origins are unclear. Hox genes are characteristically organized in clusters whose genomic organization directly reflects domains of expression along the A/P axis [1–7]; this pattern of organization is functionally important and has been conserved across the Bilateria. The central role of Hox clusters in axial patterning in animals with very different body plans, together with functional data from arthropods and chordates, has led to the assumption that much of the morphological variation seen across the animal kingdom can be directly attributed to different numbers of Hox genes or differential use of the Hox system [9–11]. For present purposes, we define a canonical Hox system as a set of closely linked and interacting homeobox genes that are directly related to the Hox classes of *Drosophila* and mammals and that, through their combined actions, are primarily responsible for patterning most or all tissues along the anterior-posterior body axis (cf. [8, 12, 13]).

Cnidarians represent a key transition in the evolution of animal complexity and are therefore critical to understanding the origins of developmental mechanisms such as the Hox system. Although they are among the simplest of true animals at the morphological level, the Cnidaria is among the most taxon-rich phyla and cnidarians have many of the genes traditionally assumed

*Correspondence: bernd.schierwater@ecolevol.de (B.S.); david.miller@jcu.edu.au (D.J.M.)

Current Biology 16, 920–926, May 9, 2006 DOI 10.1016/j.cub.2006.03.036

to have arisen in the context of vertebrate complexity [14–17]. Cnidarians have genes clearly related to a number of the key homeobox gene families of bilateral animals, such as Emx, Evx, Hex, Not, and Dlx [18], and some of these are expressed in patterns strikingly like those of their putative bilaterian orthologs (reviewed in [19]). In addition, Hox-like genes have been identified in a wide variety of cnidarians (e.g., [18]) but, in contrast to a number of other key regulatory gene types, their status is often equivocal.

In an attempt to clarify the evolutionary origins of the Hox system, we characterized the Hox-like genes in two representative cnidarians, *Eleutheria dichotoma* (Hydrozoa) and *Nematostella vectensis* (Anthozoa), in terms of sequence relationships, genomic organization, and expression patterns. *Eleutheria* is a typical cnidarian in having both polyp and medusa lifecycle stages, while *Nematostella* represents the basal cnidarian class (Anthozoa). Genes related to the anterior Hox and posterior Hox/Cdx types of bilaterians are present, but most of the Hox-like genes present in cnidarians postdate the Cnidaria/ Bilateria split. The organization of these genes differs between the two cnidarians, and we found no evidence for the clustered organization characteristic of true Hox genes. Patterns of expression of the corresponding genes also differ dramatically between *Nematostella* and *Eleutheria* and across a range of other cnidarians. The cnidarian genes therefore do not conform to the Hox paradigm in terms of structure, organization, or expression, and the simplest interpretation of these observations is that the Cnidaria predate the origins of the Hox system. Thus, contrary to expectations, a definitive Hox system is not essential for axial patterning in lower animals. Moreover, the spectacular range of morphological variation across the Cnidaria demonstrates that the canonical Hox system is not mandatory for the elaboration of a wide range of variations on a basic body plan.

Phylogenetic Analyses Reveal No Clear Orthologies to True Hox Classes

When the Gsx-type genes (which are clearly distinct) are excluded, a total of four Hox-like genes are present in *Eleutheria* [20, 21], and five genes of this type have previously been identified in *Nematostella* [22]. In the present study, we identified three novel Hox-like genes by analysis of the unassembled genomic sequence data now available for *Nematostella*. One of these, designated as anthox8A, encodes an identical homeodomain to anthox8 but differs significantly outside this. In addition, a gene related to anthox6 (anthox6A) and a highly derived gene (anthox9) were identified. The derived position of anthox9 is reflected in its position in the phylogenetic trees; note that although its predicted sequence has the Hox-like characteristic Glu residue at homeodomain position 15, the Ile residue at position 16 appears to be without precedent in the Antp superclass (only LIM and atypical homeodomains have anything other than a Leu residue at position 16 [23]).

To investigate relationships between the cnidarian Hox-like genes and the true Hox classes of higher animals, phylogenetic analyses were conducted with all of available sequences from *Nematostella* and *Eleutheria* (Figure 1 and Supplemental Data available with this article online). The results confirm several key aspects of previously published studies [7, 13, 18, 24] but do not support others. *Nematostella* anthox1/1A and anthox7/8/8A have been duplicated independently, as have several other homeobox gene loci in *Acropora* [25]. Clear support was obtained for common origin for anterior Hox genes and a class of cnidarian genes related to *Nematostella* anthox6 that also includes anthox6A and *Eleutheria* Cnox-5ed. Our analyses also confirm the high similarity between a group of cnidarian genes that includes *Eleutheria* Cnox-4ed and posterior Hox genes, although interestingly *Nematostella* appears not to have a gene of this type. Here we refer to the classes of cnidarian genes that are related to the anterior and posterior Hox/Cdx groups as the “anterior Hox-like” and “posterior Hox/Cdx-like” types, respectively. However, whereas the affinities of these gene types are well supported, other cnidarian Hox-like genes have no clear relationship to true Hox classes. Whereas orthology relationships have sometimes been suggested, a simpler interpretation is that the common ances-

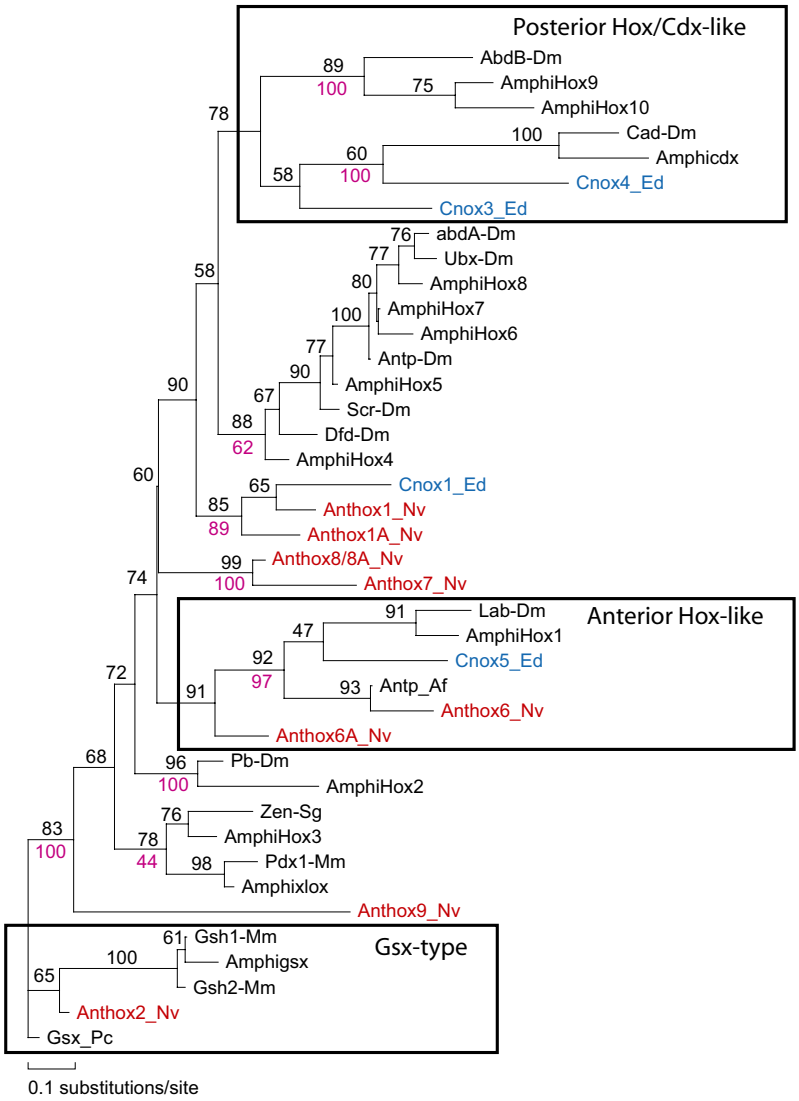


Figure 1. Phylogenetic Analyses Identify Some Cnidarian Hox-like Genes as Relatives of Anterior Hox or Posterior Hox/Cdx Types of Bilateria, but "Intermediate" Genes Are Missing

Numbers on branches reflect the percentages of 1000 ML bootstrap replicates supporting the topology shown. Bayesian posterior probability values are shown below some of the critical nodes. *Nematostella* sequences are coded in red, those from *Eleutheria* in blue. As it is not clear what might be the most appropriate outgroup, and the nature of the outgroup to some extent determines internal topology (since it affects the position of the root), analyses were unrooted. However, for clarity, the tree is shown as if rooted via the *Podocoryne carnea* Gsx sequence (Gsx_Pc; encoded by GenBank #AAG09805), which is the ortholog to *Nematostella* anthox2. Note that although some cnidarian sequences are related to the "anterior" (group 1) Hox or posterior Hox/Cdx classes of bilateral animals, true intermediate (groups 2–8) Hox genes are clearly resolved from the cnidarian sequences. Conversely, independent duplications have increased Hox-like gene complexity within the Cnidaria.

tor had genes that later gave rise to the anterior and posterior Hox genes, but that intermediate Hox genes (and hence Hox clusters) postdate the split between cnidarians and bilaterians.

Cnidarian Hox-like Genes Are Not Clustered and Their Organization Is Not Conserved within the Phylum

Because the colinear and uninterrupted structure of Hox clusters has been conserved across the Bilateria, we examined the genomic organization of the Hox-like genes in both *Eleutheria* and *Nematostella* - by fosmid cloning in the case of *Eleutheria* and by assembling genomic contigs from GenBank for *Nematostella*. Figure 2 summarizes mapping data for each of these genes.

anthox6 and *Eleutheria* Cnox-5ed are probable orthologs (Figure 1), as are anthox1 and Cnox-1ed, but all of these are flanked by unrelated non-Hox genes in the respective genomes. anthox6A and Cnox-4ed again are flanked by unrelated genes. As in another anthozoan [26], an even-skipped gene is tightly linked to anthox6 in *Nematostella*. This example of tight linkage is in contrast with the mapping data presented for other Hox-like genes and implies that if an equivalent of a Hox cluster were present, we might reasonably expect to have found other linked genes in the range of the assembled contigs. Although we found no evidence for organization characteristic of true Hox genes, the independently duplicated Hox-like genes anthox7, anthox8, and anthox8A (Figure 1) all lie within approximately 20 kb and in the same

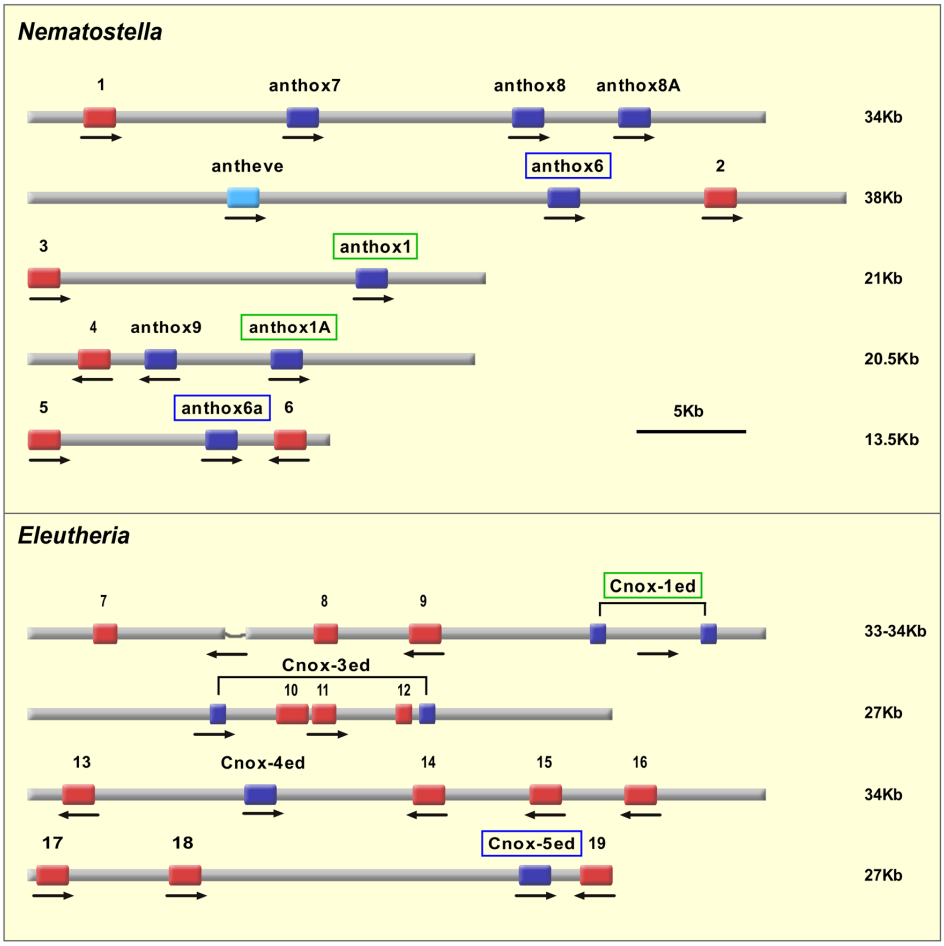


Figure 2. The Hox-like Genes of *Nematostella* and *Eleutheria* Are Organized Differently and Do Not Reflect the Clustered Patterns Characteristic of True Hox Genes

The figure summarizes schematically the genomic organization of Hox-like genes in *Nematostella vectensis* and *Eleutheria dichotoma*. In the case of *Nematostella*, contigs were assembled from GenBank whole-genome shotgun trace files, whereas *Eleutheria* contigs represent genomic fosmid clones. Rectangles show the position of the genes in the genomic context but only approximately represent the sizes of the genes. Details of intron-exon structure have been omitted except in the cases of Cnox-1ed and Cnox-3ed, both of which contain large introns. Arrows show transcriptional orientation of the genes. Hox-like genes are in dark blue with probable orthologs between *Nematostella* and *Eleutheria* framed with the same color. Hox-related genes are in light blue and non-hox genes are in red and numbered: **1**, retrotransposon; **2**, Rho/Rsa-related gene; **3**, putative metalloproteinase inhibitor; **4**, putative ANF-receptor; **5**, Phosphatidylinositol phosphate kinase; **6**, putative Iam_G domain; **7**, **8**, **10–12**, fragmented reverse transcriptases; **9**, homolog to *Danio rerio* putative protein; **13**, reverse transcriptase; **14**, resembles metabotropic glutamate receptor; **15**, two or more ORFs resembling *Homo sapiens* put. protein; **16**, last exon of RRN3; **17**, Dfp domain gene; **18**, POU gene; **19**, 1.2 Kb ORF resembling *Plasmodium* MAEBL (interrupted by vector).

orientation. Although anthox1A and anthox9 are linked, these may also be paralogs even though the latter sequence is highly derived (the Ile/Leu substitution at position 16 suggests that anthox9 might even be a pseudogene); there is no support for the alternative hypothesis that these are orthologs of different Hox classes. The linkage of paralogous homeobox genes has clear precedents [25] and should not be confused with the clustering characteristic of true Hox genes. In a true Hox cluster, no non-Hox genes lie within the cluster, so the identification of neighboring non-Hox genes implies that the Hox-like genes are not clustered in *Eleutheria* and *Nematostella*. These observations are consistent with a previous study in which the corresponding *Hydra* genes were shown not to be linked within a range of 150 kb [18]. Hence, although paralogous genes are in some cases linked, three representative cnidarians lack any equivalent of a true Hox cluster.

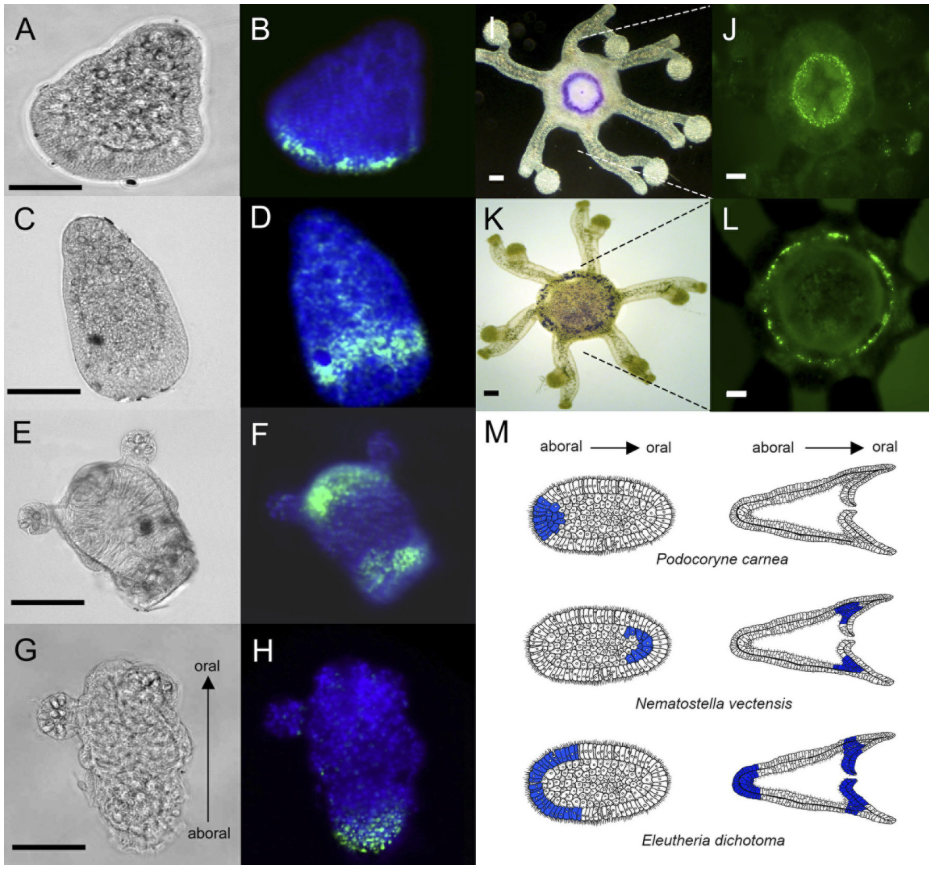


Figure 3. Expression Patterns of Related Hox-like Genes Are Heterogenous among Cnidarians and Do Not Conform to Colinearity Rules

The "anterior Hox-like" genes (= Cnox-5ed in *Eleutheria*, anthox6 in *Nematostella*, and Cnox-1pc in *Podocoryne*) provide a striking example of heterogeneity in expression patterns. In planulae, these orthologs are expressed ectodermally and aborally (*Eleutheria*; **[A–D, M]**), entodermally and orally [22] (*Nematostella*; **[M]**), or aborally in both ecto- and endoderm (*Podocoryne* [38]; **[M]**). In polyps, the corresponding genes are not expressed in *Podocoryne*, but are expressed both orally and aborally in *Eleutheria* **(E, F, M)** and at the oral end only in *Nematostella* **(M)**. **(A–F)** Cnox-5ed in 3-day-old planula larvae **(A, B)**, 4- to 5-day-old planula larvae (signal moves toward the center as an ectodermal ring; **[C, D]**), and in young primary polyps (oral and aboral; **[E, F]**). **(G and H)** Aboral expression of Cnox-4ed in a 10-day-old primary polyp.
(I and J) Ectodermal oral expression of Cnox-3ed around the manubrium of a medusa. **(K and L)** Ectodermal oral Cnox-1ed expression in the "Cnidoblast channel" of the medusa stage.
NBT/X-phosphate (I, K) or fluorescein-labeled probes (B, D, F, H, J, L). Signals in (B), (D), (F) and (H) are overlayed with DAPI staining. Morphologies are shown in light microscopy (A, C, E, G). Scale bar equals 50 mm.

Noncolinear Expression of Cnidarian Hox-like Genes

Although the cnidarian Hox-like genes are not organized in clusters, preservation of tight linkage and uninterrupted organization appears to be strictly required only for the maintenance of temporal colinearity [7, 27–29]; for example, in the derived tunicate *Oikopleura*, the remaining Hox genes are expressed in a spatially colinear fashion despite the cluster having completely fragmented [30]. If an ancestral Hox cluster had also been fragmented in cnidarians, then conservation of spatial colinearity would be predicted. To test this hypothesis, the spatial expression patterns of *Eleutheria* Hox-like genes were determined and compared with the corresponding *Nematostella* genes [22]. The most informative direct comparisons of expression patterns can be made in the developing planula larvae (which most likely reflects the phylotypic stage). In *Eleutheria*, only Cnox-5ed is expressed in planulae; as the *Podocoryne* ortholog (Cnox-1pc; Figure 3M), this "anterior Hox-like" gene is expressed at the aboral end (i.e., the front end with respect to swimming direction; Figures 3A– 3D). In the polyp, Cnox-5ed is expressed at both aboral and oral ends (Figures 3E and 3F), and

Cnox-4ed is expressed aborally (Figures 3G and 3H). Cnox-3ed and Cnox-1ed are expressed exclusively in the medusa, in ectodermal regions around the mouth and manubrium (Figures 3I–3L). The embryonic and larval expression patterns of the *Nematostella* Hox-like genes differ markedly to those of their *Eleutheria* counterparts. anthox6 corresponds to *Eleutheria* Cnox-5ed but is expressed at the opposite end of the planula, in the invaginated endoderm at the oral extremity [22]. *Nematostella* lacks a "posterior Hox/Cdx-like" gene; anthox1, which is most similar to Cnox-1ed, is expressed in the ectoderm at the aboral extremity of the *Nematostella* planula. anthox1A, 7, and 8 are expressed in the endoderm along one side of the body column [22], but not in overlapping patterns like those of true Hox genes. Moreover, as is clear not only in our analyses (Figure 1) but also in previous studies [22], those genes with axially restricted expression patterns in *Nematostella* (anthox1/1A and anthox7/ 8) have been independently duplicated in the Cnidaria, and hence any apparent similarities in expression patterns cannot reflect conservation of function with bilateral animals. In summary, expression patterns of related genes differ dramatically across the Cnidaria, and there is no evidence to support conservation of function with true Hox genes.

Implications for the Origin of Bilaterian Hox Clusters

Whereas the consensus view has been that a Hox cluster was present in the ancestral cnidarian (e.g., [13]), our analyses of sequence relationships, gene organization, and expression data indicate that definitive Hox clusters are not present in cnidarians and are therefore a synapomorphy of the Bilateria. The situation in cnidarians is therefore very different to that even in very derived members of the Bilateria. For example, whereas in urochordates the ancestral Hox cluster has fragmented, the individual genes show high levels of sequence identity and similar (A/P-restricted) patterns of expression to their orthologs in other bilaterians [30]. In cnidarians, not only are the genes dispersed, but also there are no clear relationships in terms of expression patterns or sequence identity. Cnidarians have genes related to anterior and posterior Hox/Cdx genes, but most of the Hox-like genes present are likely to postdate divergence with the bilaterian line, accounting for their unclear relationships to true Hox classes. The Hox cluster presumably arose from the outside in [31] - from a two-gene state via a series of unequal crossing-over events - and the cnidarian "anterior Hox-like" and "posterior Hox/Cdx-like" types may be derived from these ancestral two outer genes. A similar "two-gene" model of Hox cluster origin in Bilateria has been proposed recently to accommodate conflicting views about the nature of cnidarian Hox-like genes [8]. Moreover, whereas the linkage of even-skipped and Hox-like genes in anthozoans [22, 26] (Figure 2) has been interpreted as evidence for a Hox cluster in the common ancestor of Cnidaria and Bilateria [13], our data rather imply that this linkage reflects an even older array of Antp superclass gene precursors [32, 33], predating the definitive Hox system. A survey of the trace archive indicates that most of the non-Hox Antp type genes are present in *Nematostella* (e.g., EHGbox, Evx, Mox, Dlx, Msx, Emx, NK; K.K. and B.S., unpublished data). Thus, the cnidarians must have split off the lineage leading to Bilateria after the ancestors of the Antp subclasses had emerged, but before having a canonical Hox system. The alternative (less parsimonious) scenario is that in those cnidarians examined to date, an ancestral cluster of Hox genes has fragmented, and both individual sequences as well as expression patterns diverged beyond recognition.

Conclusions

The Cnidaria is among the most species-rich and diverse of phyla, indicating that neither true Hox genes nor Hox clusters are strictly required for the elaboration of morphological diversity. Moreover, the Hox system is clearly not mandatory for axial patterning. In the absence of a Hox system, other genes may be able to pattern the primary body axis - for example, *Nematostella* has the full complement of Wnt genes and these may play major roles in patterning the O/A axis [16]. It is also possible that the Hox system rep-

resents an intercalation between the “head” and “tail” domains of the common metazoan ancestor of cnidarians and bilaterians [34]. In summary, the evidence implies that true clustered Hox genes evolved in UrBilateria after the Cnidaria diverged. Although it may have facilitated morphological diversification within the Bilateria, the spectacular range of shapes and forms shown by extant cnidarians - from microscopic solitary polyps to colonial siphonophores up to 40 m long - shows the extent of variation possible even in the absence of a Hox cluster.

Experimental Procedures

Library Construction for *Eleutheria dichotoma*

Fosmid libraries were constructed with the CopyControl Fosmid Library Production Kit (Epicentre), with some steps modified to achieve a better efficiency in library construction for this AT-rich species. The detailed protocol can be obtained from the authors upon request. The libraries were screened by PCR for Cnox-1 to –5 fosmid clones and the clones were isolated from the pooled libraries by outdilution. DNA from the clones was isolated and randomly sheared to 1–2 kb fragments. The fragments were end-repaired and ligated into plasmid vectors, and the resulting subclones were sequenced on a MegaBACE1000 system with the DYEnamic ET Dye Terminator Cycle Sequencing Kit (Amersham). Sequences were assembled with the SeqMan software (DNAStar, Lasergene). Fosmid sequences were screened for potential genes by blastp and blastx.

Database Search and Contig Construction

To screen the *Nematostella* genome for new Hox-like genes, representative genes from cnidarians and *Amphioxus* were blasted against the *Nematostella* trace archive by means of Discontiguous Megablast. Positive hits were elongated and verified by retrieving and assembling (SeqMan) more sequences with Megablast. For the construction of large contigs for *Nematostella* Hox-like genes, initial contigs assembled around known sequences were elongated by searching (Megablast) the trace archive by means of the contig ends (100–300 bp segments). The ends used for elongation had to be supported by at least three overlapping sequences. The total coverage of the contigs by sequence length was 10- to 16-fold (the trace files used for the respective contigs can be obtained from the authors upon request).

In Situ Hybridization

In situ hybridization for *Eleutheria dichotoma* was carried out as previously described [35].

Phylogenetic Analyses

The homeodomains encoded by all of the known Hox-like genes from *Nematostella* and *Eleutheria* were aligned with the full suites of *Amphioxus* and *Drosophila* Hox and “Para-Hox” sequences (see [36]) and then subjected to Maximum Likelihood analyses with MolPhy version 2.3 as previously described [14] (the Dayhoff matrix was used). 1000 ML bootstrap replicates were used to test tree topology. In addition, Bayesian analyses were conducted to provide further support for aspects of the ML topology. For this purpose, we used the mixed model option in MrBayes version 3.0b4 [37] with the default setting of four Markov chains per run. The analyses were run for a total of 1.5 million generations, sampling every 1000th tree. Log Likelihood values reached a plateau after approximately 20,000 generations. One third of the resulting trees were discarded as the burn-in phase, and the remainder used to estimate posterior probabilities.

Supplemental Data

Supplemental Data include one figure, one table, and nine FASTA files of sequences and can be found with this article online at http://www.current-biology.com/cgi/content/full/16/9/920/DC1/.

Acknowledgments

K.K. was supported by an Evangelisches Studienwerk e.V. Villigst scholarship. The work was supported by grants from the German Science Foundation (DFG) and the Human Frontier Science Program (both to B.S.) and the Australian Research Council (to D.J.M. directly and via the Centre for the Molecular Genetics of Devel-

opment and the Centre of Excellence for Coral Reef Studies). Carl Hauenschild introduced us to *Eleutheria dichotoma* as a hydrozoan model system. We also thank Peter Holland, Eldon Ball, and David Hayward for fruitful discussions, and Thomas Bosch and four anonymous referees for commenting on the manuscript.

Received: December 21, 2005
Revised: March 14, 2006
Accepted: March 14, 2006
Published online: March 23, 2006

References

1. Duboule, D., and Dolle, P. (1989). The structural and functional organization of the murine HOX gene family resembles that of *Drosophila* homeotic genes. EMBO J. 8, 1497–1505.
2. Graham, A., Papalopulu, N., and Krumlauf, R. (1989). The murine and *Drosophila* homeobox gene complexes have common features of organization and expression. Cell 57, 367–378.
3. McGinnis, W., and Krumlauf, R. (1992). Homeobox genes and axial patterning. Cell 68, 283–302.
4. Duboule, D. (1992). The vertebrate limb: a model system to study the Hox/HOM gene network during development and evolution. Bioessays 14, 375–384.
5. Duboule, D. (1994). Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. Dev. Suppl., 135–142.
6. Krumlauf, R. (1994). Hox genes in vertebrate development. Cell 78, 191–201.
7. Ferrier, D.E., and Minguillon, C. (2003). Evolution of the Hox/ParaHox gene clusters. Int. J. Dev. Biol. 47, 605–611.
8. Garcia-Fernandez, J. (2005). Hox, ParaHox, ProtoHox: facts and guesses. Heredity 94, 145–152.
9. Gellon, G., and McGinnis, W. (1998). Shaping animal body plans in development and evolution by modulation of Hox expression patterns. Bioessays 20, 116–125.
10. Prince, V. (2002). The Hox paradox: more complex(es) than imagined. Dev. Biol. 249, 1–15.
11. Amores, A., Suzuki, T., Yan, Y.L., Pomeroy, J., Singer, A., Amemiya, C., and Postlethwait, J.H. (2004). Developmental roles of pufferfish Hox clusters and genome evolution in ray-fin fish. Genome Res. 14, 1–10.
12. Schierwater, B., and Desalle, R. (2001). Current problems with the zootype and the early evolution of Hox genes. J. Exp. Zool. 291, 169–174.
13. Ferrier, D.E., and Holland, P.W. (2001). Ancient origin of the Hox gene cluster. Nat. Rev. Genet. 2, 33–38.
14. Kortschak, R.D., Samuel, G., Saint, R., and Miller, D.J. (2003). EST analysis of the cnidarian *Acropora millepora* reveals extensive gene loss and rapid sequence divergence in the model invertebrates. Curr. Biol. 13, 2190–2195.
15. Fedders, H., Augustin, R., and Bosch, T.C. (2004). A Dickkopf- 3- related gene is expressed in differentiating nematocytes in the basal metazoan *Hydra*. Dev. Genes Evol. 214, 72–80.
16. Kusserow, A., Pang, K., Sturm, C., Hrouda, M., Lentfer, J., Schmidt, H.A., Technau, U., von Haeseler, A., Hobmayer, B., Martindale, M.Q., et al. (2005). Unexpected complexity of the Wnt gene family in a sea anemone. Nature 433, 156–160.
17. Technau, U., Rudd, S., Maxwell, P., Gordon, P.M., Saina, M., Grasso, L.C., Hayward, D.C., Sensen, C.W., Saint, R., Holstein, T.W., et al. (2005). Maintenance of ancestral complexity and non-metazoan genes in two basal cnidarians. Trends Genet. 21, 633–639.
18. Gauchat, D., Mazet, F., Berney, C., Schummer, M., Kreger, S., Pawlowski, J., and Galliot, B. (2000). Evolution of Antp-class genes and differential expression of *Hydra* Hox/paraHox genes in anterior patterning. Proc. Natl. Acad. Sci. USA 97, 4493–4498.
19. Ball, E.E., Hayward, D.C., Saint, R., and Miller, D.J. (2004). A simple plan—cnidarians and the origins of developmental mechanisms. Nat. Rev. Genet. 5, 567–577.
20. Kuhn, K., Streit, B., and Schierwater, B. (1996). Homeobox genes in the cnidarian *Eleutheria dichotoma*: evolutionary implications for the origin of Antennapedia-class (HOM/Hox) genes. Mol. Phylogenet. Evol. 6, 30–38.
21. Schierwater, B., and Kuhn, K. (1998). Homology of Hox genes and the zootype concept in early metazoan evolution. Mol. Phylogenet. Evol. 9,

375–381.
22. Finnerty, J.R., Pang, K., Burton, P., Paulson, D., and Martindale, M.Q. (2004). Origins of bilateral symmetry: Hox and dpp expression in a sea anemone. Science 304, 1335–1337.
23. Banerjee-Basu, S., and Baxevanis, A.D. (2001). Molecular evolution of the homeodomain family of transcription factors. Nucleic Acids Res. 29, 3258–3269.
24. Finnerty, J.R., and Martindale,M.Q. (1999). Ancient origins of axial patterning genes: Hox genes and ParaHox genes in the Cnidaria. Evol. Dev. 1, 16–23.
25. Hislop, N.R., de Jong, D., Hayward, D.C., Ball, E.E., and Miller, D.J. (2005). Tandem organization of independently duplicated homeobox genes in the basal cnidarian *Acropora millepora*. Dev. Genes Evol. 215, 268–273.
26. Miller, D.J., and Miles, A. (1993). Homeobox genes and the zootype. Nature 365, 215–216.
27. Kmita, M., Fraudeau, N., Herault, Y., and Duboule, D. (2002). Serial deletions and duplications suggest a mechanism for the collinearity of Hoxd genes in limbs. Nature 420, 145–150.
28. Spitz, F., Gonzalez, F., and Duboule, D. (2003). A global control region defines a chromosomal regulatory landscape containing the HoxD cluster. Cell 113, 405–417.
29. Chambeyron, S., Da Silva, N.R., Lawson, K.A., and Bickmore, W.A. (2005). Nuclear re-organisation of the Hoxb complex during mouse embryonic development. Development 132, 2215–2223.
30. Seo, H.C., Edvardsen, R.B., Maeland, A.D., Bjordal, M., Jensen, M.F., Hansen, A., Flaat, M., Weissenbach, J., Lehrach, H., Wincker, P., et al. (2004). Hox cluster disintegration with persistent anteroposterior order of expression in *Oikopleura dioica*. Nature 431, 67–71.
31. Gehring, W.J., Affolter, M., and Burglin, T. (1994). Homeodomain proteins. Annu. Rev. Biochem. 63, 487–526.
32. Castro, L.F., and Holland, P.W. (2003). Chromosomal mapping of ANTP class homeobox genes in amphioxus: piecing together ancestral genomes. Evol. Dev. 5, 459–465.
33. Pollard, S.L., and Holland, P.W. (2000). Evidence for 14 homeobox gene clusters in human genome ancestry. Curr. Biol. 10, 1059–1062.
34. Meinhardt, H. (2004). Different strategies for midline formation in bilaterians. Nat. Rev. Neurosci. 5, 502–510.
35. Jakob, W., Sagasser, S., Dellaporta, S., Holland, P., Kuhn, K., and Schierwater, B. (2004). The Trox-2 Hox/ParaHox gene of *Trichoplax* (Placozoa) marks an epithelial boundary. Dev. Genes Evol. 214, 170–175.
36. Brooke, N.M., Garcia-Fernandez, J., and Holland, P.W. (1998). The ParaHox gene cluster is an evolutionary sister of the Hox gene cluster. Nature 392, 920–922.
37. Huelsenbeck, J.P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.
38. Yanze, N., Spring, J., Schmidli, C., and Schmid, V. (2001). Conservation of Hox/ParaHox-related genes in the early development of a cnidarian. Dev. Biol. 236, 89–98.

Accession Numbers

Sequences of the *Eleutheria* genomic fosmids have been deposited into GenBank with the following accession numbers: Cnox-1ed DQ451870; Cnox-3ed DQ451871; Cnox-4ed DQ451872; Cnox-5ed DQ451873. Sequences of the *Nematostella* genomic contigs were assembled from NCBI trace files and are available in Document S2 with the Supplemental Data online.

Ancient Complexity of the Non-Hox ANTP Gene Complement in the Anthozoan *Nematostella vectensis*. Implications for the Evolution of the ANTP Superclass

Kai Kamm[1] and Bernd Schierwater[1,2]*

[1]ITZ, Ecology and Evolution, TiHo-Hannover, D-30559 Hannover, Germany
[2]Department of Molecular Cellular and Developmental Biology, Yale University, New Haven, Connecticut 06520-8104, USA

*Correspondence to: B. Schierwater, ITZ, Ecology and Evolution, TiHo Hannover, Bünteweg 17d, 30559 Hannover, Germany. E-mail: bernd.schierwater@ecolevol.de

This is the author's version of a work published by Wiley. Reprinted by permission of Wiley-Liss Inc. a subsidiary of John Wiley & Sons, Inc. The article has been originally published in:

Journal of Experimental Zoology (Mol Dev Evol) 306B:589–596 (2006)
Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jez.b.21123.

The original article can be found online at: www3.interscience.wiley.com/cgi-bin/jhome/102521318

Ancient Complexity of the Non-Hox ANTP Gene Complement in the Anthozoan *Nematostella vectensis*. Implications for the Evolution of the ANTP Superclass

Kai Kamm[1] and Bernd Schierwater[1,2]*

[1]ITZ, Ecology and Evolution, TiHo-Hannover, D-30559 Hannover, Germany
[2]Department of Molecular Cellular and Developmental Biology, Yale University, New Haven, Connecticut 06520-8104, USA

Abstract

The origin and evolution of ANTP superclass genes has raised controversial discussions. While recent evidence suggests that a true Hox cluster emerged after the cnidarian bilaterian split, the origin of the ANTP superclass as a whole remains unclear. Based on analyses of bilaterian genomes, it seems very likely that clustering has once been a characteristic of all ANTP homeobox genes and that their ancestors have emerged through several series of cis-duplications from the same genomic region. Since the diploblastic Cnidaria possess orthologs of some non-Hox ANTP genes, at least some steps of the expansion of this hypothetical homeobox gene array must have occurred in the last common ancestor of both lineages - but it is unknown to what extent. By screening the unassembled *Nematostella* genome, we have identified unambiguous orthologs to almost all non-Hox ANTP genes which are present in Bilateria - with the exception of En, Tlx and (possibly) Vax. Furthermore, *Nematostella* possesses ANTP genes that are missing in some bilaterian lineages, like the rough gene or NK7. In addition, several ANTP homeobox gene families have been independently duplicated in *Nematostella*. We conclude that the last cnidarian/bilaterian ancestor already harboured the almost full complement of non-Hox ANTP genes before the Hox system evolved.

Introduction

Homeobox genes are a large and diverse family of transcription factors characterized by the presence of a conserved 180 bp sequence encoding a DNA binding motif, the homeodomain. All homeobox genes play important roles in metazoan development and cell proliferation, making them important tools for studying the evolution of genomes, bauplan and developmental patterns. Thus, unravelling the genealogy of these genes is a crucial effort. Homeobox genes can be subdivided into distinct classes: LIM, POU, Atypical, Paired and the so-called ANTP superclass (Galliot et al., 1999; Gauchat et al., 2000; Banerjee-Basu and Baxevanis, 2001). The latter is comprised of the Hox, extended Hox and the NK-like (NKL) genes and seems to be restricted to the Metazoa (Holland, 2001; Holland and Takahashi, 2005).

The origin of Hox genes has been controversially discussed (reviewed in Garcia-Fernandez, 2005b). This subclass of ANTP genes is organized in tight clusters in the genomes of most higher metazoans and their products provide positional information along the primary body axis (reviewed in Ferrier and Minguillon, 2003). Since cnidarians possess genes related to Hox genes of higher Metazoa (e.g., Schierwater et al., 1991; Kuhn et al., 1996, 1999), it has been speculated

Supplemental material for this article can be found on the JEZ website at www.mrw.interscience.wiley.com/suppmat/1552-5007/suppmat/jez.b.21123.html.

Grant sponsor: Evangelisches Studienwerk e.V. Villigst Scholarship; Grant sponsor: Deutsche Forschungsgemeinschaft; Grant numbers: DFG Schi 277/10-2, Schi 277/20-1; Grant sponsor: Human Frontier Science Program; Grant number: HF-SPRGP0221/2001/-M.

*Correspondence to: B. Schierwater, ITZ, Ecology and Evolution, TiHo Hannover, Bünteweg 17d, 30559 Hannover, Germany. E-mail: bernd.schierwater@ecolevol.de

Published online 12 July 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jez.b.21123.

in different directions to what extent metazoans share this important set of genes (e.g., Slack et al., 1993; Ferrier and Holland, 2001; Schierwater and Desalle, 2001). Most importantly, recent evidence suggests that Cnidaria predate the origin of the Hox system (Kamm et al., 2006). Therefore, cnidarian genes that are related to bilaterian Hox genes should rather be regarded as Hox-like but not as true Hox genes.

Whereas it seems likely that Cnidaria split off the lineage leading to Bilateria before a true Hox cluster developed, the origin of the whole ANTP superclass remains unclear. Recent data indicate that clustering might once have been a characteristic of all ANTP homeobox genes (reviewed in Garcia-Fernandez, 2005a). The need for spatiotemporal expression of Hox genes during development might be the constraining force that keeps the Hox cluster intact. In contrast, these constrains seem to be much lower for non-Hox ANTP genes and thus they have become scattered around in different genomes (Ferrier and Minguillon, 2003; Luke et al., 2003; Garcia-Fernandez, 2005a). Nevertheless, remnants of clusters exist in many bilaterian animals (Coulier et al., 2000; Pollard and Holland, 2000; Luke et al., 2003). Moreover, sequence relationships among the ANTP genes, combined with comparative gene mapping in the genomes of *Drosophila*, human, mouse and *Amphioxus* suggest that the ancestors of all ANTP gene families have once evolved together on an ANTP mega array (Coulier et al., 2000; Pollard and Holland, 2000; Castro and Holland, 2003). Earlier studies revealed that cnidarians have clear orthologs to some non-Hox ANTP genes such as Evx, Emx, Hex, Not or Dlx (Gauchat et al., 2000). In Placozoa, which are probably more basal (cf. Ender and Schierwater, 2003; Miller and Ball, 2005; Schierwater, 2005), we find a very low diversity of these genes (Monteiro et al., 2006). Therefore, at least some steps of the expansion of the ANTP superclass must have taken place in the cnidarian bilaterian ancestor; but only an almost complete assessment of cnidarian non-Hox ANTP genes can provide insights into the origin and evolution of these genes. It has recently been shown that the common ancestor of cnidarians and bilaterians already had a very complex genome, which has been retained in basal cnidarians like *Nematostella* and *Acropora* (Kortschak et al., 2003; Miller et al., 2005; Technau et al., 2005). Hence we screened the full, though unassembled, *Nematostella* genome for non-Hox ANTP superclass genes and inferred phylogenetic relationships between *Nematostella* non-Hox ANTP genes and their bilaterian counterparts. By taking advantage of existing linkage information from the literature and genome databases, we deduced a possible ANTP gene complement of the last common ancestor of Cnidaria and Bilateria.

Materials and Methods

Sequence acquisition

Representatives of bilaterian non-Hox ANTP gene families were blasted against the *Nematostella* whole genome shotgun trace archive at NCBI. Positive hits were elongated and corrected by retrieving more sequences representing the same genomic region. Open Reading Frames containing the homeoboxes were translated into homeodomains. Bilaterian ANTP protein sequences, as well as sequences for *Nematostella* NK1, 2, 3, 4 and Mox2, were obtained from the NCBI protein database. For *Nematostella* Mox1 and *Nematostella* Evx (anth-eve), where only incomplete homeoboxes are deposited in the databases, the available sequences were elongated with the trace files. In the case of NvEvx, a 38Kb genomic contig, containing also Anthox6, has been reported in Kamm et al. (2006).

Phylogenetic analyses and linkage patterns

The 60 amino acid residues of the homeodomains were aligned with ClustalW (Thompson et al., 1994) implemented in MEGA v3.1 (Kumar et al., 2004) and Bayesian analysis with MrBayes v3.1.1 was conducted to infer tree topologies (Ronquist and Huelsenbeck, 2003). Four chains and two independent simultaneous runs over 500,000 generations were used, sampling every 25th tree. The likelihoods of the generations were examined to estimate the beginning of the stationary phase and the trees after the first 125,000 generations were used to create a

consensus tree. We used the mixed model of amino acid substitution, assuming the presence of invariant sites and using a gamma distribution approximated by four different rate categories to model rate variation between sites. For a Neighbour Joining analysis, the alignment was tested for an appropriate model of amino acid substitution with ProtTest v1.2.6 (Abascal et al., 2005) using six rate categories for estimation of gamma. A NJ-Tree was then inferred using MEGA with the JTT model of amino acid substitution (Jones et al., 1992) and a gamma distribution for rate variation among sites. The tree was tested with 2,000 bootstrap replicates. Tree topologies were also assessed by Maximum Likelihood using PhyML v.2.4.4 (Guindon and Gascuel, 2003). The JTT model and the gamma value estimated with ProtTest were used. A consensus tree was calculated from 1,000 bootstrap resamplings. Trees were examined using the TreeExplorer implemented in MEGA and rooted with the *Nematostella* Hox-like gene Anthox6. Sequence accession numbers and alignment are available as supplementary material.

Linkage patterns of ANTP superclass genes were obtained from the literature (Coulier et al., 2000; Pollard and Holland, 2000; Castro and Holland, 2003; Luke et al., 2003; Minguillon et al., 2005), and by blasting ANTP genes against the current genome releases at NCBI.

Results and Discussion

A total of 34 unambiguous non-Hox ANTP genes from *Nematostella* were identified and used for analyses of phylogenetic relationship to their bilaterian counterparts. These include seven sequences already published in NCBI plus 27 new sequences we retrieved from the *Nematostella* genome. Another seven ANTP-related sequences were not used in the final dataset, because they were highly derived and no clear orthologs could be found in the databases and initial analyses. Likewise, three (presumably) derived BarH genes were excluded (supplemental data). Table 1 summarizes the identified non-Hox ANTP gene complement in *Nematostella*.

Table 1. The non-Hox ANTP gene complement in *Nematostella*

ANTP gene family	*Nematostella* orthologs
Mox/Meox/Btn	NvMox1-4
HB9/HlxB9/Mnx	NvHB9
Gbx/Unp	NvGbx
En/engrailed	-
Evx/Eve	NvEvx (anth-eve)
Rough	NvRough
Emx/Ems	NvEmx1-4
Vax	? (see text)
Not	NvNot1-4
Msx/Msh	NvMsx
Dlx/Dll	NvDlx
NK1/Sax/Slou	NvNK1
Hmx	NvHmx
NK6/Gtx	NvNK6
NK7	NvNK7
NK3/Bap	NvNK3
NK2 (NK4/NK2.5/tinman - NK2.1/scarecrow - NK2.2)	NvNK4, NvNK2, NvNK2b-e
Hhex	NvHhex
Lbx/ladybird	NvLbx
Tlx/C15	-
BarH	NvBarH1-2 (and 3 presumably related
Hlx/H2.0	NvHlx
Unassigned	7 unassigned sequences

Phylogenetic analyses revealed that *Nematostella* possesses the majority of the non-Hox ANTP genes present in Bilateria, comprising 19 gene families. Furthermore, some families contain paralogs and have thus been independently duplicated in cnidarians, a phenomenon which has been described before in the coral *Acropora* (Hislop et al., 2005). In three cases (see below) no clear orthologs to bilaterian ANTP families could be identified. The reasons could be: (1) gene loss, (2) these genes are an invention of the Bilateria, (3) highly derived sequences. Favouring any of the alternatives would be premature.

Bayesian, NJ and ML analyses revealed the same gene families; only the relation between particular families differed slightly (Fig. 1 and supplemental figures). We would like to note that there is no evidence that all of these genes are functional. This does not affect the subsequent conclusions, however.

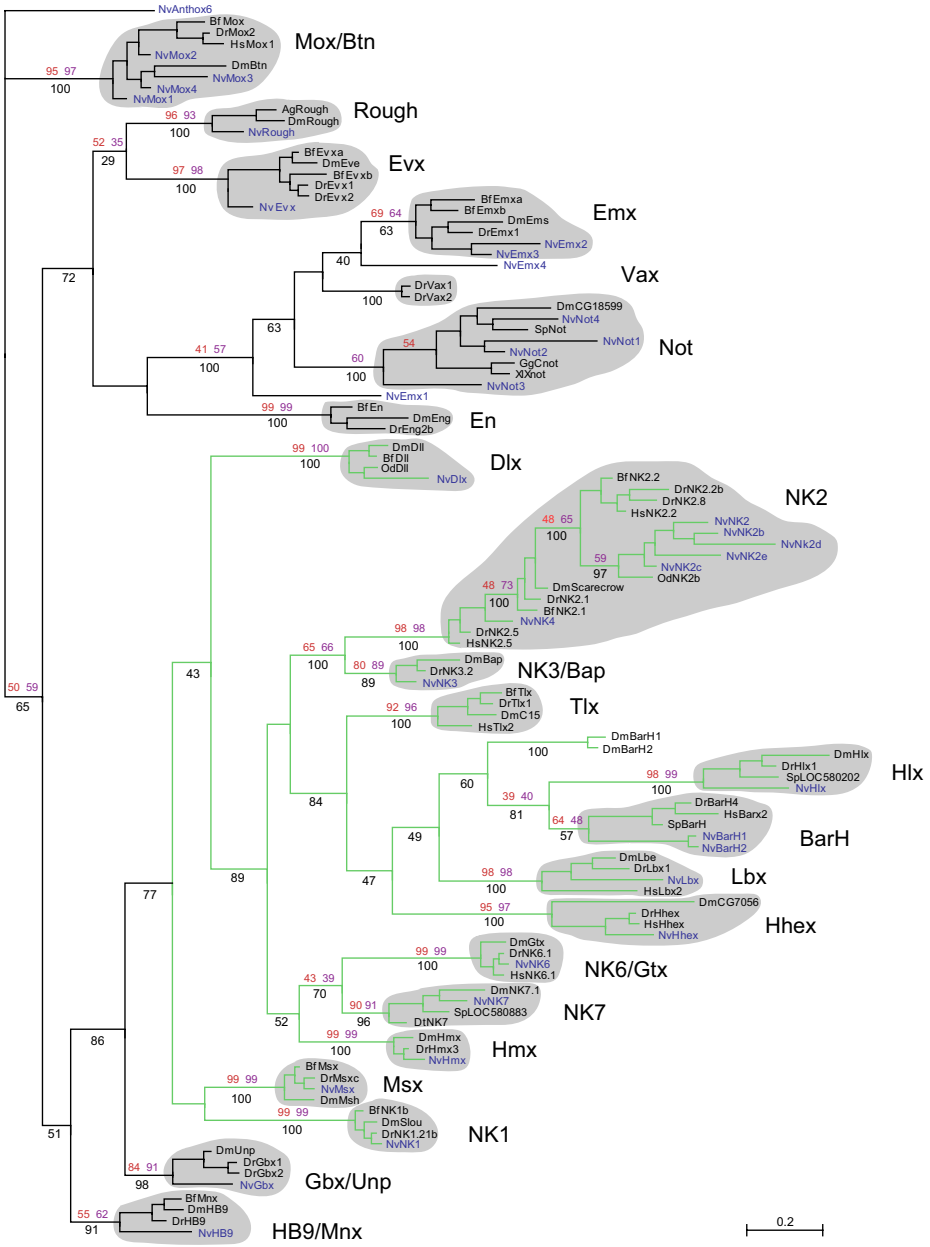


Figure 1. Phylogenetic tree of 111 metazoan non-Hox ANTP superclass genes including 34 *Nematostella* sequences. The tree was inferred using Bayesian analysis with MrBayes. Numbers below branches reflect the posterior probabilities supporting the topology shown. Percentages of 2,000 bootstrap replicates from the NJ analysis (red) and from 1,000 bootstrap resamplings of the ML consensus tree (purple) are shown above critical branches. *Nematostella* sequences are coded in blue. Branches of the NKL subclass are shown in green. (Ag=*Anopheles gambiae*, Bf=*Branchiostoma floridae*, Dm=*Drosophila melanogaster*, Dr=*Danio rerio*, Dt=*Discocelis tigrina*, Gg=*Gallus gallus*, Hs=*Homo sapiens*, Nv=*Nematostella vectensis*, Od=*Oikopleura dioica*, Sp=*Strongylocentrotus purpuratus*, Xl=*Xenopus laevis*).

The „extended Hox“ ANTP genes: Mox, Evx, Rough, Eng, HB9, Gbx, Emx, Vax and Not

We here follow the classification system from other studies (Gauchat et al., 2000; Pollard and Holland, 2000; Banerjee-Basu and Baxevanis, 2001; Minguillon and Garcia-Fernandez, 2003) although the exact composition of the extended Hox differs somewhat in the literature. For example, Emx is sometimes regarded as a member of the extended Hox (Banerjee-Basu and Baxevanis, 2001), and sometimes as an NKL gene (Gauchat et al., 2000; Pollard and Holland, 2000).

With the exception of engrailed (En) and possibly Vax genes, *Nematostella* has one or even multiple orthologs of these genes. NvMox1, NvMox2 and NvEvx have been previously identified and we could retrieve two more Mox genes

in *Nematostella* (NvMox3-4). Thus the Mox family seems to be a good example for independent duplications of homeobox genes in Cnidaria. Former analyses with Mox and Evx genes from human, mouse and *Amphioxus* suggested a monophyletic group and thus common ancestry for both genes (Minguillon and Garcia-Fernandez, 2003).

However, none of our analyses reveal support for a monophyletic origin of these genes. In contrast, the closest relatives to even-skipped seem to be the rough genes, for which we found one ortholog in *Nematostella*. Screening the genomes and ESTs of *Homo*, *Gallus*, *Danio* and *Ciona*, we could not retrieve a rough ortholog, whereas we could find a putative ortholog in the sea urchin *Strongylocentrotus* (accession #AAGJ01115187). Obviously, since the rough gene is absent in chordates, it has not been included in previous analyses.

In our analyses, the Mox genes, together with Mnx/HB9, are more closely related to the Hox-like gene Anthox6, while Evx, Rough, Gbx, En, Emx, Vax and Not occupy an intermediate position between Hox-like and NKL genes (Fig. 1). The Gbx genes themselves are possibly closely related to Mnx/HB9, which is not surprising since these genes build a homeobox cluster in vertebrate genomes (Pollard and Holland, 2000; Castro and Holland, 2003). The third member of this vertebrate mini-cluster (which has been termed the EHGbox-cluster), the engrailed family, is far more derived from the latter two, which is consistent with a previous analysis (Banerjee-Basu and Baxevanis, 2001).

Emx, Vax (could not be unambiguously identified in *Nematostella*) and Not genes form a monophyletic group in our analyses (Fig. 1). Four Emx-related genes (NvEmx1-4) could be identified in *Nematostella*, of which only two seem to be clear orthologs to bilaterian Emx genes. NvEmx1 and 4 are more derived but also clearly fall into the Emx/Vax/Not clade (Fig. 1). It may be possible that one of the “derived” Emx genes represents a descendant of the Vax gene ancestor. At least in the NJ and ML analyses NvEmx1 groups with the Vax clade, although with a bootstrap support of only 41% and 32% ,respectively (supplemental figures). Four unambiguous Not orthologs could be identified (except maybe NvNot3). Mammals seem to have lost Not genes (Martinelli and Spring, 2004) while birds (*Gallus*) and amphibians (*Xenopus*) still have it. It is noteworthy that the two Emx and Vax paralogs in human are linked once on chromosome 2 (Fig. 2A) - together with Tlx2 and Lbx2 - and once on chromosome 10 - together with Hhex, Tlx1 and Lbx1 on one side, and Hmx2 and Gtx (NK6) on the opposite side (Pollard and Holland, 2000; NCBI-MapViewer). In zebrafish, the Not ortholog floating head is closely linked to Emx1 on chromosome 13 and in chicken Cnot1 lies only 100Kb apart from an Emx ortholog on chromosome 4 (NCBI-MapViewer). Thus the chromosomal arrangement of Emx, Vax and Not in the respective genomes seems to reflect phylogenetic relationships.

The NK and NKL genes

This group includes the Dlx (Dll), Msx, (Msh), NK1 (Sax), NK2, NK3 (Bap), Hmx, NK6, NK7, Tlx, Lbx (ladybird), BarH, Hlx and Hhex genes (Gauchat et al., 2000; Pollard and Holland, 2000; Banerjee-Basu and Baxevanis, 2001; Holland, 2001). With the exception of Tlx, we could find unambiguous orthologs of all of the remaining NKL families.

The NK1 ortholog of *Nematostella* has been previously identified and clearly falls into one clade with its bilaterian counterparts. The NK2 family is more diverse in metazoans and can be further subdivided into NK2.1 (scarecrow), NK2.2 and NK2.5 (NK4/tinman). Six *Nematostella* genes belong to this group: two previously identified *Nematostella* genes, NvNK2 and NvNK4, together with four new genes (termed NvNK2 b-e). None of these can be clearly assigned to the subfamilies in bilaterians in our analyses (Fig. 1). NvNK4 has affinities to NK2.5, as well as to NK2.1, whereas NvNK2 and NvNK2b-e are more related to NK2.2, their closest bilaterian counterpart being the tunicate NK2 gene from *Oikopleura*. The latter five *Nematostella* genes are also closely related to each other, which points to a paralogous relationship. The diversification of the NK2 family might thus have happened independently in cnidarians and bilaterians, but the precursor(s) of

the NK2 family must have been present in their last common ancestor. The previously reported gene NvNK3 belongs to the NK3 family, which seems to have a monophyletic origin with the NK2 genes.

Hmx, NK6 and NK7 fall into one clade in Bayesian analysis (without Hmx in the NJ and ML analyses, however). Interestingly, the NK7 family is not represented in vertebrates (NCBI genome BLAST), which must have lost it because it is already present in insects (Coulier et al., 2000) and in the echinoderm *Strongylocentrotus* (Fig. 1).

Hhex, Tlx and Lbx form one clade, together with BarH and Hlx in Bayesian and ML analyses and a separate clade in the NJ analysis (supplemental figures). In the *Drosophila* genome, both ladybird paralogs are close neighbours of Tlx (C15), Hhex lies close by (350 Kb) on the opposite side (Fig. 2A). The same picture applies to human: Tlx1 and Lbx1 are close neighbours on chromosome 10, although Hhex is about 8Mb apart here (Pollard and Holland, 2000; Luke et al., 2003; NCBIMapViewer). Hence it is evident

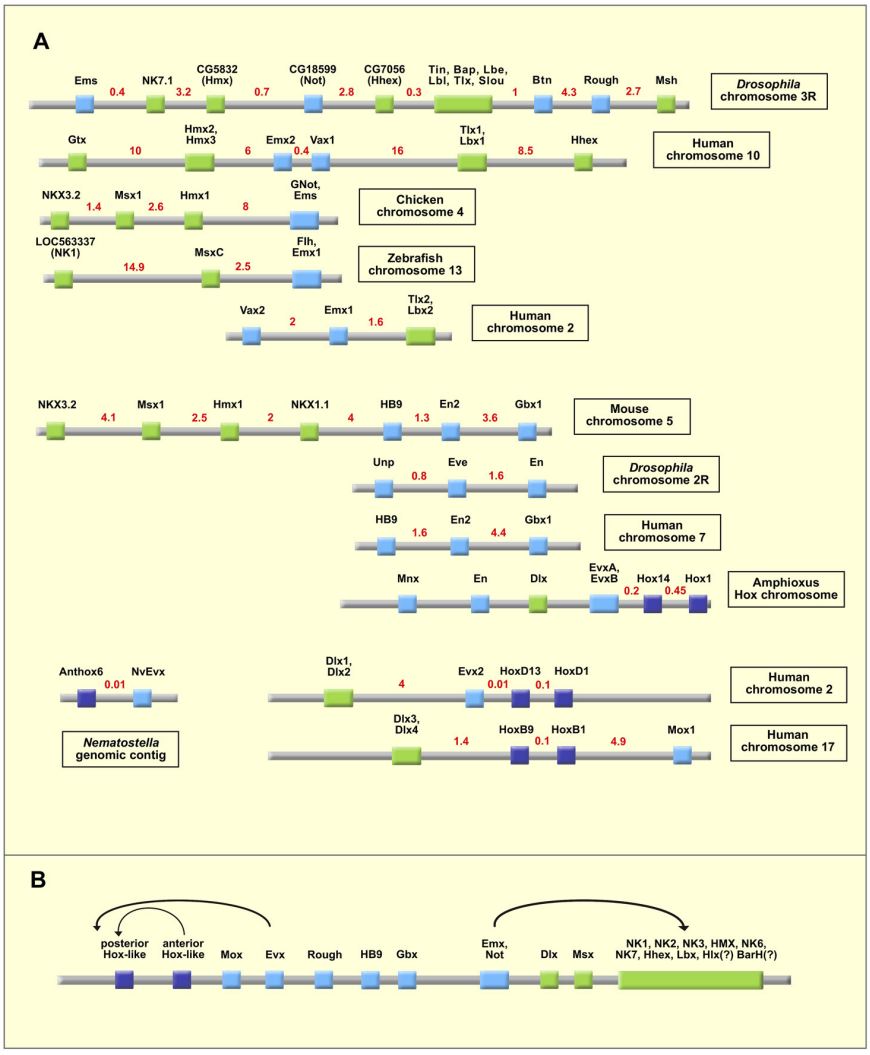


Figure 2. (A) Sample key regions of metazoan genomes with linkage of ANTP genes. The chromosomal regions are not drawn to scale. Intergenic distances are given in megabase pairs (in red) where possible. Linkage data were deduced from the literature (Coulier et al., 2000; Pollard and Holland, 2000; Castro and Holland, 2003; Luke et al., 2003; Minguillon et al., 2005) and from current releases of the respective genomes at NCBI (NCBI-MapViewer). Hox genes (or Hox-like genes in Cnidaria) are shown in dark blue, extended Hox genes in light blue and NKL genes in green. **(B)** Possible composition of an ancestral ANTP array before the cnidarian bilaterian split as deduced from phylogenetic analyses and observed patterns of gene linkage in metazoans. The order of single genes is not mandatory (it cannot be unambiguously deduced) as it is evident from sequence relationships and comparative linkage data that numerous chromosomal rearrangements, gene losses and gene duplications must have taken place. Arrows show examples for ambiguities in gene order (see text).

that Hhex, Lbx and Tlx have been neighbours in the bilaterian ancestor.

From the present survey, *Nematostella* has one ortholog of Hlx and two of BarH (maybe five, cf. Table 1). In our analyses, the *Nematostella* BarH genes are orthologs of the deuterostome BarH family. Our analyses also suggest a monophyletic origin of BarH and Hlx. In the human genome BarH and Hlx are both dispersed and show no linkage to other ANTP genes (Popovici et al., 2001). It is noteworthy that the *Drosophila* BarH genes do neither fall into the BarH nor into the BarH/Hlx clade.

The Dlx and Msx gene families have been found to be closely related in previous analyses (Gauchat et al., 2000; Banerjee-Basu and Baxevanis, 2001; Minguillon and Garcia-Fernandez, 2003). In our Bayesian analysis, there is no support for this clade (Fig. 1) and only low support in the NJ and ML analyses (supplemental figures). Independent of whether both groups have a monophyletic origin, they clearly belong to the NKL subclass.

Evolution of the ANTP superclass

As has been first pointed out by Pollard and Holland (2000) for vertebrates, and subsequently later for other lineages (Coulier et al., 2000; Castro and Holland, 2003; Garcia-Fernandez, 2005a), the linkage pattern of ANTP genes in metazoan genomes strongly points to a common origin of ANTP superclass genes from the same genomic region through several series of cis-duplications. Figure 2A depicts several key regions of linkage observed in metazoan genomes: Some members of the extended Hox genes show linkage to Hox clusters in bilaterians, others to remnants of NKL clusters. For example, Emx and Not genes are linked to NK genes in many bilaterian genomes and the Mox and even-skipped genes are linked to the Hox clusters in vertebrates, while the *Drosophila* Mox ortholog Btn and the rough gene are linked to an NK cluster on Chromosome 3R (Coulier et al., 2000; Pollard and Holland, 2000; Castro and Holland, 2003; NCBI-MapViewer). Dlx and Msx are probably closely related and both are clearly NKL genes. While Msx is linked to other NK genes, Dlx shows linkage to chordate Hox clusters (Pollard and Holland, 2000; Castro and Holland, 2003; NCBI-MapViewer). EHGbox genes also show linkage to chordate Hox clusters (Pollard and Holland, 2000; Castro and Holland, 2003). In *Nematostella* we find to date only one example of ancient linkage: like in the coral *Acropora* (Miller and Miles, 1993), the anterior Hox-like gene Anthox6 is tightly linked to NvEvx (Kamm et al., 2006).

The combined evidence from phylogenetic analyses and linkage data clearly suggest a common origin for the ANTP superclass genes. Deducing gene order in a hypothetical ancient gene array, however, is very difficult (Fig. 2B), because the ancient gene array must have undergone several rounds of rearrangements (segment inversion or translocation, cluster breakage), segment duplications and gene losses in the different lineages. For example, one of the rare mapping data for cnidarians comes from the even-skipped orthologs in *Acropora* and *Nematostella*, both of which are linked closely to the anterior Hox-like genes AfAntp and NvAnthox6, respectively (Miller and Miles, 1993; Kamm et al., 2006). In vertebrates, the two even-skipped orthologs are both tightly linked to the posterior genes of two out of the four Hox clusters, while the two Mox orthologs occupy a position adjacent to the anterior genes of two of the Hox clusters (Pollard and Holland, 2000). In *Drosophila*, even-skipped is flanked on one side by unplugged (Gbx) and on the opposite side by engrailed, while Btn lies adjacent to the tight NK cluster on chromosome 3R (Coulier et al., 2000; NCBI-MapViewer). Likewise, from the linkage pattern observed in vertebrates (Fig. 2A), we could deduce that the Emx/Vax/Not clade belongs to the NKL genes, which is not clearly supported by phylogenetic analyses (e.g., Banerjee-Basu and Baxevanis, 2001; this work). There are also oddities to resolve within chordates: Whereas in *Amphioxus* (Castro and Holland, 2003), human (Pollard and Holland, 2000), chicken and rat (NCBI-MapViewer), the EHGbox genes are linked to the same chromosome(s) as the Hox cluster(s), this is not the case in the mouse (Coulier et al., 2000; NCBI-MapViewer). Instead, in mouse HB9, En2 and Gbx1 are linked close to four NKL genes on chromosome

5 (Fig. 2A). However, early primate and rodent ancestors seem to have had a highly conserved genome organization (Murphy et al., 2005), the rate of chromosomal rearrangement increased dramatically thereafter (especially in the rodent lineage). Possibly we here see a “hot spot” region of rearrangement, with different breakpoints in the respective lineages.

It is evident that *Nematostella* already possesses the almost full complement of non-Hox ANTP genes (Table 1; Figs. 1 and 2B), including genes that have been lost in different bilaterian lineages, like NK7 and the rough gene. This observation is consistent with previous studies which identified ancestral complexity in cnidarian genomes (Kortschak et al., 2003; Miller et al., 2005; Technau et al., 2005). Most likely the last common ancestor of cnidarians and bilaterians had a gene region harbouring the almost full non-Hox ANTP superclass gene complement. Recent evidence also suggests that cnidarians represent a preHox state, with only anterior and posterior Hox-like genes (Kamm et al., 2006). We therefore conclude that the expansion of the non-Hox ANTP gene families took place before a true Hox system evolved. With the example of the Mox, Emx, Not, NK2 and BarH gene families, we can also support that the independent duplication of homeobox gene loci is a general phenomenon in Cnidaria (cf. Hislop et al., 2005).

Acknowledgements

We thank Peter Holland and David J. Miller for inspiring discussions.

K.K. was supported by an Evangelisches Studienwerk e.V. Villigst

Literature Cited

Abascal F, Zardoya R, Posada D. 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21: 2104–2105.

Banerjee-Basu S, Baxevanis AD. 2001. Molecular evolution of the homeodomain family of transcription factors. Nucleic Acids Res 29:3258–3269.

Castro LF, Holland PW. 2003. Chromosomal mapping of ANTP class homeobox genes in amphioxus: piecing together ancestral genomes. Evol Dev 5:459–465.

Coulier F, Popovici C, Villet R, Birnbaum D. 2000. MetaHox gene clusters. J Exp Zool 288:345–351.

Ender A, Schierwater B. 2003. Placozoa are not derived cnidarians: evidence from molecular morphology. Mol Biol Evol 20:130–134.

Ferrier DE, Holland PW. 2001. Ancient origin of the Hox gene cluster. Nat Rev Genet 2:33–38.

Ferrier DE, Minguillon C. 2003. Evolution of the Hox/ParaHox gene clusters. Int J Dev Biol 47:605–611.

Galliot B, de Vargas C, Miller D. 1999. Evolution of homeobox genes: Q50 paired-like genes founded the Paired class. Dev Genes Evol 209:186–197.

Garcia-Fernandez J. 2005a. The genesis and evolution of homeobox gene clusters. Nat Rev Genet 6:881–892.

Garcia-Fernandez J. 2005b. Hox, ParaHox, ProtoHox: facts and guesses. Heredity 94:145–152.

Gauchat D, Mazet F, Berney C, Schummer M, Kreger S, Pawlowski J, Galliot B. 2000. Evolution of Antp-class genes and differential expression of *Hydra* Hox/paraHox genes in anterior patterning. Proc Natl Acad Sci USA 97: 4493–4498.

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704.

Hislop NR, de Jong D, Hayward DC, Ball EE, Miller DJ. 2005. Tandem organization of independently duplicated homeobox genes in the basal cnidarian *Acropora millepora*. Dev Genes Evol 215:268–273.

Holland PW. 2001. Beyond the Hox: how widespread is homeobox gene clustering? J Anat 199:13–23.

Holland PW, Takahashi T. 2005. The evolution of homeobox genes: Implications for the study of brain development. Brain Res Bull 66:484–490.

Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8:275–282.

Kamm K, Schierwater B, Jakob W, Dellaporta SL, Miller DJ. 2006. Axial patterning and diversification in the Cnidaria predate the Hox system. Curr Biol 16:920–926.

Kortschak RD, Samuel G, Saint R, Miller DJ. 2003. EST analysis of the cnidarian *Acropora millepora* reveals extensive gene loss and rapid sequence divergence in the model invertebrates. Curr Biol 13:2190–2195.

Kuhn K, Streit B, Schierwater B. 1996. Homeobox genes in the cnidarian *Eleutheria dichotoma*: evolutionary implications for the origin of Antennapedia-class (HOM/Hox) genes. Mol Phylogenet Evol 6:30–38.

Kuhn K, Streit B, Schierwater B. 1999. Isolation of Hox genes from the scyphozoan *Cassiopeia xamachana*: implications for the early evolution of Hox genes. J Exp Zool 285: 63–75.

Kumar S, Tamura K, Nei M. 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 5:150–163.

Luke GN, Castro LF, McLay K, Bird C, Coulson A, Holland PW. 2003. Dispersal of NK homeobox gene clusters in amphioxus and humans. Proc Natl Acad Sci USA 100: 5292–5295.

Martinelli C, Spring J. 2004. Expression pattern of the homeobox gene Not in the basal metazoan *Trichoplax adhaerens*. Gene Expr Patterns 4:443–447.

Miller DJ, Ball EE. 2005. Animal evolution: the enigmatic phylum placozoa revisited. Curr Biol 15: R26–R28.

Miller DJ, Miles A. 1993. Homeobox genes and the zootype. Nature 365:215–216.

Miller DJ, Ball EE, Technau U. 2005. Cnidarians and ancestral genetic complexity in the animal kingdom. Trends Genet 21:536–539.

Minguillon C, Garcia-Fernandez J. 2003. Genesis and evolution of the Evx and Mox genes and the extended Hox and ParaHox gene clusters. Genome Biol 4:R12.

Minguillon C, Gardenyes J, Serra E, Castro LF, Hill-Force A, Holland PW, Amemiya CT, Garcia-Fernandez J. 2005. No more than 14: the end of the amphioxus Hox cluster. Int J Biol Sci 1:19–23.

Monteiro AS, Schierwater B, Dellaporta SL, Holland PW. 2006. A low diversity of ANTP class homeobox genes in Placozoa. Evol Dev 8:174–182.

Murphy WJ, Larkin DM, Everts-van der Wind A, Bourque G, Tesler G, Auvil L, Beever JE, Chowdhary BP, Galibert F, Gatzke L, Hitte C, Meyers SN, Milan D, Ostrander EA, Pape G, Parker HG, Raudsepp T, Rogatcheva MB, Schook LB, Skow LC, Welge M, Womack JE, O'Brien S J, Pevzner PA, Lewin HA. 2005. Dynamics of mammalian chromosome evolution inferred from multispecies comparative maps. Science 309:613–617.

NCBI-MapViewer. http://www.ncbi.nlm.nih.gov/mapview/

Pollard SL, Holland PW. 2000. Evidence for 14 homeobox gene clusters in human genome ancestry. Curr Biol 10:1059–1062.

Popovici C, Leveugle M, Birnbaum D, Coulier F. 2001. Homeobox gene clusters and the human paralogy map. FEBS Lett 491:237–242.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Schierwater B. 2005. My favorite animal, *Trichoplax adhaerens*. Bioessays 27:1294–1302.

Schierwater B, Desalle R. 2001. Current problems with the zootype and the early evolution of Hox genes. J Exp Zool 291:169–174.

Schierwater B, Murtha M, Dick M, Ruddle FH, Buss LW. 1991. Homeoboxes in cnidarians. J Exp Zool 260:413–416.

Slack JM, Holland PW, Graham CF. 1993. The zootype and the phylotypic stage. Nature 361:490–492.

Technau U, Rudd S, Maxwell P, Gordon PM, Saina M, Grasso LC, Hayward DC, Sensen CW, Saint R, Holstein TW, Ball EE, Miller DJ. 2005. Maintenance of ancestral complexity and non-metazoan genes in two basal cnidarians. Trends Genet 21:633–639.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673–4680.

Ancient Linkage of a POU Class 6 and an Anterior Hox-Like Gene in Cnidaria: Implications for the Evolution of Homeobox Genes

Kai Kamm[1]* and Bernd Schierwater[1,2]

[1]ITZ, Ecology and Evolution, TiHo Hannover, D-30559 Hannover, Germany
[2]Dept. of Molecular Cellular and Developmental Biology, Yale University, New Haven, CT 06520-8104, USA

*Correspondence to: Kai Kamm, ITZ, Ecology and Evolution, TiHo Hannover, Bünteweg 17d, 30559 Hannover, Germany. E-mail: kai.kamm@ecolevol.de.

The following article is a preprint version as initially submitted to the Journal of Experimental Zoology (Mol Dev Evol)

submitted: December 19, 2006
revised: April 24, 2007
accepted: July 26, 2007
published online: August 20, 2007

Journal of Experimental Zoology (Mol Dev Evol)
Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jez.b.21196.

The revised and accepted version of the article is available online at: www3.interscience.wiley.com/cgi-bin/jhome/102521318

Ancient Linkage of a POU Class 6 and an Anterior Hox-Like Gene in Cnidaria: Implications for the Evolution of Homeobox Genes

Kai Kamm[1]* and Bernd Schierwater[1,2]

[1]ITZ, Ecology and Evolution, TiHo Hannover, D-30559 Hannover, Germany
[2]Dept. of Molecular Cellular and Developmental Biology, Yale University, New Haven, CT 06520-8104, USA

Abstract

Linkage analyses in metazoan genomes suggest two super-paralogons for the majority of homeobox genes. The related homeobox genes and chromosomal regions which are dispersed in extant species possibly derived from only two single common ancestor regions. One super-paralogon, designated as ANTP mega-array, contains most of the ANTP class homeobox genes; the second, named the contraHox, would consist of the classes PRD, POU, LIM, CUT, prospero, TALE and SIX. Here we report the tight linkage of a POU class 6 gene to an anterior Hox-like gene in the hydrozoan *Eleutheria dichotoma* and discuss its possible significance for the evolution of homeobox genes. POU class 6 genes also seem to be ancestrally linked to the HoxC and A clusters in vertebrates, despite POU homeobox genes belonging to the contraHox paralogon. Hence, the much tighter linkage of a POU class 6 gene to an anterior Hox-like gene in a cnidarian is possibly the evolutionary echo of an ancestral genomic region from which most metazoan homeobox classes emerged.

Introduction

Homeobox genes are important transcription factors characterized by the presence of a conserved 180bp sequence encoding a DNA binding motif - the homeodomain (Gehring, 1985). Since the discovery of the homeobox a vast variety of them has been found in animals, plants and fungi, all of which play important roles in development and cell differentiation (Galliot et al., 1999; Banerjee-Basu and Baxevanis, 2001; Holland and Takahashi, 2005). The homeobox genes present in Metazoa can be mainly assigned to the ANTP, PRD, POU, LIM, CUT, prospero, TALE and SIX classes, most of which seem to be restricted to the animal kingdom (Galliot et al., 1999; Banerjee-Basu and Baxevanis, 2001; Holland and Takahashi, 2005). Homeobox genes of the atypical TALE-homeobox class, however, are also found in plants and fungi (Burglin, 1997), while a homeobox gene possibly related to the metazoan LIM-homeobox class is present in the slime mold *Dictyostelium* (Galliot et al., 1999).

By far the most diverse group of homeobox genes is the ANTP class which comprises Hox, ParaHox, NK, and other related genes (Garcia-Fernandez, 2005a). Within the ANTP class Hox genes have attracted particular attention because they show a remarkable similarity across Bilateria not only in sequence identity but also in genomic organization and function (Garcia-Fernandez, 2005a; Garcia-Fernandez, 2005b; Ferrier and Minguillon, 2003). Bilaterian Hox

*Correspondence to: Kai Kamm, ITZ, Ecology and Evolution, TiHo Hannover, Bünteweg 17d, 30559 Hannover, Germany. E-mail: kai.kamm@ecolevol.de.

Grant information: KK was supported by an Evangelisches Studienwerk e.V. Villigst scholarship. We acknowledge support from the Deutsche Forschungsgemeinschaft (DFG Schi 277/10-2, Schi 277/20-1) and the Human Frontier Science Program (HF-SPRGP0221/2001/-M).

genes pattern the anterior-posterior axis during development and their defining characteristic is that they are organised in clusters in which genomic organization directly reflects domains of expression along the anterior-posterior body axis: genes at one end of the cluster pattern the anterior end of the embryo, genes at the opposite end pattern the posterior end. This is referred to as spatial collinearity; in vertebrates, which show very tight clusters, coordinated expression also has a temporal aspect: the genes are turned on successively from one end to the other, reflecting enhancer sharing and common regulatory mechanisms.

The POU class of homeodomain proteins was first identified in the mammalian transcription factors **P**it-1, **O**ct-1 and Oct-2 and in the nematode **U**nc-86 factor (**POU**) which shared a novel domain N-terminal to the well known homeodomain (Ryan and Rosenfeld, 1997; Phillips and Luisi, 2000). The POU domain is thus bipartite, consisting of the N-terminal POU specific domain (POU_S / ~75aa) connected by a highly variable linker to the POU homeodomain (POU_H / 60aa), and both sub-domains are involved in recognizing and binding of target DNA sequences (Ryan and Rosenfeld, 1997; Banerjee-Basu and Baxevanis, 2001; Phillips and Luisi, 2000). POU domain proteins also interact with transcription factors of the same or unrelated families like the SOX proteins, thereby controlling gene expression from various DNA enhancers in a combinatorial fashion which is thought to provide a higher level of functional diversity (Remenyi et al., 2004). POU genes play important roles in many developmental systems, notably in the nervous system (Latchman, 1999; Zhou et al., 1996), and they have been detected in virtually all metazoan taxa, including Cnidaria and sponges (Seimiya et al., 1997; Larroux et al., 2006; Shah et al., 2000; Ryan et al., 2006). Based on sequence similarity of the POU domain six subclasses are commonly recognized (Ryan and Rosenfeld, 1997).

Phylogenetic analyses based on sequence identity suggest that the ANTP and PRD classes are closer related to each other than to the remaining classes, whereas, for example, LIM and POU are more diverged (Galliot et al., 1999; Banerjee-Basu and Baxevanis, 2001; Holland and Takahashi, 2005). A possible scenario has been proposed (Galliot et al., 1999) in which a LIM-like ancestor, possibly related to the *Dictyostelium* Wariai HD, gave rise to the ancestors of LIM and PRD classes (Galliot et al., 1999). Besides other discriminating characteristics, all LIM homeobox proteins share a glutamine residue at homeodomain position 50 (Q50) - like the basal members of the PRD class, whereas several independent transitions from glutamine to serine and lysine have occurred in more derived PRD proteins. Close to the base of PRD origin the ANTP class could have originated, and almost all of its members have retained a glutamine residue at position 50 (Galliot et al., 1999). The remaining classes may have originated even earlier in this scenario - and one of the defining characteristics of the POU class in particular is the unique possession of a cysteine residue at position 50 (Banerjee-Basu and Baxevanis, 2001) which seems to be essential for the discrimination of target sequences (Stepchenko et al., 1997). Hence, the founding event of the POU class seems to be correlated with the transition of a Q50 to a C50 residue in an ancestral homeodomain and the fusion to a POU specific domain.

Linkage patterns of genes can aid to unravel their evolutionary history. If similar genes are linked in a genome, co-evolution or common origin by duplication seems likely. This is evident for the Hox genes in Bilateria which are clustered and thought to have originated by successive duplications from a single ancestor gene (Garcia-Fernandez, 2005a; Garcia-Fernandez, 2005b). However, if there were no functional constrains for maintaining of linkage - like the coordinated expression of Hox cluster genes - then trans duplication events, gene loss and chromosomal breakage are likely to occur which might disperse genes that once had coevolved. Hence, ancient linkage of genes is not easy to assess in extant species, especially for genes that are less closely related and have diverged for a longer time - like Hox and NK genes, or even Hox and POU genes. Nevertheless, comparative mapping of related genes and chromosomal regions across many taxa have aided to

resolve several questions about homeobox gene evolution. For example, comparative gene mapping has revealed that remnants of NK clusters exist in metazoan species and that homeobox genes closely related to NK genes are often either linked to other NK genes (like Msx) or to Hox genes (like Dlx) in chordates (Coulier et al., 2000; Pollard and Holland, 2000; Luke et al., 2003; Castro and Holland, 2003). Gene mapping has thus shown that most likely the ANTP class genes have emerged and coevolved once on a single ancestral genomic region - the proposed ANTP mega-array - before they were separated. Similarly, gene linkage and analyses of paralogous chromosomal regions in human and mouse also suggest that most of the remaining metazoan homeobox classes possibly derived from a single common ancestor region - which was consequently named the contraHox superparalogon (Popovici et al., 2001).

The question is whether there has been a link between the two proposed super-paralogons of metazoan homeobox genes - one containing most of the ANTP class genes, the other containing the PRD, LIM, POU, SIX, CUT, TALE and prospero classes? From analyses of sequence relatedness we might expect such a link because PRD and ANTP classes are closely related. However, the founding event of the ANTP class might as well have been a trans duplication that translocated a putative founder gene to a different chromosomal location, thereby separating the evolution of the ANTP class from that of the remaining classes.

We have recently reported about the genomic organization of Hox-like genes in the hydrozoan *Eleutheria dichotoma* and the anthozoan *Nematostella vectensis* (Kamm et al., 2006) and briefly mentioned that the *Eleutheria* anterior Hox-like gene Cnox-5 (Kuhn et al., 1996) is linked to a putative POU gene. A closer look at this linkage might contribute to our understanding of metazoan homeobox diversification. Hence, we further investigated the linked POU gene by means of 3' and 5' RACE, phylogenetic sequence analyses and by comparing linkage patterns of putative orthologs in other metazoan genomes.

Materials and Methods

Molecular Methods

A fosmid containing the Hox-like gene Cnox-5 (Kuhn et al., 1996) and the putative POU gene has been initially described in (Kamm et al., 2006) (accession DQ451873). To obtain the full length cDNA sequence for the POU gene whole RNA was isolated from ~70 *Eleutheria* medusae using the Trizol Kit (Invitrogen) and 3' and 5' RACE was conducted using the GeneRacer Kit along with SuperScript II Reverse Transcriptase (Invitrogen). Gene specific primers were constructed from conserved regions lying in open reading frames of the genomic sequence and used together with the adapter primers of the GeneRacer Kit for PCR amplification of the 3' and 5' ends. PCR fragments were separated by gel electrophoresis and the resulting bands purified and cloned into the pGEM-T vector (Promega). Plasmids were sequenced on a MegaBACE 1000 system (Amersham) and obtained sequences assembled with SeqMan (Lasergene). Sequences have been deposited to GenBank (accession numbers EU072102-EU072103).

Phylogenetic analyses and linkage patterns

Phylogenetic analyses were conducted using the POU specific domain together with the 60 amino acids of the POU homeodomain or the POU homeodomain alone (to allow the inclusion of LIM homeodomains as outgroup). Amino acids were aligned using the MEGA3.1 (Kumar et al., 2004) software package and corrected by eye. In the case of the POU_S domain, the conserved F residue at position 16 was used as a reference point (Munoz-Marmol et al., 1998) while an E/D residue at position 79 marked the end of the sub-domain (note that POU class 4 proteins contain a conserved insertion of the motif PGV after position 43 which is treated as an alignment gap in the analyses). The variable linker between POU_S and POU_H was removed from the alignment for the analysis of both domains. Alignments were tested for an appropriate model of amino acid substitution with ProtTest v1.2.6 (Abascal et al., 2005) and Bayesian analyses (MrBayes v3.1.1) were conducted to infer phylogenetic relationships (Ronquist and

Huelsenbeck, 2003). Four chains and two independent simultaneous runs over 1,000,000 generations were used, sampling every 50th tree. The likelihoods of the generations were examined to estimate the beginning of the stationary phase and trees after the first 300,000 generations were used to create a consensus tree. We used the JTT model of amino acid substitution, assuming the presence of invariant sites and using a gamma distribution approximated by four different rate categories to model rate variation between sites. Maximum Likelihood analyses were conducted using PHYML v2.4.4 (Guindon and Gascuel, 2003) with the JTT model of amino acid substitution and a gamma distribution approximated by four rate categories. A consensus tree was calculated from 1,000 bootstrap resamplings. Tress were examined using the TreeExplorer implemented in MEGA and rooted on midpoint in the case of the full POU domain. Accession numbers for protein sequences from NCBI databases are as follows: AmPOUIII, ABD97868; BfPOUIII, AAL85498; CiPOU-IV, NP_001027972; DmCG11641, AAF59100; DmI-POU, CAA41342; Dmnubbin, NP_476659; DmVVL, AAO39521; DrPOU12, AAH58318; DrPOU1F1,NP_998016;DrPOU47,NP_571235; DrPOU5F1, NP_571187; DrPOUC, NP_571188; GgPOU1F1, NP_989650; GgPOU2F1, NP_990803; HsLHX1, AAH20470; HsLHX2, P50458; HsLHX3, Q9UBR4; HsPOU1F1, NP_000297; HsPOU2F2, AAH06101; HsPOU3F1, NP_002690; HsPOU3F2, AAH51699; HsPOU4F1, NP_006228; HsPOU4F3, AAC06203; HsPOU5F1, NP_002692; HsPOU6F1, AAH51326; HsPOU6F2, EAL23992; NvLHX15, ABG67841; SpPOU6, XP_785865; XlPOU3F2, AAH41298; Xlxoct91, AAA49999.

Linkage of POU class 6 genes was assessed by blasting known orthologs against existing genome releases at NCBI.

Results and Discussion

3' and 5' RACE revealed a fusion transcript

Unexpectedly, 3' and 5' RACE resulted in a ~2600bp cDNA containing the POU gene with the POU_S and POU_H domains, together with the coding sequence for a complete Phosphopantothenoylcysteine-Synthetase (PPCS) 5' of the POU domain, all in a single continuous open reading frame (Figure 1A). A putative start co-

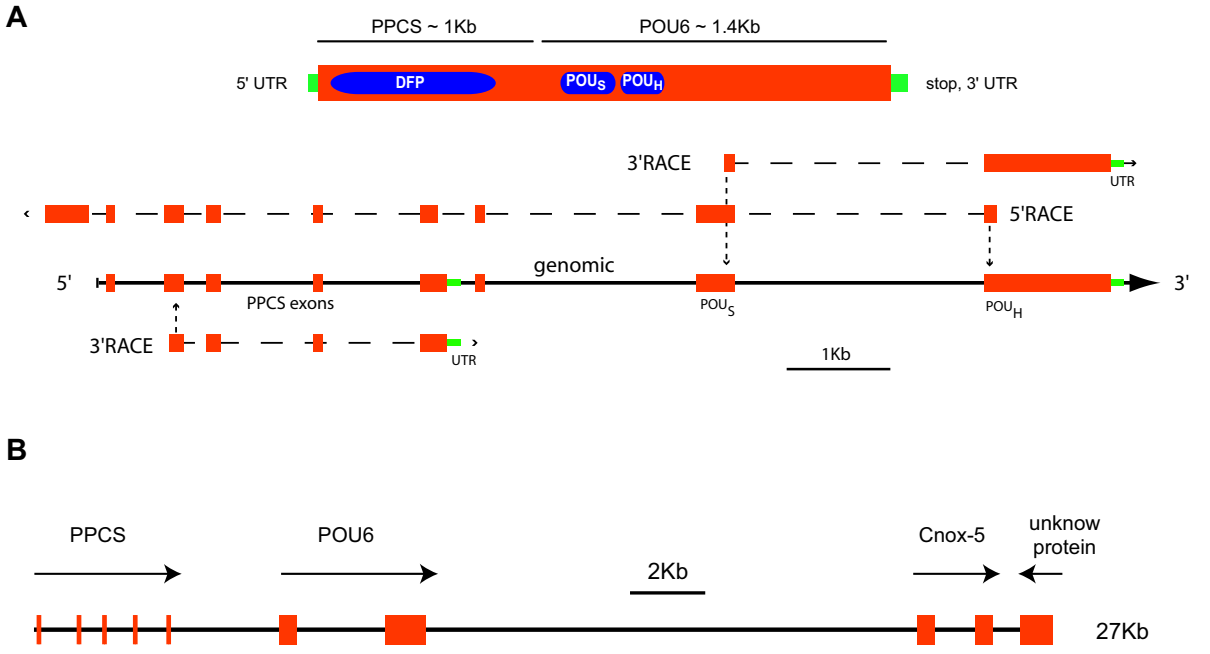


Figure 1. A cDNA genomic alignment - RACE experiments revealed a fusion transcript consisting of the *Eleutheria* POU6 gene and a PPCS gene. 3' RACE starting from a PPCS exon revealed an alternative transcript. **B** Genomic organization of the *Eleutheria* POU6 and Cnox-5 genes.

don is present at position 22-24 and a stop at 2476-78 after which the 3'UTR follows. Alternative transcripts could not be revealed by 5' RACE. The PPCS exons also lie 5' of the POU exons on the genomic fosmid (Figure 1A & B) and the sequence of the transcript is completely alignable with the genomic sequence, hence it is virtually impossible that the fusion transcript could have resulted from accidental ligation of two transcripts during ligation of the GeneRacer RNA oligo. Conversely, 3' RACE with a gene specific primer binding to one of the exons of PPCS resulted in a transcript without the POU exons. Additionally, the last PPCS exon was some 60bp longer in this transcript, after which the stop codon and the 3'UTR followed. The alignment of both transcripts with the genomic sequence revealed that all exons bear acceptor and donor splice sites. This suggests that PPCS itself is differentially transcribed and alternatively spliced.

Polycistronic transcripts with unrelated genes and/or hybrid mRNAs have been revealed in several cases in the phylum Nematoda (von Mering and Bork, 2002) or, in some cases, even in mammals (Mayer and Floeter-Winter, 2005). Moreover, in *Hydra vulgaris* a proportion of mRNAs seems to receive leader sequences which in nematodes are known to be essential for the procession of polycistronic transcripts via a spliced-leader RNA trans-splicing mechanism (Stover and Steele, 2001). However, nothing has been known about an operon-like organization of particular *Hydra* genes. The *Eleutheria* case reported here, i.e. a seemingly completely spliced fusion transcript, seems to be without precedents. Its functional significance remains unknown.

Phylogenetic analyses of POU homeobox genes

The full POU_S and POU_H domains of the *Eleutheria* POU gene were determined by RACE and phylogenetic analyses were conducted with members of all six bilaterian POU classes (Figure 2A & B). With the exception of an *Acropora millepora* POU class 3 gene, other cnidarian POU genes have not been included in the analyses because no complete POU domains have been isolated yet (Shah et al., 2000;

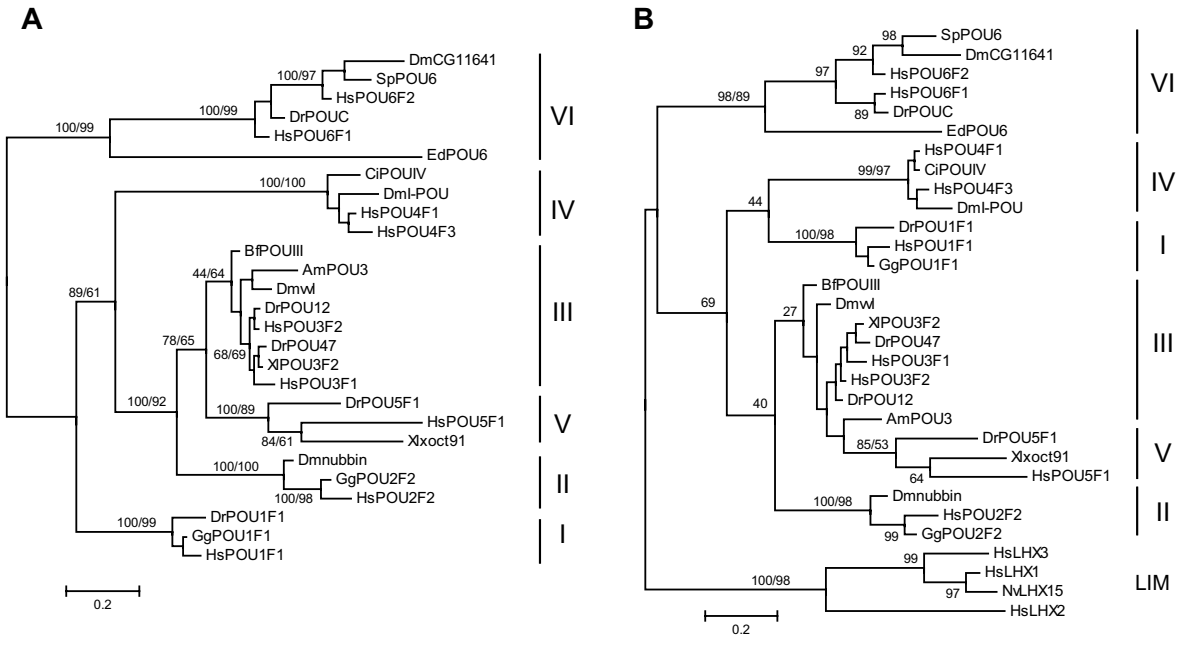


Figure 2. A Bayesian phylogenetic tree of complete POU domains, rooted on midpoint. Numbers at branches reflect Bayesian posterior probabilities on the left and percentages of 1,000 bootstrap resamplings of the ML analysis on the right. **B** Bayesian phylogenetic tree with POU homeodomains rooted with LIM homeodomains. Numbers at branches reflect Bayesian posterior probabilities. Percentages of 1,000 bootstrap resamplings from an ML analysis are given on the right at critical branches (Am *Acropora millepora*, Bf *Branchiostoma floridae*, Ci *Ciona intestinalis*, Dm *Drosophila melanogaster*, Dr *Danio rerio*, Ed *Eleutheria dichotoma*, Gg *Gallus gallus*, Hs *Homo sapiens*, Nv *Nematostella vectensis*, Sp *Strongylocentrotus purpuratus,* Xl *Xenopus laevis*).

Ryan et al., 2006). Analyses with both, the complete POU domain (Figure 2A) as well as with the homeodomain alone (Figure 2B), supported the same POU classes, although the latter revealed some ambiguity for the relation between POU class 3 and class 5 proteins which is also visible in the relatively low support for the class 3 clade in the analyses of the complete POU domain.

Phylogenetic analyses show that the *Eleutheria* POU protein segregates with the POU 6 class. Compared to representatives from Bilateria it appears to be relatively derived - whereas the other cnidarian sequence, the *Acropora* POU3 protein, seems to be less derived from the POU consensus. Initial analyses (not shown) with the POU homeodomain alone showed that the *Nematostella* POU 1, 3 and 4 class genes (Ryan et al., 2006) also seem to be less derived from the POU consensus. However, the homeodomain of a *Nematostella* POU class 6 gene (Ryan et al., 2006) as well as a putative POU6 gene from *Condylactis* (with only partial POU_S and partial POU_H domain) (Shah et al., 2000) show similar deviations. For example, at homeodomain position 44 all cnidarian class 6 genes show a threonine residue, whereas all other POU genes (including bilaterian class 6) have a valine residue at this position. Hence this deviation seems to be specific for cnidarian POU class 6 genes, indicating that they are all more diverged from the POU consensus than the remaining classes present in Cnidaria.

Phylogenetic analyses with the complete POU domain as well as with the homeodomain only suggest that POU class 6 genes are the sister clade of the remaining POU classes (Figure 2A & B). This has to be taken cautiously, however, because the outcome of the different POU classes might be strongly affected by the outgroup, or, in the case of midpoint rooting, by the composition of the dataset. Moreover, in the ML analyses of the POU homeodomain we observed some ambiguity between the ML bootstrap consensus tree and the bootstrapped ML tree: while the first also suggests POU class 6 as the sister clade to the remaining classes, the latter suggests POU class 4 instead - and it cannot be excluded that the three amino acids longer POU_S domain of POU class 4 proteins (see above) represents the ancestral condition.

Linkage patterns of POU class 6 genes

The *Eleutheria* POU class 6 gene is tightly linked to the anterior Hox-like gene Cnox-5 (Kamm et al., 2006) (Figure 1B) although POU and ANTP class genes are not closely related (Galliot et al., 1999; Banerjee-Basu and Baxevanis, 2001). If linkage reflects relatedness, we would expect other ANTP class members to be linked closely to an anterior Hox-like gene. However, it is possible that a hypothetical ancestral homeobox gene array, containing ANTP gene members in addition to other more distantly related homeoboxes, had diverged to such an extent that only the linkage of an anterior Hox-like and a POU class 6 gene were retained. It is also possible that the observed linkage in *Eleutheria dichotoma*, though extremely tight, is the result of genome rearrangements specific to this lineage, and thus mere chance.

In order to discriminate between the two possibilities we examined current assemblies of genomes deposited at NCBI. The results clearly favour the possibility that the linkage of *Eleutheria* Cnox-5 and POU6 is the remnant of ancient linkage of homeobox genes from distantly related classes. At least for vertebrates there is strong evidence that these also have retained linkage of POU class 6 genes to their Hox clusters (Figure 3): In human POU6F1 is linked to the HoxC cluster on chromosome 12 (12q13.13 / 12q13.3) and POU6F2 to the HoxA cluster on chromosome 7 (7p13-p14 / 7p14-p15). In the case of the HoxC cluster this makes a distance of app. 2.7Mb, in the latter case app. 12Mb. In both cases, the distances are either equal or significantly less than the distances of Mox orthologs to the HoxA and B clusters (12 and 5Mb in human, respectively) - and Mox orthologs are commonly associated with Hox cluster evolution in chordates (Minguillon and Garcia-Fernandez, 2003). Also in mouse the POU6F1 gene is only 2.4Mb apart from the HoxC cluster on chromosome 15; the mouse POU6F2 gene, however, is not linked to the HoxA cluster but this could be due to mouse specific genome rearrangements (Nadeau and Sankoff, 1998). In the unfinished

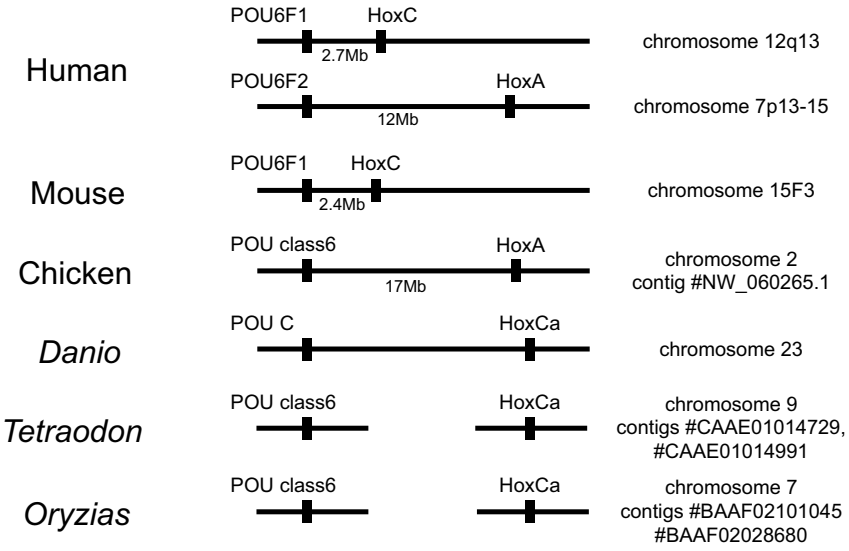


Figure 3. Linkage of POU class 6 genes to Hox genes in vertebrates. Distances are given where possible.

chicken genome we identified a POU class 6 ortholog on the same contig (#NW_060265.1) that also harbours the HoxA cluster, in a similar distance of 17Mb. In teleost fish, the *Danio rerio* class 6 ortholog POUC has been mapped to chromosome 23 where the HoxCa cluster is located (Woods et al., 2000); the same holds true for current *Tetraodon nigroviridis* and *Oryzias latipes* genome assemblies: a POU class 6 gene and the HoxCa cluster are located on two contigs that both map to chromosome 9 in the case of *Tetraodon* (#CAAE01014729 and #CAAE01014991, respectively) or chromosome 7 in the case of *Oryzias* (#BAAF02101045 and #BAAF02028680, respectively).

Conclusions

Cnidaria seem to have split off the lineage leading to Bilateria before a mature Hox cluster, consisting of all canonical classes, developed (Kamm et al., 2006; Chourrout et al., 2006) - while they already possess the majority of the remaining ANTP class genes (Kamm and Schierwater, 2006; Ryan et al., 2006). But some of the Hox-like genes present in Cnidaria - the anterior Hox-like and posterior Hox/Cdx-like genes - likely share the same ancestor with bilaterian anterior and posterior Hox genes (Kamm et al., 2006). Hence, it appears that the linkage of POU class 6 genes to Hox or Hox-like genes has preceded the cnidarian bilaterian split and can thus be regarded as ancestral linkage of distantly related homeobox genes.

We cannot decide between the possibilities that the observed linkage is either the direct remnant of an ancient homeobox mega-array, or whether a founder gene of the ANTP class and a POU class 6 gene were transposed away from this region. However, the observation that POU class 6 genes could be the sister clade of the remaining POU classes, together with the fact that cnidarian POU class 6 orthologs seem to be more derived from the POU consensus than other cnidarian POU genes, supports the second alternative. It favours an early division of POU class 6 genes from the remaining POU classes, which is also consistent with the mapping of most other POU classes to the contra-Hox paralogon (Popovici et al., 2001). POU class 4 genes, however, map to the same chromosomes as ParaHox genes in human, though with distances of more than 50Mb in two out of three cases. The exception is POU4F3 which maps only 4Mb apart from CDX1 (Popovici et al., 2001; NCBI-MapViewer). It is tempting that this provides support for the paralogous relationship of the predecessors of Hox and ParaHox clusters (whether originated from a two or three gene condition, or by cis- or trans-duplication (Garcia-Fernandez, 2005a)). Indeed, we have no evidence that the ancestors of POU class 4

and class 6 genes could have originated by duplication from one predecessor along the same process that might have produced distinct Hox and ParaHox clusters, because both subclasses show no close relationship (in this context one should also note the unique PGV motif in the POU_S domain of class 4 genes). In either case, the observed linkage of POU class 6 genes to Hox or Hox-like genes in two distantly related phyla suggests that at least the founder genes of most metazoan homeobox classes were once linked on a single ancestral genomic region.

Acknowledgements

We thank Peter Holland, David J. Miller and Allen Collins for comments and discussions.

Literature Cited

Abascal F, Zardoya R, Posada D. 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21(9):2104-2105.

Banerjee-Basu S, Baxevanis AD. 2001. Molecular evolution of the homeodomain family of transcription factors. Nucleic Acids Res 29(15):3258-3269.

Burglin TR. 1997. Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals. Nucleic Acids Res 25(21):4173-4180.

Castro LF, Holland PW. 2003. Chromosomal mapping of ANTP class homeobox genes in amphioxus: piecing together ancestral genomes. Evol Dev 5(5):459-465.

Chourrout D, Delsuc F, Chourrout P, Edvardsen RB, Rentzsch F, Renfer E, Jensen MF, Zhu B, de Jong P, Steele RE, Technau U. 2006. Minimal ProtoHox cluster inferred from bilaterian and cnidarian Hox complements. Nature 442(7103):684-687.

Coulier F, Popovici C, Villet R, Birnbaum D. 2000. MetaHox gene clusters. J Exp Zool 288(4):345-351.

Ferrier DE, Minguillon C. 2003. Evolution of the Hox/ParaHox gene clusters. Int J Dev Biol 47(7-8):605-611.

Galliot B, de Vargas C, Miller D. 1999. Evolution of homeobox genes: Q50 Paired-like genes founded the Paired class. Dev Genes Evol 209(3):186-197.

Garcia-Fernandez J. 2005a. The genesis and evolution of homeobox gene clusters. Nat Rev Genet 6(12):881-892.

Garcia-Fernandez J. 2005b. Hox, ParaHox, ProtoHox: facts and guesses. Heredity 94(2):145-152.

Gehring WJ. 1985. The homeo box: a key to the understanding of development? Cell 40(1):3-5.

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52(5):696-704.

Holland PW, Takahashi T. 2005. The evolution of homeobox genes: Implications for the study of brain development. Brain Res Bull 66(4-6):484-490.

Kamm K, Schierwater B. 2006. Ancient complexity of the non-Hox ANTP gene complement in the anthozoan *Nematostella vectensis*. Implications for the evolution of the ANTP superclass. J Exp Zoolog B Mol Dev Evol 306(6):589-596.

Kamm K, Schierwater B, Jakob W, Dellaporta SL, Miller DJ. 2006. Axial patterning and diversification in the cnidaria predate the Hox system. Curr Biol 16(9):920-926.

Kuhn K, Streit B, Schierwater B. 1996. Homeobox genes in the cnidarian *Eleutheria dichotoma*: evolutionary implications for the origin of Antennapedia-class (HOM/Hox) genes. Mol Phylogenet Evol 6(1):30-38.

Kumar S, Tamura K, Nei M. 2004. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief Bioinform 5(2):150-163.

Larroux C, Fahey B, Liubicich D, Hinman VF, Gauthier M, Gongora M, Green K, Worheide G, Leys SP, Degnan BM. 2006. Developmental expression of transcription factor genes in a demosponge: insights into the origin of metazoan multicellularity. Evol Dev 8(2):150-173.

Latchman DS. 1999. POU family transcription factors in the nervous system. J Cell Physiol 179(2):126-133.

Luke GN, Castro LF, McLay K, Bird C, Coulson A, Holland PW. 2003. Dispersal of NK homeobox gene clusters in amphioxus and humans. Proc Natl Acad Sci U S A 100(9):5292-5295.

Mayer MG, Floeter-Winter LM. 2005. Pre-mRNA trans-splicing: from kinetoplastids to mammals, an easy language for life diversity. Mem Inst Os-

waldo Cruz 100(5):501-513.

Minguillon C, Garcia-Fernandez J. 2003. Genesis and evolution of the Evx and Mox genes and the extended Hox and ParaHox gene clusters. Genome Biol 4(2):R12.

Munoz-Marmol AM, Casali A, Miralles A, Bueno D, Bayascas JR, Romero R, Salo E. 1998. Characterization of platyhelminth POU domain genes: ubiquitous and specific anterior nerve cell expression of different epitopes of GtPOU-1. Mech Dev 76(1-2):127-140.

Nadeau JH, Sankoff D. 1998. Counting on comparative maps. Trends Genet 14(12):495-501.

NCBI-MapViewer. http://www.ncbi.nlm.nih.gov/mapview/.

Phillips K, Luisi B. 2000. The virtuoso of versatility: POU proteins that flex to fit. J Mol Biol 302(5):1023-1039.

Pollard SL, Holland PW. 2000. Evidence for 14 homeobox gene clusters in human genome ancestry. Curr Biol 10(17):1059-1062.

Popovici C, Leveugle M, Birnbaum D, Coulier F. 2001. Homeobox gene clusters and the human paralogy map. FEBS Lett 491(3):237-242.

Remenyi A, Scholer HR, Wilmanns M. 2004. Combinatorial control of gene expression. Nat Struct Mol Biol 11(9):812-815.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19(12):1572-1574.

Ryan AK, Rosenfeld MG. 1997. POU domain family values: flexibility, partnerships, and developmental codes. Genes Dev 11(10):1207-1225.

Ryan JF, Burton PM, Mazza ME, Kwong GK, Mullikin JC, Finnerty JR. 2006. The cnidarian-bilaterian ancestor possessed at least 56 homeoboxes. Evidence from the starlet sea anemone, *Nematostella vectensis*. Genome Biol 7(7):R64.

Seimiya M, Watanabe Y, Kurosawa Y. 1997. Identification of POU-class homeobox genes in a freshwater sponge and the specific expression of these genes during differentiation. Eur J Biochem 243(1-2):27-31.

Shah D, Aurora D, Lance R, Stuart GW. 2000. POU genes in metazoans: homologs in sea anemones, snails, and earthworms. DNA Seq 11(5):457-461.

Stepchenko AG, Luchina NN, Pankratova EV. 1997. Cysteine 50 of the POU H domain determines the range of targets recognized by POU proteins. Nucleic Acids Res 25(14):2847-2853.

Stover NA, Steele RE. 2001. Trans-spliced leader addition to mRNAs in a cnidarian. Proc Natl Acad Sci U S A 98(10):5693-5698.

von Mering C, Bork P. 2002. Teamed up for transcription. Nature 417(6891):797-798.

Woods IG, Kelly PD, Chu F, Ngo-Hazelett P, Yan YL, Huang H, Postlethwait JH, Talbot WS. 2000. A comparative map of the zebrafish genome. Genome Res 10(12):1903-1914.

Zhou H, Yoshioka T, Nathans J. 1996. Retina-derived POU-domain factor-1: a complex POU-domain gene implicated in the development of retinal ganglion and amacrine cells. J Neurosci 16(7):2261-2274.