Hox cluster genes and collinearities throughout the tree of animal life

STEPHEN J. GAUNT*
Department of Zoology, University of Cambridge, Cambridge, U.K.

ABSTRACT  The discovery of Hox gene clusters, first in *Drosophila* (a protostome) and then as homologues in vertebrates (deuterostomes), was a major step in our understanding of both developmental and evolutionary biology. Hox genes in both species perform the same overall function: that is, organization of the body along its head-tail axis. The conclusion is that the protostome-deuterostome ancestor, founder of 99% of all described animal species, must already have had this same basic Hox cluster, and that it probably used it in the same way to establish its body plan. A striking feature of Hox genes is the spatial collinearity rule: that order of the genes along the chromosome corresponds with the order of their expression domains along the embryo. For vertebrates, though not *Drosophila*, there is also the temporal collinearity rule: that order of genes along the chromosome corresponds with timing of Hox expressions in the embryo. Although Hox genes are clearly recognized in pre-bilaterians (Cnidaria), it is only in bilaterians that the characteristic clustered Hox arrangement and function is commonly found. Spatial collinearity in expression is conserved widely throughout Bilateria but temporal collinearity is so far limited to vertebrates, cephalochordates, and some arthropods and annelids. In addition to conserved use of Hox genes to pattern the head-tail axis, some animal groups, particularly lophotrochozoans, have extensively co-opted Hox genes, outside collinearity rules, to regulate development of novel structures. Satisfactory understanding of Hox cluster function requires better understanding of the bilaterian last common ancestor (Urbilateria). Xenacoelomorpha may provide useful living models of the ancestral bilaterian condition.

KEY WORDS: embryo, development, evolution, phylogeny

Introduction

Studies in *Drosophila* indicated that Hox genes are determinants of the body plan along the head-tail axis, and that they are clustered such that the order of the genes along the chromosome corresponds with the order of their expressions along the body (Lewis, 1978). Lewis named this correspondence ‘collinearity’ (Lewis, 1985), though it is now usually called ‘spatial collinearity’. Lewis proposed that the Hox genes are expressed in a series of partially overlapping domains along the head-tail axis and that each region along the body expresses a different combination of Hox genes. Subsequent molecular studies (Gehring, 1985, Harding et al., 1985): 1) confirmed Hox gene clustering and the validity of the collinearity rule, 2) showed that Hox genes are indeed commonly expressed in partially overlapping domains, and 3) revealed that all Hox genes contain a 180bp conserved DNA motif, the homeobox. The homeobox encodes the homeodomain, a DNA sequence-specific binding domain which enables the Hox protein to fulfil its role as a transcription factor. *Drosophila* homeobox sequences were then used as probes to isolate Hox genes from vertebrates. Remarkable findings were: 1) that vertebrate genes are also clustered and obey the collinearity rule in their expressions (Gaunt et al., 1988), 2) that the entire set of Hox genes in *Drosophila* is homologous with each of four Hox gene clusters in amniotes (Boncinelli et al., 1988, Duboule and Dolle, 1989, Graham et al., 1989), and 3) that the Hox genes perform similar roles in specification of body regions along the head-tail axis (Mallo et al., 2010). The most likely explanation is that the Hox cluster, its collinear expression, and its role in positional specification along the head-tail axis were already present in the last common ancestor of *Drosophila* and vertebrates. This

*Address correspondence to: Stephen J. Gaunt, Department of Zoology, University of Cambridge, Downing Street, Cambridge, U.K., CB2 3EJ.Tel: +44-1223-768917 Fax: +44-1223-336676. E-mail: sg397@cam.ac.uk - https://orcid.org/0000-0001-6038-2272

Submitted: 19 March, 2018; Accepted: 26 June, 2018.
In vertebrates, the body develops in an anterior to posterior temporal progression, with new structures developing from cells that emerge from a posterior growth zone. Correspondingly, the Hox genes of vertebrates are first expressed with this same temporal progression. This is known as ‘temporal collinearity’ because the timing of first expression corresponds with the ordering of the genes along the chromosome (Izpisua-Belmonte et al., 1991). There are alternate views on whether temporal collinearity dictates need for spatial collinearity, or whether spatial dictates temporal collinearity. Either way, this is an important difference from Drosophila where Hox genes all activate at the same time, without temporal collinearity.

**Hox genes envisaged as a bunch of keys**

A Hox gene and its protein can be thought of as a key that unlocks the developmental potential of a distinct body zone along the head-tail axis (Fig. 2). Like a key, the protein homeodomain carries specificity, enabling it to bind to and activate an appropriate variety of 'downstream' genes which, together, specify development of an anatomical structure (e.g. a leg in the case of the Antp gene, or a ribbed vertebra in the case of Hoxc6). Like keys on a keyring, the Hox genes of both Drosophila and vertebrates are clustered together. To complete the analogy, the keys are arranged on the keyring such that they are collinear with the order of the body zones whose developmental programs they unlock (Fig. 2).

The bunch of keys is substantially the same in both Drosophila and vertebrates. Supporting this, mouse Hoxb6 and Hoxa5 genes can mimic, respectively, Antp and Scr functions when expressed in Drosophila embryos, and this 'unlocking' of Antp- and Scr-directed programs is due to their activation of the appropriate downstream genes (Malicki et al., 1990, Zhao et al., 1993). Similarly, human Hoxd4 can substitute for a normal regulatory function of Drosophila Dfd (McGinnis et al., 1990). Differences in anatomy between Drosophila and vertebrates are, therefore, largely due to differences in the variety and function of downstream genes activated by the Hox keys. A structure can change its morphology over evolutionary time by mutations which change the mix of downstream genes activated by its regulating Hox protein. This is by gain, loss, or modification of the Hox protein binding motifs within regulatory regions of downstream genes (Gaunt and Paul, 2012). A structure can change its position over evolutionary time by a change in the location of the Hox gene expression domain (Gaunt, 1994, Martin et al., 2016).

Since both Drosophila and mouse use a similar and related set of Hox keys we can infer that this was probably also used by their last common ancestor (Fig. 1). The question then arises as to whether there is a universal bunch of keys that regulates development in all animals. This review summarizes organization of Hox clusters and patterns of expression (spatial and temporal collinearities) as they have been found throughout the tree of animal life.

**The tree of animal life**

The modern tree of animal life (the ‘new phylogeny’) (Figs. 3-5), constructed in 1997 (Aguinaldo et al., 1997) and thereafter, relies upon comparison of DNA sequences. Random mutations cause DNA sequence to change gradually over time, so species that share similar sequences are deemed to be more closely related than are species with more diverged sequences. Some shared
anatomical features, considered below, are consistent with the new phylogeny. The two great groups of bilaterally symmetrical animals (Bilateria) are the protostomes and the deuterostomes. These groups account for 99% of all described animal species (DuBuc et al., 2012). The protostome-deuterostome last common ancestor (P-DLCA) is estimated to have lived between about 550 and 650 million years ago (Cunningham et al., 2017).

Deuterostomes, though related in DNA sequence, form a morphologically diverse collection of animal phyla (Chordata, Hemichordata and Echinodermata) (Fig. 3). However, unlike protostomes, they share a distinct morphology at the eight-cell embryo stage: four of the cells sit directly on top of the other four cells following a process known as radial cleavage (Holland, 2011). Among protostomes (Fig 4), most lophotrochozoan embryos at the eight-cell stage have four upper cells rotated relative to the four lower cells. This is called a spiral cleavage pattern, and the group as a whole is sometimes called Spiralia. Early embryos of Ecdysozoa are not characterized by either spiral or radial cleavage.

Deuterostomes are also distinguished from protostomes in how they form the mouth (Holland, 2011). During gastrulation, invagination into the early deuterostome embryo produces the anus, while the mouth must form secondarily (deutero: second; stome: mouth). For protostomes, in contrast, early invagination into the embryo produces the mouth (proto: first) and sometimes also the anus. Not all animals conform with the above definitions (Martin-Duran et al., 2016) but the terms deuterostome and protostome have been retained for reasons of familiarity and convenience.

Within deuterostome phylogeny (Fig. 3) (Tassia et al., 2016), the chordates possess, at least during part of their life cycle, pharyngeal clefts, a hollow dorsal nerve cord, a notochord, and a post-anal tail. Hemichordates are worm-like marine invertebrates which have a tripartite division of the body and some chordate like features: pharyngeal gill clefts, dorsal nerve cord and a notochord-like structure. Echinoderms are bilaterally symmetrical at the larval stage but this is lost at metamorphosis when they develop radial (usually five-fold) symmetry.

Within protostome phylogeny (Fig. 4), the ecdysozoans are distinct from lophotrochozoans in their common property of growth by moulting, a feature that may have arisen only once during evolution (Aguinaldo et al., 1997).

Xenacoelomorpha (Fig. 3) are a grouping of at least three phyla, Acoelomorpha, Nemertodermatida and Xenoturbellida, which are bilaterally symmetrical marine flatworms. They lack some features common to most other bilaterians such as an anus, nephridia, and a circulatory system. Even after intensive DNA sequence analysis, there remains uncertainty about whether these phyla are living basal bilaterians, forming a sister group to all other bilaterians (‘Nephrozoa’) (Bourlat and Hejnol, 2009, Cannon et al., 2016, Rouse et al., 2016), or whether they are degraded forms of deuterostomes (Philippe et al., 2011) (Fig 3).

Among pre-bilaterian phyla (Fig. 5), cnidarians are the most complex with a two-layered body wall, radial symmetry, and being the only members to possess Hox genes. These features, together with DNA analyses, have commonly led to their placement as the sister group to Bilateria (Cannon et al., 2016). However, branching orders among pre-bilaterians and the earliest bilaterian remain uncertain (Fig. 5).

**Fig. 3. Hox gene arrangements in deuterostomes.** Genes are numbered and coloured according to their homology groups: those that share the same numbers and colours are orthologues, most recently related by descent. Arrows indicate directions of transcription. Species shown are according to their homology groups: those that share the same numbers and colours are orthologues, most recently related by descent. Arrows indicate directions of transcription. Species shown are
Hox cluster structures and expressions are now described for representative members in the tree of animal life. Prominence is given to species where there is both gene mapping and expression data.

**Hox cluster genes in deuterostome bilaterians**

The vertebrates are the only chordates that show Hox cluster duplications (giving four paralogous clusters in amniotes) (Figs. 1, 3). Vertebrate clusters are largely intact and the genes are expressed with spatial (Fig. 6A) and temporal collinearities in both ectoderm- and mesoderm-derived tissues (Gaunt , 1988, Izpisua-Belmonte et al., 1991).

*Oikopleura dioica* is a urochordate (tunicate) of the larvacean type, which means that it retains a larval morphology throughout life. It is tadpole-like, 1-8 mm. long with 9 Hox genes. These do not include the usual central Hox genes but a full vertebrate-like set of posterior genes is present (Fig. 3). The genes display expression boundaries which are ‘spatially collinear’ as expected from their corresponding genes in vertebrates, but the genes are now dispersed with no remnant of the ancestral clustering (Seo et al., 2004). Duboule describes this as ‘trans-collinearity’ (Duboule, 2007). The urochordate *Ciona intestinalis* (an ascidian) develops to an adult form resembling a leather bottle (a sea squirt). Its Hox genes are only partially dispersed (Pascual-Anaya et al., 2013, Sasakura and Hozumi, 2018). *Ciona* shows residual spatial collinearity in the developing larval nervous system and in the juvenile gut during metamorphosis (Ikuta et al., 2004, Nakayama et al., 2016). Knock-down of *Ciona* Hox genes shows that they play only minor roles in larval development but major roles during subsequent metamorphosis (Ikuta et al., 2010, Sasakura and Hozumi, 2018). Neither of the above urochordate species displays obvious temporal collinearity in Hox gene expression, and expressions are reported in ectoderm, mesoderm and endodermal tissues (Ikuta et al., 2004, Seo et al., 2004).

The amphioxus *Branchiostoma floridae* is a cephalochordate. Its body is translucent, fish-like without paired fins, and about 5 cm. in length. It has a single cluster of 15 Hox genes (Fig. 3) (Garcia-Chávez et al., 2012). These include the full set of Hox genes found in each vertebrate cluster. In general, amphioxus Hox genes are seen to be expressed in the embryo with spatial and temporal collinearity, though *Hox6* may be expressed anteriorly to *Hox*4 in the European amphioxus (Pascual-Anaya et al., 2012). Expression is reported in both ectodermal and mesodermal tissues (Pascual-Anaya et al., 2012).

*Saccoglossus kowalevskii* is an acorn worm, the best known type of hemichordate. Acorn worms are usually a few cm. long, are worm shaped with an anterior proboscis, and they live in burrows where they filter food particles from sea water passing through their pharyngeal slits. The Hox cluster retains much of the ancestral, clustered arrangement (Fig. 3) (Freeman et al., 2012). The Hox genes are generally expressed during development with spatial but not temporal collinearity, and in ectodermal rather than mesodermal tissues (Aronowicz and Lowe, 2006). *S. kowalevskii* is a direct-developing hemichordate, which means that it does not develop via a larval stage (Gonzalez et al., 2017). Indirect-developing hemichordates, such as *Schizocardium californicum*, hatch to a free swimming larval stage. Metamorphosis proceeds by addition and development of a more posterior trunk region. Larvae without trunks have been described as ‘swimming heads’, and the trunk develops later under the influence of Hox genes, expressed in ectodermal tissues with spatial but no obvious temporal collinearity (Gonzalez et al., 2017). Hox genes are not expressed in early larvae. Indirect development, with a prolonged larval stage, has been regarded as a more primitive mode of development which has independently transformed to direct development in multiple animal groups (Peterson et al., 1997).

*Strongylocentrotus purpuratus*, the purple sea urchin, is an indirect-developing echinoderm. Typical of sea urchins the adult is globular with a rigid and spiny calcareous skeleton. The sea urchin Hox cluster is characterized by re-organization from the ancestral arrangement (Fig. 3). Only 2 of 11 Hox genes are clearly expressed during formation of the free swimming bilaterian larva, and even these are probably not required for regional embryonic specification (Arenas-Mena et al., 1998). However, the posterior group of Hox genes (*Hox7 to Hox13*) displays spatial collinearity in expression in the mesoderm-derived posterior coeloms during the establishment of the adult five-fold radially symmetrical body plan (Arenas-Mena et al., 2000, Aronowicz and Lowe, 2006). There is no clear temporal collinearity in Hox gene expression and only limited expression, without spatial collinearity, in ectodermal tissues (Arenas-Mena et al., 2000, Arenas-Mena et al., 1998).

Apart from the sea urchins, echinoderms show a variety of other adult forms. For example, starfish have five (usually) stiff arms; sea cucumbers lie on their side so that in addition to five-fold radial symmetry they appear to have bilateral symmetry with anterior and posterior ends, and sea lilies have feathery tentacles around their mouth and an underside attached to the substratum by a stalk. The extensive Hox cluster re-arrangement shown by the sea urchin (Fig 3), which includes loss of *Hox*4, is likely also present in the sea cucumber *Apostichopus japonicus* (Byrne et al., 2016). However, the more distantly related starfish *Acanthaster planci* has an intact cluster (Baughman et al., 2014, Byrne et al., 2016) disproving an earlier hypothesis that five-fold symmetry is caused by the Hox gene rearrangement. Hence, the ancestral echinoderm Hox cluster was likely intact. Although their genomic Hox arrangements are uncertain, a sea cucumber *A. japonicus* (Kikuchi et al., 2015) and a sea lily *Metacrinus rotundus* (Hara et al., 2006) both show apparent spatial but not temporal collinearity in Hox gene expressions in their bilaterally symmetrical larvae (David and Mooi, 2014). In both of these species, as in sea urchin, collinearly-expressing structures include mesoderm-derived posterior coeloms.

**Hox cluster genes in protostome bilaterians**

In the arthropod *Drosophila melanogaster* the ancestral Hox cluster has become split into two (Fig. 4). However, the position of this split varies between different *Drosophila* species without obvious difference in body plan, suggesting that cluster integrity is not essential for function (Negre and Ruiz, 2007). The genes show spatial but not temporal collinearity. They regulate both ectodermal (Lewis, 1978) and mesodermal development (Greig and Akam, 1993, Michelson, 1994). Amongst other arthropods, *Tribolium castaneum*, the red flour beetle, has all Hox genes in an intact cluster but the cluster can be split without any adverse effects upon Hox gene function (Shippy et al., 2008). *Parhyale hawaiensis*, a shrimp-like crustacean, shows both spatial (Fig. 6B) and temporal collinearities (Serano et al., 2016). In the chelicerate
group of arthropods, a spider (Cupiennius salei) and a scorpion (Centruroides sculpturatus) show spatial collinearity in expression, and also extensive duplication of Hox cluster genes (Schwager et al., 2007, Sharma et al., 2014). The duplication events likely occurred independently (Kenny et al., 2016). In butterflies and moths, the zen (Hox3) gene has undergone tandem duplications to form a linear array of Hox-derived (Shx) genes located between Hox3 and 4 which, in at least one species, are expressed without collinearity in extra-embryonic tissues (Ferguson et al., 2014).

Outside Arthropoda, but within Ecdysozoa, are two smaller phyla Onychophora and Tardigrada. Members of both have stubby limbs and soft cuticles. The integrity of their Hox clusters is not yet established but both display apparent spatial collinearity in Hox gene expression (Fig. 6C) (Janssen et al., 2014, Smith et al., 2016). While onychophorans possess many leg bearing segments of similar structure, tardigrades possess only four pairs of legs due to apparent loss from the ancestral condition of more posterior leg-bearing segments. The Hox genes that specified the lost segments (Antp, Ubx, abdA) are now absent from the tardigrade genome (Smith et al., 2016).

The Nematoda are unsegmented worms which are round in cross section (roundworms). Caenorhabditis elegans is a transparent nematode about 1mm. in length and free-living in soil. Like all nematodes it lacks respiratory and circulatory systems. It is the nematode most studied with respect to Hox genes. It has only six remaining Hox genes in a widely spread cluster (Fig. 4) and only three are required during embryogenesis (Van Auken et al., 2000). Expression of the Hox genes is specified more by cell lineage than A-P position but there remains at least partial spatial collinearity and function in anteroposterior patterning (Van Auken et al., 2000). This is in spite of the fact that the most anteriorly expressed Hox1 gene has now transposed to be surrounded by more posteriorly expressed genes (Tihanyi et al., 2010). C. elegans Hox genes regulate cell fate in some ectodermal and mesodermal tissues (Liu and Fire, 2000, Tihanyi et al., 2010). The reduced role of C. elegans Hox genes in head-tail patterning may reflect a shift from a regulative to a lineage-dependent, deterministic mode of development (Aboobaker and Blaxter, 2003, Duboule, 1992). Some nematodes such as Ascaris suum retain two additional genes from the ancestral Hox complement even though their lineage is similar to that of C. elegans (Aboobaker and Blaxter, 2003).

The Platyhelminthes are the flatworms, lacking a true coelom. The parasitic flatworm Schistosoma mansoni has a dispersed Hox gene set (Fig. 4) (Pierce et al., 2005). The free-living planarian Schmidtea mediterranea has 13 Hox genes which include representatives of all the ancestral genes (Currie et al., 2016). The genes are probably dispersed in the genome. At least 5 are expressed in axially restricted zones along the head-tail axis, mostly overlapping posteriorly, but with only limited evidence of spatial collinearity (Currie et al., 2016). Other Hox genes display tissue-specific, rather than axially-restricted, expressions (Currie et al., 2016). In the free-living planarian Dugesia japonica a posterior Hox gene is expressed anterior to a middle Hox gene, breaking spatial collinearity (Nogi and Watanabe, 2001). Hox expressions are found in both ectodermal and mesodermal tissues, and some genes are apparently co-opted to tissue-specific and radially arranged expression roles (Currie et al., 2016, Nogi and Watanabe, 2001).

The Annelida are the segmented worms. Capitella teleta is a marine segmented polychaete (many-bristled) worm that, like many other annelids, shows continued adult growth by addition of segments at a posterior growth zone. It has 11 Hox genes, at least 8 of which are grouped together in a large Hox cluster (Fig. 4) (Simakov et al., 2013). The arrangement of these genes is conserved relative to the ancestral configuration. Capitella Hox genes are expressed in developing larvae with much spatial (Fig. 6D) and temporal collinearities (Frobius et al., 2008). Larval stages of the polychaete Chaetopterus (Irvine and Martinlade, 2000, Peterson et al., 2000) also display both spatial and temporal collinearities.

![Fig. 4. Hox gene arrangements in protostomes.](image) Hox genes are presented as in Fig. 3. Posterior (AbdB-like) Hox genes are labelled 9-13 to indicate their homeobox homologies with the posterior genes of vertebrates. Labelling of genes 6 to 8 does not necessarily imply proven orthology with vertebrate genes: they could represent independent expansion of the central genes (Fig. 1). Gene arrangements are from the following sources. D. melanogaster and C. elegans (Lemons and McGinnis, 2006; Tihanyi et al., 2010); S. mansoni (Lemons and McGinnis, 2006); C. teleta and L. gigantea (Frobius et al., 2008; Simakov et al., 2013).
However, larvae of the polychaete Alitta (Nereis) virens (Bakalenko et al., 2013, Kulakova et al., 2007) and embryos of the leech Helobdella (Kourakis et al., 1997) have shown spatial but not temporal collinearities. Differences in temporal collinearities may be linked to whether or not the axis sub-divides into distinct morphological regions (tagmata). Capitella is tagmatized, with tagma boundaries that must align with Hox expression boundaries (Frombius et al., 2008). Alitta is non-tagmatized over most of its length, and Hox expression boundaries at later stages regress posteriorly along the segment series (Bakalenko et al., 2013). Most, but not all, of the expression described in these annelids is ectodermal rather than mesodermal.

The Mollusca members vary greatly in appearance. The aculiferans (worm-like molluscs and the eight-part-shelled chitons) form a separate group to the conchiferans (clam, limpet, snail, slug, squid, and octopus) (Fritsch et al., 2015). Lottia gigantea (a limpet) has an intact cluster of 11 Hox genes which are collinear with the ancestral cluster (Fig. 4) (Simakov et al., 2013). Acanthochitona crinita, a chiton, has 7 Hox genes expressed with spatial collinearity along the head-tail axis of the larva but not in molluscan-specific structures such as the shell or foot (Fritsch et al., 2015). Expression was noted in ecto-, endo- and mesodermal tissues. In contrast, conchiferans such as snails Gibbula varia and Haliothis asinina examined at larval stages, and embryos of the squid Euprymna scolopes have been found to show limited spatial collinearity in expression only within the nervous system, and there is secondary recruitment of Hox genes into novel structures without any evidence for spatial collinearity (Hinman et al., 2003, Lee et al., 2003, Samadi and Steiner, 2010). This secondary co-option may help to explain how conchiferans have acquired a diverse array of body structures and designs (Fritsch et al., 2015). A recent study indicates that spatial, though not temporal, collinearity may be more widespread amongst molluscs than previously thought (Wolleisen et al., 2018).

**Hox genes in Xenacoelomorpha**

The P-DLCA ancestor must, as we have seen in Fig. 1, have had a cluster of at least 7 Hox genes. This large complement probably indicates that it already had a complex body plan. It has been suggested that this was the first bilaterian (Urbilateria) (De Robertis, 2008), but we cannot be certain of this. The first bilaterian may have had fewer Hox genes, and a simpler body plan. It is suggested, with some controversy, that Xenacoelomorpha may provide a living model of early bilaterians (Bourlat and Hejnl, 2009, Cannon et al., 2016). Symasgitifera roscoffensis is a acel flatworm, up to 15mm long, and green in colour due to algae incorporated as a source of photosynthetic energy. It has 3 Hox genes which represent the anterior, middle and posterior groups of other bilaterians. The genes have lost the ancestral clustering, being now dispersed onto different chromosomes (Fig. 3), and they are expressed in nested domains which show spatial collinearity along the embryo axis (Moreno et al., 2009). Convolutriloba longifisssura, an acel with a similar set of 3 Hox genes, shows spatial but not temporal collinearity (Hejnl and Martindale, 2009). Xenoturbella bocki, a xenoturbellid, has 5 Hox genes which include anterior, middle and posterior groups (Fritsch et al., 2008).

**Hox genes in pre-bilaterians**

Most authors agree that cnidarians have Hox genes of the anterior (Hox1 and Hox2) classes, and they also recognize at least one gene of either the posterior or middle class (Chiori et al., 2009, DuBuc et al., 2012). Strict orthologies with Hox genes of bilaterians have, however, been questioned (Kamm et al., 2006). In the coral Acropora, all three of these Hox genes are linked in a single cluster, but in some other species of Cnidaria there is either no cluster or only clustering of the anterior genes (Fig. 5) (DuBuc et al., 2012). The number of Hox genes often varies between different cnidian species, and the particular genes present may also vary. The cnidarian/bilaterian ancestor probably possessed a cluster containing at least one of each of the following Hox gene types: group 1, group 2, and a middle or posterior gene (DuBuc et al., 2012).

The sea anemone Nematostella vectensis (Fig. 5) has a Hox1 gene that is expressed orally, while a middle/posterior gene is expressed aborally. This, together with functional studies, led to hypotheses that 1) the oral-aboral axis of a cnidian is homologous with the head-tail axis of bilaterians, and 2) the Hox code was already established in the cnidian/bilaterian common ancestor (DuBuc et al., 2018, Finnerty et al., 2004). These proposals are not supported by Hox expression analyses on several other cnidian species which have shown no consistent evidence for collinearity, no consistency in expression patterns of orthologous Hox genes in different species, and therefore no evidence for a common cnidian Hox code (Chiori et al., 2009, Kamm et al., 2006, Reddy et al., 2015). Although Hox gene duplication had already likely occurred in the cnidian/bilaterian ancestor it remains possible that spatial collinearity in Hox gene expression first arose, or at least only flourished, in bilaterians, where it may have been a crucial factor in development of complexity along the head-tail axis.

Among pre-bilaterians, Hox genes are found only in Cnidaria.

---

**Fig. 5. Hox gene arrangements in pre-bilaterian animals.** Genes are presented as in Figs. 3, 4. N. vectensis genes are as described earlier (DuBuc et al., 2012).
Putative paraHox genes are, however, reported in Porifera (some calcisponges) (Fortunato et al., 2014) and in Placozoa (Ferrier, 2016, Mendivil Ramos et al., 2012), though not in Ctenophora. Since paraHox genes are thought to have originated along with Hox genes by duplication of an ancestral protoHox cluster (Brooke et al., 1998) Porifera and Placozoa may have had a Hox cluster ancestrally, and then this was lost secondarily (Ferrier, 2016). Further work is needed to evaluate this possibility.

**The compactness of Hox gene clusters**

Figs. 3-5 show whether or not Hox genes are clustered, but they do not accurately represent the compactness of clustering. The vertebrate clusters are the most compact. This is seen in an overall size for each cluster of 100-170kb, and also in the absence of any interspersed non-Hox genes (Duboule, 2007, Pace et al., 2016). The amphioxus cluster has an overall size of at least 450kb (Duboule, 2007). The purple sea urchin cluster extends over more than 500kb (Arenas-Mena et al., 2000). The two parts of the *Drosophila* cluster extend over 712kb (392kb ANT-C plus 320kb BX-C) (Negre and Ruiz, 2007). The single cluster of *Tribolium* extends over 756kb (Shippy et al., 2008). Other arthropods too have loose clusters (Pace et al., 2016). Four core Hox members of *C. elegans* are spread over 300kb with two additional *AbdB* genes located 4 to 6 Mb away on the same chromosome (Aboobaker and Blaxter, 2003, Gutierrez et al., 2003, Van Aucken et al., 2000).

In contrast to the compact clusters of vertebrates, the more loose clusters of *C. elegans, Drosophila* and other arthropods contain interspersed non-Hox genes (Gutierrez et al., 2003, Pace et al., 2016). In at least some members of Urochordata, Platyhelminthes and Xenacoelomorpha the Hox genes have become extensively dispersed from their ancestral clustered arrangement (Figs. 3,4).

Compactness of vertebrate Hox clusters is associated with presence of ‘global’ regulatory elements identified mainly beyond each of the two cluster ends (Duboule, 2007). These elements are in addition to more local regulatory elements positioned within the clusters and regulating mainly nearby Hox genes. One proposal is that the vertebrate clusters acquired and maintained compactness so that clustered Hox genes could be co-ordinately regulated by not-too-distant global regulatory elements (Duboule, 2007, Spitz 2008). Such regulatory elements are in addition to more local regulatory elements positioned within the clusters and regulating mainly nearby Hox genes. One proposal is that the vertebrate clusters acquired and maintained compactness so that clustered Hox genes could be co-ordinately regulated by not-too-distant global regulatory elements (Duboule, 2007, Spitz 2008). Such regulatory elements are in addition to more local regulatory elements positioned within the clusters and regulating mainly nearby Hox genes.

Further work is needed to evaluate this possibility.

**The core ancestral cluster of Hox genes arose successfully only once during bilaterian evolution**

The Hox gene cluster expanded probably one gene at a time by the process of tandem gene duplication (Lewis, 1998). At each event, which followed an error during meiotic crossover, both copies of a particular Hox gene (maternal and paternal) came to lie in tandem on the same chromosome. One copy could then mutate to acquire a new expression boundary and function, thereby permitting development of a new structural feature along the head-tail axis. This may have conferred a selective advantage, or not, according to the selective pressure of the environment.
to the rules of natural selection.

It is assumed that the particular Hox gene that duplicates, and how it mutates, are random processes, and so it might reasonably be expected that many different collinear Hox gene clusters could have evolved (Gaunt and Gaunt, 2016). Surprisingly, however, the same core set of Hox genes, often in the same transcriptional orientation, has been found throughout protostomes, deuterostomes and, perhaps in its juvenile form, in Xenacoelomorpha (Figs. 3-5). In all of these bilaterians a gene that is structurally Hox1 typically specifies anterior embryonic parts and a gene that is structurally Hox9-13 specifies posterior parts. The most likely explanation for these findings is that the Hox cluster of bilaterians evolved successfully only once. That is, bilaterians arose from a single ancestor. Its selective advantage may have been its collinear Hox cluster, though it may have been some other novelty such as acquisition of the mesoderm germ layer or bilater symmetry.

Conservations in spatial and temporal collinearities

Spatial collinearity in Hox expression is seen from the above descriptions to be widespread throughout bilaterians (Fig. 6). In some animals the cluster has undergone partial or complete disruption (e.g. in some members of Urochordata, Platyhelminthes and Xenacoelomorpha) but the Hox genes nevertheless maintain patterns of expression reminiscent of their clustered ancestral arrangement. Although spatial collinearity is common, many exceptions are known. This may be for genes that remain in the expected position in their cluster but have acquired an unexpected expression pattern (e.g. Hox6 in the European amphioxus) (Pascual-Anaya et al., 2012). Or, it may be for genes that retain an expected pattern of expression but have acquired an unexpected cluster position (e.g. Hox1 in C. elegans) (Tihanyi et al., 2010). Temporal collinearity is seen to be confined to vertebrates, cephalochordates, some arthropods, and some annelids.

The significance of Hox gene collinearities

Three proposals for the role of collinearity are mentioned here. First, the body of the ancestral bilaterian may have developed in an anterior to posterior temporal progression, with new structures developing from cells that emerged from a posterior growth zone and which displayed temporal collinearity (Ferrier and Holland, 2002, Monteiro and Ferrier, 2006). This would have been as is seen today in vertebrates, some arthropods and annelids. It is further suggested that progressive activation of Hox genes within the growth zone is due to a time-regulated, progressive opening in chromatin structure along the cluster (Duboule, 1994). This would explain the need for Hox gene clustering, and also why gene order on the chromosome must be collinear with the initial time (temporal collinearity) and position (spatial collinearity) of gene expressions in the embryo.

This ‘chromatin opening’ model is supported by the fact that species showing temporal collinearity have so far been found to develop from a posterior growth zone, and to have substantially intact Hox clusters without gene inversions or interspersed non-Hox genes: for example, vertebrates and an annelid (Duboule, 2007, Frobius et al., 2008). However, more species are needed to test this correlation further. Studies on Parhyale, though incomplete, already indicate that Hox genes need not necessarily be compacted in the cluster for temporal collinearity (Serano et al., 2016). Supporters of the chromatin opening model suggest that many animal groups devised alternative ways to set up their Hox expression patterns, so that they no longer required either temporal collinearity or an intact Hox cluster. This may have been to achieve more rapid embryonic development as in Drosophila (Ferrier and Minguillon, 2003), or to adopt a largely lineage-dependent embryonic strategy as in C. elegans (Duboule, 1992, Duboule, 2007). Several observations, recently reviewed (Gaunt, 2015), on Hox genes transposed within the cluster, Hox/lacZ transgender expressions, and discrepancies in timing between chromatin opening and Hox expressions have not readily supported the chromatin opening model.

The ‘gene segregation’ model provides a second possible explanation for spatial collinearity. For a partially overlapping array of Hox gene expressions, as is found in most species (Fig. 6), spatial collinearity results in the minimum number of boundaries (that is, maximum segregation) between the active and inactive genes of a Hox gene cluster (Gaunt, 2015, Gaunt and Gaunt, 2016). The proposal here is that boundaries are prone to accidental leakage of the Hox-active and Hox-inactive chromatin states, and that spatial collinearity evolved to minimize this risk. A third ‘chromatin closing’ model notes that spatial collinearity results in maximal contiguity between inactive genes of a Hox gene cluster, and suggests that this may be essential if the repressed chromatin state must spread from one Hox gene to its neighbours (Gaunt 2015). In terms of models two and three, temporal collinearity in species that develop by posterior extension is viewed as a consequence of spatial collinearity. The ancestral clustering, which first arose as a result of the Hox gene tandem duplication mechanism, is maintained over evolutionary time by constraining forces such as by need to share enhancer elements (Graham et al., 1989) or chromatin repression (Gaunt 2015), by need to contain Hox genes within the same nuclear locality (Bantignies et al., 2011, Chan et al., 2015), or perhaps by secondary development of a chromatin opening mechanism. Hox gene clusters have become extensively dispersed in at least some members of Urochordata, Platyhelminthes and Xenacoelomorpha (Figs. 3,4). In terms of model two, this dispersion may have conferred selective advantage by further minimizing leakage of active and inactive states between Hox genes that were initially adjacent.

Ancestral and derived strategies in Hox gene function

Comparisons of Hox gene expression in cnidarians and bilaterians show that spatial collinearity flourished only with the advent of bilateral symmetry, although this does not exclude the possibility that it may have first arisen in a pre-bilaterian (DuBuc et al., 2018). Since spatial collinearity in Hox expression is widely conserved between animal groups, it is a common view that Hox genes retain an ancestral function among bilaterians to specify organization along the head-tail axis. That is, in terms of Fig. 2, there is a universal bunch of Hox keys that regulates head-tail organization in all of these animals. This view is certainly consistent with results from arthropods (Martin et al., 2016), vertebrates (Mallo et al., 2010) and, by inference, their P-DLCA ancestor. However, more Hox gene function analyses are needed to test how widely this mechanism has been conserved in other bilaterians, particularly those that develop without segmentation, and those that favour a lineage-dependent mode of development.
While most or all animals apparently retain at least part of the ancestral Hox gene function, observations such as the following indicate that alternative, derived mechanisms have been commonly adopted to contribute to body designs. 1) Knockdown experiments on sea squirt C. intestinalis Hox genes indicate that they may not all play a role in larval development (Ikuta et al., 2010), even though they do regulate development during subsequent metamorphosis (Sasakura and Hozumi, 2018). 2) Most of the Hox genes in sea urchin S. purpuratus are not expressed during formation of the free swimming bilaterian larva (Areñas-Mena et al., 1998). 3) Three of the six C. elegans Hox genes can be lost without preventing development to fertile adults (Van Auken et al., 2000). 4) Hox genes in many species, especially lophotrochozoans, have been extensively co-opted to facilitate development of novel structures without apparent compliance with ancestral collinearity rules. Examples include shell and apical organ formation in snails (Fritsch et al., 2015, Hinman et al., 2003, Samadi and Steiner, 2010), the light organ in squid (Lee et al., 2003), and the foot in a rotifer (Frobius and Funch, 2017).

Points 1 to 3 above show that a bilaterian animal may develop, at least to a large extent, without use of the ancestral head-tail Hox patterning mechanism. It has been suggested that this may be facilitated by lineage-dependent (deterministic) rather than regulative development, and that the former is more commonly utilized in nematodes, molluscs, annelids and some deuterostomes (Aboobaker and Blaxter, 2003, Areñas-Mena et al., 1998, Duboule, 1992, Duboule, 2007, Seo et al., 2004). While this may be so, clear-cut distinction between these two modes of development has been questioned (Lawrence and Levine, 2006).

It was proposed that Hox-cluster genes of the ancestral bilaterian regulated regionalization only in neuructoderm (García-Fernandez, 2005) and that Hox head-tail patterning was later co-opted to mesodermal structures in segmented animals such as annelids, arthropods and vertebrates (Samadi and Steiner, 2010). Supporting this, Hox expression is principally neural in some conchiferan molluscs (Hinman et al., 2003, Samadi and Steiner, 2010). However, mesodermal expression of Hox genes is reported in unsegmented acelifers (Fritsch et al., 2015) and a platyhelmint (Nogi and Watanabe, 2001); and both neural (Hejnol and Martindale, 2009) and ‘parenchymal’ expressions (Moreno et al., 2009) are reported in acellos.

A major difference in developmental strategy between species lies in whether or not their embryos use a posterior growth zone in order to elongate the head-tail axis. In vertebrates, for example, the head end develops before the tail end, whereas in Drosophila all parts develop at the same time. It might be expected that all species that develop morphologically distinct axial regions from a posterior growth zone will also show temporal collinearity in Hox gene expressions. Some authors suggest that development from a posterior growth zone (Gold et al., 2015) and temporal collinearity (Ferrier and Holland, 2002, Monteiro and Ferrier, 2006) are ancestral bilaterian conditions, but this is uncertain. More research is needed to shed light on the likely nature of the ancestral bilaterian. Of potential value is the possibility that this might be represented today by the Xenacoelomorpha. Acellos do not display temporal collinearity (Hejnol and Martindale, 2009), and therefore presumably have no apparent need for the chromatin opening mechanism, or for a posterior growth zone. Spatial collinearity might therefore be the ancestral bilaterian condition, with temporal collinearity evolving secondarily, and perhaps on multiple occasions. Hopefully, we shall soon understand more clearly whether acellos may, or may not, provide a useful model of the ancestral bilaterian condition.

Acknowledgement

I thank Rob Asher for support.

References


BYRNE, M., MARTINEZ, P. and MORRIS, V. (2016). Evolution of a pentamerical body plan was not linked to translocation of anterior Hox genes: the echinoderm Hox cluster revisited. Evo. Dev. 18: 137-143.


DAVID, B. and MOOI, R. (2014). How Hox genes can shed light on the place of


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

**The significance of Hox gene collinearity**
Stephen J. Gaunt  
*Int. J. Dev. Biol.* (2015) 59: 159-170  
https://doi.org/10.1387/ijdb.150223sg

**Function and specificity of Hox genes.**
Foronda, D., De Navas, L.F., Garaulet, D.L. and Sanchez-Herrero, E.  
*Int. J. Dev. Biol.* (2009) 53: 1409-1419  
doi: 10.1387/ijdb.072462df

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

**The origins of axial patterning in the metazoa: how old is bilateral symmetry?**
John R. Finnerty  
*Int. J. Dev. Biol.* (2003) 47: 523-529

**Evolutionary shifts of vertebrate structures and Hox expression up and down the axial series of segments: a consideration of possible mechanisms**
S J Gaunt  
http://www.intjdevbiol.com/web/paper/10761855

**The bithorax complex: the first fifty years.**
Lewis, E. B.  

**Conservation in the Hox code during morphological evolution**
Stephen J. Gaunt  
*Int. J. Dev. Biol.* (1994) 38: 549-552

5 yr ISI Impact Factor (2016) = 2.421