Reassessing the Role of *Hox* Genes during Vertebrate Development and Evolution

Moisés Mallo*

Since their discovery *Hox* genes have been at the core of the established models explaining the development and evolution of the vertebrate body plan as well as its paired appendages. Recent work brought new light to their role in the patterning processes along the main body axis. These studies show that *Hox* genes do not control the basic layout of the vertebrate body plan but carry out region-specific patterning instructions loaded on the derivatives of axial progenitors by *Hox*-independent processes. Furthermore, the finding that *Hox* clusters are embedded in functional chromatin domains, which critically impacts their expression, has significantly altered our understanding of the mechanisms of *Hox* gene regulation. This new conceptual framework has broadened our understanding of both limb development and the evolution of vertebrate paired appendages.

**The Vertebrate Basic Body Plan Is Laid Out Independently of *Hox* Genes**

Vertebrates display a remarkable variety of body sizes and shapes, typically involving features along their main body axis and their paired appendages (see Glossary). However, the basic developmental principles generating the various structures are largely shared among vertebrates. The various body attributes along the main body axis are assembled sequentially in a head-to-tail sequence as the embryo extends at its posterior end. This results from the activity of dedicated axial progenitors producing the raw material that eventually forms the various embryonic tissues [1,2]. Although continuous, axial extension can be divided into three major steps each regulated by a distinct gene network [3–5]. During the first step, typically known as head development, the embryo generates the brain and heart primordia together with musculoskeletal structures of the head and neck. This is followed by formation of the trunk, which holds most of the internal organs involved in vital and reproductive functions. The final step of axial extension is devoted to tail formation, essentially comprising vertebrae, muscles, and a variable amount of neural tissue. Differences in the overall extent of body elongation during development, as well as the portions of this axial growth devoted to making head, trunk, or tail structures, are among the most relevant parameters generating anatomical diversity in vertebrates. For this reason both the mechanisms controlling these processes and their role in vertebrate evolution have attracted plenty of attention over the past few decades.

Since their discovery *Hox* genes (Box 1) have been considered the major drivers of morphological evolution in the animal kingdom [6]. Comparative expression analyses in embryos of vertebrates with clearly distinct body distributions revealed a close correlation between the expression boundaries of particular *Hox* genes and specific landmarks in the axial skeleton [7,8], suggesting key roles for these genes in setting up the basic vertebrate body plan. A large variety of genetic studies, mostly in the mouse, confirmed the relevant contributions of *Hox* genes to patterning processes along the main body axis [9,10]. For instance, *Hox4* genes are involved in proper patterning of the neck vertebrae [11], *Hox5*, *Hox6*, and *Hox9* genes control

**Highlights**

The basic layout of the vertebrate body is outlined in the axial progenitors mostly through *Hox*-independent mechanisms.

*Hox* genes are carriers of patterning information loaded onto the derivatives of axial progenitors that guides the production of body structures congruent with their axial level.

Selective target inactivation allows the shutting down of a subset of *Hox*-dependent functions while keeping others active in the same domain. This increases the flexibility of evolutionary processes.

The processes regulating *Hox* gene expression in the proximal and distal regions of the limb buds occur in two alternative functional chromatin domains.

Modification of *Hox* regulatory processes within chromatin domains might have played a role in the evolution of vertebrate paired appendages.

---

1 Instituto Gulbenkian de Ciencia, Rua da Quinta Grande 6, 2780-156 Oeiras, Portugal

*Correspondence: mallo@igc.gulbenkian.pt (M. Mallo).
various aspects of ribcage development [12,13], and Hox10 and Hox11 genes are essential for the formation of the lumbar and sacral areas of the axial skeleton, respectively [14,15]. However, these studies consistently failed to show significant changes in the basic head/trunk/tail distribution of the body, even in cases where alterations in Hox gene activity or expression produced extreme phenotypes in the axial skeleton. Only Hox13 genes have been found to contribute to this process, by playing a role in determining the final length of the tail region [16,17]. Overall these data indicated that the basic head/trunk/tail structure of the embryo is most likely not under direct Hox regulation, which seemed to contradict the prevailing model for the evolution of the vertebrate body plan along the main body axis.

**Gdf11 and Oct4 Are Upstream of Hox Genes in the Body-Patterning Hierarchy**

Recent findings brought new light to this issue, identifying Oct4 and Gdf11 signaling as major players in the establishment of the basic body plan, acting upstream of Hox genes. Gdf11, a member of the Tgfβ family expressed in the posterior embryonic area starting at mid-gestation [18,19], was shown to be a key activator of the trunk-to-tail transition [20], a process in which it shows partial redundancy with Gdf8 [21]. Mice with impaired Gdf11 signaling have longer trunks resulting from delayed activation of this transition from the early stages of development, as reflected in the significantly more posterior position of the hindlimbs and cloaca, which mark the posterior end of trunk-associated structures such as the lateral mesoderm and the endodermal tissues contributing to the internal organs [19,20]. Conversely, premature activation of Gdf11 signaling resulted in more anterior induction of the trunk-to-tail transition, which placed the hindlimb next to the forelimb bud and thus led to embryos without a trunk [20]. Gdf11 expression in different vertebrates seems to give further support for the role of this signaling in the evolution of vertebrate trunk length [22,23].

Other studies revealed that the pluripotency factor Oct4 plays a somewhat complementary role in the layout of the basic body plan, as it promotes extension through the trunk region. Such new role for Oct4, consistent with its expression dynamics during mouse development [24,25], was suggested by genetic studies. In particular, conditional Oct4 inactivation in mouse embryos at early stages of trunk formation resulted in embryos without a trunk but that still contained recognizable tails [26]. A different set of studies showed that Oct4 is also sufficient to extend the vertebrate trunk [23]. Prolonging the period of Oct4 activity in mouse embryos resulted in longer trunks at the expense of the tails. In addition, molecular analyses in snake embryos indicate that their remarkably long trunks might be the result of an increased period of Oct4 activity during embryonic development [23]. Altogether, various lines of evidence place the balance between Oct4 and Gdf11 activities at the top of the hierarchy of regulatory processes controlling the basic features of the vertebrate body plan by playing fundamental roles in determining the relative contributions of the different body sections to the animal’s anatomy.

**Where Do Hox Genes Fit In This Scheme?**

Expression analyses indicated that in mouse embryos with modified Gdf11 or Oct4 activity, 5’ Hox genes became activated at axial levels congruent with the new position of the hindlimb bud and, thus, the trunk-to-tail transition [19,20,23,27,28]. Interestingly, some Hox genes of 3’ groups showed a complementary behavior, best seen in embryos with longer trunks (i.e., Gdf11 mutants or transgenics with sustained Oct4 expression), where their expression spread into more posterior embryonic areas [19,20,23,28] (Figure 1). The same global patterns of Hox gene expression were observed in the natural setting presented by the snake embryo [8,29]. These studies thus place Hox gene expression downstream of Oct4 and Gdf11 signaling. A
Box 1. Basic Concepts of Vertebrate Hox Gene Organization and Expression

In mammals, which provide a paradigm to outline the basic concepts of Hox gene organization (Figure I), Hox genes are distributed in four clusters, named HoxA, HoxB, HoxC, and HoxD, thought to result from two successive duplications of an ancestral cluster. Hox genes are subdivided in 13 groups (from Hox1 to Hox13) based on sequence homology and their position in the cluster. Hox genes with lower numbers are located at the 3’ side of the cluster and those with higher numbers toward the 5’ end. In general, Hox gene activation is sequential, following their order in the cluster in a 3’-to-5’ direction, a property known as collinearity. Because Hox gene activation is concurrent with axial extension and limb growth, the different areas of the body and limbs express distinct combinations of Hox genes. Other vertebrates have different numbers of Hox clusters. For instance, the zebrafish has seven clusters, known as HoxAa, HoxAb, HoxBa, HoxBb, HoxCa, HoxCb, and HoxDa, resulting from an additional duplication followed by the loss of one whole cluster.

Figure I. Basic Structure of the Mammalian Hox Clusters.

similar hierarchy was observed at the cell/tissue level, as Oct4 and Gdf11 regulate patterning directly on the axial progenitors (and thus at the top of the cellular hierarchy controlling body formation) whereas Hox gene activity is most important in tissues derived from those progenitors [20,23]. Hence, the global picture emerging from these studies (Figure 2) is that, in a first step, the basic layout of the vertebrate body is outlined in the axial progenitors mostly through Hox-independent mechanisms. As an integral part of these basic processes, the different cell types produced by these progenitors become loaded with patterning information corresponding to the axial level where they will differentiate. Hox genes thus belong to this second patterning layer providing axial identity to mesodermal, neural, and maybe also endodermal derivatives of the axial progenitors [9,30,31]. Thus, their role in the evolution of the vertebrate anatomy is exerted in these tissues where specific combinations of Hox gene expression determine regional variations in the main body domains. Again, snakes provide a natural example to illustrate this hypothesis. In particular, the morphology of the trunk vertebrae of different snake species revealed that this region is not uniform as was classically considered but regionalized along the anterior–posterior axis following species-specific patterns, most likely resulting from specific variations in Hox gene expression, set independently of their trunk length [32].

So far little is known about how Oct4 and Gdf11 signaling controls Hox gene expression. The two-way complementarity of Hox regulation by Oct4 and Gdf11 (i.e., that 5’ Hox genes are repressed by Oct4 and activated by Gdf11 whereas 3’ Hox gene expression expands posteriorly following extended Oct4 activity or on Gdf11 inactivation) might provide clues to understand this process. The regions of the Hox cluster under differential regulation by Oct4 and Gdf11 roughly correspond to their distribution within the two **topologically associating domains** (TADs) covering the HoxA and HoxD clusters [33,34]. TADs demarcate chromatin territories facilitating regulatory interactions [35,36] and the TAD structure associated with the
Hox clusters has been shown to be relevant for the regulation of Hox gene expression during limb bud development \cite{33,37}. In particular, as mentioned below, as the limb bud grows distally the Hox regulatory landscape switches from the 3’ to the 5’ TAD coincident with the activation of Hox genes at the 5’ end of the cluster (Figure 2). Hox gene regulation in the major body axis might also fit in a similar general scheme involving a regulatory switch between TADs orchestrated by the balance between Oct4 and Gdf11 signaling activities. Involvement of the TAD structure in Hox gene regulation in the main body axis is supported by recent findings showing

Figure 1. Oct4 Promotes Trunk Elongation and Gdf11 Activates the Trunk-to-Tail Transition (Marked by the Hindlimb Position). As part of these activities, they regulate Hox gene expression. 3’ Hox genes are kept inactive at trunk levels by Oct4 and become activated when Gdf11 signaling takes over. Regulation of several 3’ Hox genes follows a complementary pattern. Hox gene expression in the areas of the limb bud generating the arm and forearm is under the control of regulatory interactions in the 3’ topologically associating domain (TAD) (in blue). The regulatory landscape switches to the 5’ TAD in the distal limb (purple) to produce hand-specific Hox gene expression. Does TAD organization also impact Oct4/Gdf11-mediated Hox regulation in the main body axis?

Figure 2. Patterning in the Main Body Axis Occurs in Two Consecutive Stages. In the first stage, the final size of the head/trunk/tail region is determined in the axial progenitors mostly through Hox-independent mechanisms, including Oct4 and Gdf11 activities. The transition between these regions can be identified in the developing embryo by the positions of the forelimb (FL) and hindlimb (HL). The second stage occurs in the derivatives of the axial progenitors by loading them with axial level-specific patterning information (including Hox genes) that guides their differentiation into the appropriate body structures.
that Wnt3/Wnt3a signaling and Cdx proteins sequentially activate 3' HoxA genes in the epiblast through specific regulatory interactions occurring in the 3’ TAD [38,39]. Additional findings suggest that Oct4 activity might also fit within this regulatory scheme. In particular, it has been shown that embryonic stem cells Oct4 primes Hoxa1 and Hoxb1 for activation on exposure to retinoic acid [40], a physiological activator of Hox gene expression. These observations indicate that Oct4 might indeed be involved in activation of 3’ Hox genes. In addition, Oct4 has been reported to bind the HoxA cluster at the intersection of the TADs, and this binding might be important for proper chromatin structure and Hox gene expression in cell lines [41]. Even less is known about Hox regulation by Gdf11 signaling. Currently, the only connection between Gdf11 signaling and Hox gene expression was suggested by the study of an enhancer in the Hoxd11 3’ UTR required for proper activation of Hoxd11 [42]. This enhancer contains a phylogenetically conserved Smad binding site essential for its activity in transgenic assays [43]. Whether this enhancer plays a more general role in the regulation of the HoxD cluster remains to be determined.

Selective Inactivation of Specific Hox Targets Adds Evolutionary Flexibility

Although a large part of Hox-dependent morphological variations in the axial skeleton might result from changes in Hox gene expression, modulation of specific downstream aspects of Hox activity can also have a relevant impact on this process. An interesting example is the origin of the large rib numbers of Paenungulata (including elephants and manatees), extending further posteriorly than in other mammals to cover most of their presacral skeleton [44]. This feature looks surprisingly similar to the skeletal phenotype of mutant mice totally lacking Hox10 activity [14]. Since Hox10 genes play other essential roles in mammalian development (e.g. [45]) it is unlikely that Paenungulata just happened to lose these genes. However, the genome of animals belonging to this clade contain an SNP in an enhancer mediating Hox10 activity during rib development [13] that precludes binding of Hox10 proteins [46], thus simulating a functional Hox10 inactivation restricted to the developing axial skeleton. Interestingly, the same polymorphism was found in snakes, where it is also very likely to interfere with Hox10 rib-repressing activity. In particular, expression of the snake Hoxa10 gene, which blocks rib formation when tested in mouse embryos [46], extends well into rib-forming somites of the snake embryo without apparent negative effects on rib development [8]. This example illustrates well how regulation of Hox gene activity by selective interference with specific downstream targets generates substantial evolutionary flexibility, as it can affect a subset of the protein's functions while keeping other relevant activities in the same domain or even allowing the acquisition of additional functions in that area.

The Chromatin Structure of Hox Clusters Impacts Limb Morphogenesis . . .

Hox genes also play essential roles in the morphogenesis and evolution of vertebrate paired appendages. Their expression in the tetrapod limb bud occurs in two sequential phases, the first associated with proximal limb segments (stylopod and zeugopod) and the second with the distal domain, the autopod [47]. Recent data indicate that the regulation of these two phases of Hox gene expression is closely connected to the 3D chromatin topology covering the Hox clusters [33,34,37]. Interaction analyses performed on the HoxA and HoxD clusters, the major Hox players in limb development, revealed that during the first expression phase Hox genes in the 3’ TAD (up to Hox1) are under the control of regulatory regions in this TAD [33,48]. This regulation is maintained in the proximal limb domain at later developmental stages. However, in the developing autopod the Hox regulatory landscape switches drastically. In this region Hox genes in the 5’ TAD (those of the Hox9 to Hox13 groups), change their functional interactions to become regulated by enhancers in this topology domain [33] (Figure 2). These latter interactions are required for the autopod-type Hox gene expression, including the
‘reverse collinearity’ [49] (i.e., the strength and extension of Hox gene expression decreases in a 5’-to-3’ direction) and, importantly, the inactivation of Hoxa11 in the autopod [50], which, as discussed below, is directly linked to Hox13 gene activity. This regulatory switch creates a ‘Hox-free’ area between the proximal and distal limb domains thought to be required for wrist/ankle joint development [33,47].

Although the mechanisms involved in this regulatory switch remain incompletely understood, it is clear that Hox13 genes are a key component of the process. Consistent with this, Hox13 mutants are unable to elicit the switch and as a consequence characteristic features of the proximal limb bud extend into the prospective autopod region, hindering proper morphogenesis in this area [51,52]. Hox13 genes play several roles in this process. Besides disconnecting regulatory activities that involve the 3’ TAD [51], Hox13 proteins also interact with specific enhancers in the 5’ TAD to promote the late phase of Hox gene expression in the limbs [51,52]. Finally, Hox13 proteins control Hoxa11 expression as well, by activating an enhancer that promotes the transcription of an antisense Hoxa11 transcript (Hoxa11as) that silences in-cis Hoxa11 gene expression [50].

... and the Evolution of Paired Appendages
Progress in understanding the control of Hox activity during limb development also provided new insights into how the tetrapod limb might have evolved from paired fins of fishes. While the early stages of fin and limb development are similar, they clearly differ during the formation of their distal domains. The distal-most limb domain, the autopod, is dominated by mesenchymal tissue that eventually provides the substrate for digit formation, whereas the distal fin domain comprises an epithelial structure, the fin fold, that eventually holds the fin radials. The origins of the differences between distal fin and limb development remain not totally understood, but recent data suggest that they might involve the acquisition by amniotes of novel regulatory regions involved in the second phase of Hox gene expression in the limbs. 3D structure and interaction analyses indicate that Hox clusters of teleosts and amniotes largely share both their TAD structure and the preferred contacts between Hox13 genes and genomic regions in the 5’ TAD [37,53]. Despite this, functional analyses revealed different regulatory potentials for the 5’ TAD sequences of teleosts and mammals. In particular, when pufferfish BAC clones containing the HoxAa, HoxAb, or HoxDa cluster together with the adjacent 5’ region were introduced into mice, they activated fish Hox gene expression in proximal limb domains, failing to extend into the autopod [37]. Consistent with this, the zebrafish 5’ TADs lack several key regulatory elements required for autopod-type Hox gene expression, including regions controlling Hox13 activity [54,55] and the Hox13-responsive enhancer promoting Hoxa11as transcript expression [50]. Typical features of fin Hox gene expression, like overlapping Hoxa11 and Hoxa13 expression domains [56], non-detectable Hoxa11as transcript [50], and expression of 5’ Hox genes with no sign of reverse collinearity [57], are consistent with the lack of those enhancer elements.

Interestingly, at least some of the 5’ regulatory elements absent from teleosts can activate reporter expression in the zebrafish fin bud with appropriate spatial distribution [50,54], suggesting that teleost paired appendages contain the molecular machinery required to control those enhancers. This observation, together with the similar chromosomal architecture and interaction profiles of fish and amniote Hox clusters, indicate that the teleost fin could easily acquire an autopod-type regulatory landscape on incorporation of the relevant enhancers into their 5’ TADs. This scenario has been suggested to have contributed to the fin-to-limb evolutionary transition [37]. Interestingly, the gar HoxA region seems to have a regulatory structure somewhere between teleosts and amniotes, as it contains at least one of the
enhancers absent from zebrafish that is able to reproduce autopodal Hoxa13 gene expression when tested in mouse embryos [54].

A Central Role for Hox13 Genes in Distal Development of Paired Appendages

A variety of functional experiments suggest that the specific expression features of Hox13 genes, including their expression levels, are relevant for the differential developmental characteristics of the distal limb and fin. Gradual reduction of the Hox13 dosage in the mouse correlated with the progressive acquisition of several fin-like features, including the distal extension of Hoxa11 expression that invades the Hox13 domain, digit shortening, and the absence of the small skeletal elements characteristic of the zeugopod/autopod joint [58,59]. Conversely, increasing Hox13 levels in zebrafish fins expanded the mesodermal core at the expense of the ectodermal fold and promoted the expression of some typical distal limb bud markers [55].

Hox13 genes are also required for distal fin development. Cell-tracing studies indicate that distal radials derive from Hox13-positive mesenchyme entering the fin fold and genetic experiments revealed that those structures fail to form in the absence of Hox13 activity [60]. These observations reopen the old discussion about the homology between digits and distal radials. Such homology could actually help to explain the larger number of radials versus digits on the basis of quantitative and qualitative aspects of Hox gene expression. The characteristic Hoxa11 and Hox13 expression overlap in fin buds [56] could be part of the mechanism, as experimental Hoxa11 activation in the Hox13 expression domain resulted in polydactyl limbs [50]. In addition, reduced Hox13 activity in a Gli3-null context (thus with reduced hedgehog activity) leads to a significant increase in digit number [59]. It was suggested that under these conditions Hox13 activity would determine the wavelength of the Turing-type mechanism controlling digit number [59,61], with this number increasing with the reduction of Hox13 activity. It is thus possible that a Turing system acting in the relatively low Hox13 context of the fin could contribute to the large radial fin numbers, although the precise components of this system might not be the same as in mammals.

Can Snakes Blame Hox Genes for Having Lost Their Limbs?

In addition to the fin-to-limb evolutionary transition, Hox genes have been suggested to play a role in the loss of paired appendages by snakes [62], although the precise mechanisms or even whether this is indeed true awaits direct experimental proof. It has been reported that Tbx5, a key regulator of forelimb bud induction [63], is activated in the lateral mesoderm by Hox4 and Hox5 genes [64], suggesting a possible mechanism by which changes in Hox gene expression might have contributed to the loss of forelimbs characteristic of snakes. However, regardless of whether these Hox genes are essential for forelimb induction [11,12], expression studies showed that members of those Hox groups as well as Tbx5 itself are expressed in the lateral mesoderm of snake embryos [8,29], indicating that the lack of forelimbs in snakes is most likely to require alternative explanations.

Still, modification of Hox gene activity might account for the inability of the hindlimb buds of ancient snakes to develop into full-grown limbs. It has long been shown that expression of Shh, a key player in limb bud growth and patterning [65], is defective in the snake limb buds [62]. Shh limb expression relies on a distal enhancer that is functionally compromised in snakes [66,67]. Interestingly, the modifications found in the python Shh enhancer render it unable to respond to Hox proteins [67]. Considering that Hox gene activity is required to initiate and/or maintain Shh expression in the mouse limb buds [68], it is likely that the lack of response of the Shh enhancer
to Hox proteins has played a fundamental role in the developmental arrest of the snake hindlimb bud.

Concluding Remarks and Future Perspectives

Over the past years a considerable amount of work has changed the place held by Hox genes in the overall hierarchy of the patterning cascade regulating the vertebrate body structure and their variations among taxa. The resulting new model not only describes how patterning information can traverse the various hierarchical levels controlling axial patterning and the production of functional body structures, but also sheds light on some unexplained aspects of Hox mutant phenotypes. So far it is unclear how Oct4 and Gdf11 control Hox genes (see Outstanding Questions). The effect of these two factors on Hox gene expression suggests a mechanism relying on the modulation of regulatory activities organized in TADs, akin to the model recently described in the limb buds. However, direct experimental evaluation is needed to elucidate whether this is indeed the case or whether these factors operate according to entirely different principles. It might also be interesting to determine whether Gdf11/Gdf8 signaling participates in Hox gene regulation in the limb buds, as both Gdf11 and Gdf8 are expressed in these appendages \[18,69\] and their simultaneous inactivation led to strong limb malformations \[21\]. Another interesting question is the extent to which other aspects of Hox regulation identified in the limb buds, like those involving Hox13 genes, also operate in the main body axis. The subsequent findings will show the extent to which basic mechanisms of Hox gene regulation are conserved throughout developmental territories. Finally, it will be important to understand how the mechanisms of Hox gene regulation coordinate with other relevant features associated with Hox gene activation, like the progressive opening of the Hox clusters \[70\], as well as their relationship with other factors known to regulate patterning processes in Hox-expressing embryonic areas, like the main body axis or paired appendages.

Acknowledgments

The author thanks Ana Casaca, Rita Aires, and Andre Dias for insightful comments on the manuscript. Work in M.M.’s laboratory is supported by grants PTDC/BEX-BID/0899/2014 Fundação para a Ciência e a Tecnologia (FCT, Portugal) and SCML-MC-60-2014 (from Santa Casa da Misericordia de Lisboa, Portugal).

References

11. Horan, G.S.B. et al. (1995) Compound mutants for the paralogous hoxa-4, hoxb-4, and hoxd-4 genes show more complete homeotic transformations and a dose-dependent increase in the number of vertebrate transformed. Genes Dev. 9, 1667–1677
42. Zikán, J. et al. (1997) Deletion of a HoxD enhancer induces transcriptional heterochromy leading to transposition of the saurcom. EMBO J. 16, 4393–4402