

Organizing Axes in Time and Space; 25 Years of Colinear Tinkering

Marie Kmita and Denis Duboule*

During vertebrate development, clustered genes from the *Hox* family of transcription factors are activated in a precise temporal and spatial sequence that follows their chromosomal order (the "*Hox* clock"). Recent advances in the knowledge of the underlying mechanisms reveal that the embryo uses a variety of strategies to implement this colinear process, depending on both the type and the evolutionary history of axial structures. The search for a universal mechanism has likely hampered our understanding of this enigmatic phenomenon, which may be caused by various and unrelated regulatory processes, as long as the final distribution of proteins (the HOX code) is preserved.

In the course of animal embryogenesis, distinct morphological identities are established along the body axes. For example, mammals have thoracic vertebrae that bear ribs, whereas cervical vertebrae do not, and digits are eventually positioned at the distal ends of our limbs, rather than elsewhere. The genetic mechanism underlying this patterning system was uncovered by studying mutations in *Drosophila* where correct structures were wrongly positioned. Drastic alterations, such as the outgrowth of a limb instead of an antenna or of a wing in place of a haltere, are associated with the misexpression of gene members of the *Hox* family of transcription factors.

Because HOX proteins are at work in animals displaying a variety of morphologies, they likely act as developmental switches, rather than as specific stonework of the body architecture. Twenty-five years ago, Lewis (1) showed that *Hox* genes were clustered along the chromosome, colinear with their domains of action in the thorax and abdomen of flies. This observation was subsequently extended to vertebrates and other animals (2–4), leading to the suggestion that morphological diversity along the body axis was generated by a combinatorial distribution of HOX products [the HOX code (5)]. How do such proteins differentially instruct cohorts of cells about their fates, and how are their functional domains established in time and space?

Decoding the HOX Code; Posterior Prevalence at Work

In flies, eight *Hox* genes belong to the *Antennapedia* and *Bithorax* clusters [the HOM complex (6)]. In many instances, these proteins act by regulating a few downstream effectors sufficient to trigger alternative de-

velopmental pathways, with the discriminative help of various cofactors (7). By these means, an embryonic field can be specified as a whole, such as the limb imaginal disc, where suppression of the *Distalless* limb-promoting function by the *Hox* genes *Ubx* and *abdA* prevents the appearance of limbs in the abdomen (8). This suppression does not occur in the thorax despite the expression of *Antp*. HOX proteins can also affect the determination of a particular cell fate, as has recently been substantiated by the effect of *abdA* on *Rhomboid*, during oenocyte differentiation (9) or that of *Dfd* on the apoptotic gene *reaper* during head formation (10).

This qualitative type of regulation is perhaps less obvious in vertebrates, where quantitative parameters must be considered. First, *Hox* clusters have been duplicated twice so as to generate 39 genes (6, 11), consisting of groups of paralogous genes, that are highly related to each other in sequence. The resulting functional overlaps between paralogous proteins suggest that the developmental pathways concerned may rely on strong quantitative parameters. Second, the mere meristic nature of some vertebrate structures suggests that differential target selection in anteroposterior (AP) patterning is not the sole underlying mechanism. For example, the subtle morphological differences between lumbar and sacral vertebrae likely result from the slight quantitative modifications in a complex equilibrium of target genes, rather than from the critical regulation of a novel effector component (12).

In vertebrates, successively more caudal body levels tend to show an increasing amount and diversity of HOX products, resulting from the expression strategy (Fig. 1). Yet segmented structures do not become more elaborate toward the caudal end of the embryo, nor do they display a greater potential for variation after gene-inactivation experiments, thus excluding a strict combinatorial input. The patterning information delivered at one particular body level primarily

relies on one HOX protein (or a group of paralogs), rather than on a combination of proteins (Fig. 1). The most posteriorly expressed gene usually imposes its function over that of more anterior genes through a suppressive mechanism that does not involve transcriptional repression (13). This posterior prevalence (14, 15) explains why the phenotypes induced by vertebrate *Hox* mutations are restricted either to a few body segments or to the upper morphological window in which a given group of paralogs is at work (16, 17). Large overlapping expression domains are merely another way to produce discrete functional domains (Fig. 1D).

Posterior prevalence is an interesting property for morphological evolution, given that an anterior shift in the expression of a caudal gene would lead to the functional inactivation of more rostral components (such as genes A and M in Fig. 1D). Therefore, the functional interplay between HOX proteins is the result of their colinear distribution along the body and is the essential constraint of the system. Consequently, any mechanism(s) generating this protein distribution may have been evolutionarily selected and implemented in the numerous instances in which this strategy is used.

The Hox Clock

Vertebrate development follows a rostral-to-caudal temporal progression, best exemplified by somitogenesis (18). Initially, *Hox* genes are activated along and in the neighborhood of the primitive streak (19) in an area where precursors of all germ layers are produced. Because the first cells produced will contribute to anterior structures—unlike cells produced subsequently, which will end up at more caudal body levels—the time of activation of a given *Hox* gene was proposed to determine the position of its expression boundary along the body axis. In this way, temporal information could be translated into spatial coordinates. However, the simple observation of a *Hox*-stained embryo reveals that boundaries can be very different in various cell layers such as the neural tube and somitic mesoderm (Fig. 2A), making an explanation based only on time unlikely. Instead, it suggests that various mechanisms are at work in different contexts.

Furthermore, the main experimental approaches used to decipher this mechanism in vertebrates have tempered this interpretation.

Department of Zoology and Animal Biology and National Center of Competence in Research "Frontiers in Genetics," University of Geneva, Sciences III, Quai Ernest Ansermet 30, 1211 Geneva 4, Switzerland.

*To whom correspondence should be addressed. E-mail: denis.duboule@zoo.unige.ch

Spatial expression of a given *Hox* gene does not seem to depend critically on its location either within the cluster or as a randomly integrated transgene (20–22), even though a time delay was often scored in the latter situation. However, whereas targeted rearrangements of a cluster in vivo have uncovered a cluster-intrinsic timing device (23, 24), the deregulation of the posterior gene *Hoxd13* observed after the positioning of an anterior gene nearby is at odds with this view (25). Taken together, these results suggest that the colinearity in temporal activation of *Hox* genes may not be the sole cause of their future spatial transcript distribution. If so, what is the rationale of this *Hox* clock, and how does it work? Unlike the situation in flies, where the activation mechanism depends on factors unequally distributed by the segmentation process, three classes of mechanisms have been evoked to implement colinearity in vertebrates, alone or in combination (Fig. 2).

The End Justifies the Means

The first mechanism relies on the progressive transcriptional availability of *Hox* genes, from one end of the cluster to the other, a process that may or may not be independent of their own transcription. For example, repressive or silencing factors (26, 27) could be released through a passive transition in chromatin states (24). Alternatively, transcription of the genes themselves could help remodel chromatin to allow the next gene to be accessed. The failure of the posterior *HoxD* cluster to efficiently repress the early expressed *Hoxb1/lacZ*

transgene (25) supports the latter possibility and suggests that an early gene can still recruit the necessary factors to be activated in a timely manner, even when positioned within a “closed” domain. In this view, a chromatin-dependent colinear process would involve a transcriptional entry point at one side of the cluster that triggers the processing from a closed to an open configuration through a proximity effect. This would allow for progressively more genes to be transcribed to-

ward the other extremity of the cluster (Fig. 2B). The comparison between the transcriptional availability of genes within and outside their cluster (28) supports this view.

The second scenario proposes that colinear activation in time and space is orchestrated by the integration of locally cis-acting control sequences. A gradient of signaling molecules [such as Fgf (29, 30)] could ultimately be read by a series of upstream sequences showing increasing or decreasing affinities for the effector molecule, all along the cluster. Because local enhancer functions are shared among subsets of neighboring genes, they ultimately provide distinct expression features through unequal partitioning of their activities on the genes they control (Fig. 2C) (31). Although this strategy accounts for the rather precise activation of randomly integrated transgenes, it may not be a key factor in tightly maintaining genes in clusters. Hence, it is likely not a primary mechanistic basis for the *Hox* clock.

The third possibility involves the existence of global enhancer sequences, located outside the clusters, which can regulate several genes in a relatively promoter-unspecific manner (Fig. 2D). The positions of these enhancers, close to either end of a *Hox* complex, introduce an intrinsic regulatory asymmetry that can be subsequently translated into a colinear mechanism. For example, the cycling expression of *Hoxd* genes in the presomitic mesoderm, in coordination with segmentation, involves a regulatory element located outside the cluster, which can act over several genes at different times (32). This regulatory element may be an outcome of the segmentation process, setting up the pace of the *Hox* clock and thereby keeping these two key aspects of patterning in phase with each other (32, 33).

Likewise, colinearity in developing limbs was recently shown to rely on the existence of a global digit enhancer element located far upstream on the other side of the cluster (34). Sequence-specific enhancer tropism, as well as promoter competition, eventually induces the terminal genes to be expressed in the most distal structures, the digits, with a progressively decreasing efficiency (35). In this case, colinearity is determined by the action of a global enhancer, and the necessity of gene clustering for obtaining the observed patterns can be readily understood.

Colinear Tinkering

It is also possible that gene clustering is required to maintain the colinear pattern, rather than to establish it. Indeed, correct transcriptional initiation must be followed by the persistence of the transcript domain to affect the morphology. Premature gene activation induced only subtle or transient variations in the AP level of the expression pattern (24). Conversely, a delay in gene acti-

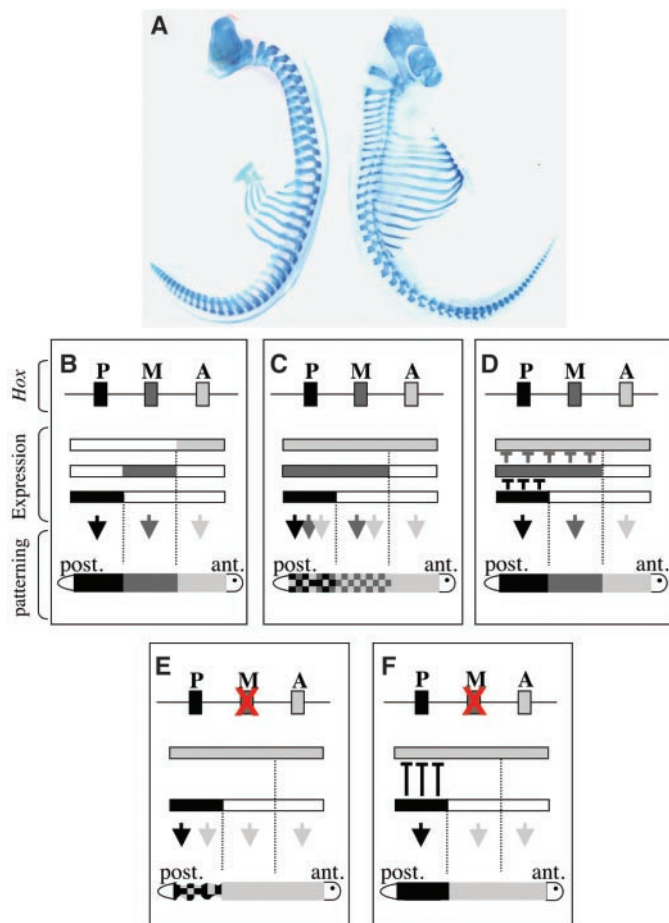


Fig. 1. Axial patterning relies upon a given partition of *Hox* functional domains along the developing AP axis. Various body architectures can derive from the modulation of HOX protein distribution. (A) The axial skeletons of a chick (left) and a mouse (right) embryos stained with alcian blue. The same set of *Hox* genes are differently implemented in time and space, leading to distinct vertebral formula (36, 37). [(B) to (D)] Alternative modes to produce a functional HOX code along the AP axis are illustrated with an anterior (A), middle (M), and posterior (P) gene. Genes can be expressed in discrete domains (B), thus generating a transcription-based partition of HOX functions, as exemplified by *labial* and *Deformed* in *Drosophila* head, or by *Hoxa13* and *Hoxa11* expression in vertebrate limb buds. In many cases, however, transcript domains are nested within each other, increasing HOX protein diversity toward the caudal end of the axis. In these cases, the functional readout can reflect either the combination of proteins present at a given level (C), or the prominent function of the most caudal protein, through the suppression of other protein functions (D). The *Hox* code model (C) predicts that inactivation of gene M will generate morphological alterations until the most posterior part of the embryo (E), whereas the posterior prevalence model (D) restricts the phenotype to a stripe (F), similar to the situation in (B). post., posterior; ant., anterior.

variation did not induce a stable change in the spatial distribution of transcripts (23). Therefore, one may wonder whether a causal link

exists between temporal gene activation and successive spatial domains, as was initially proposed. It could be that the processes serve

older one. For instance, the design of a novel global enhancer-based colinear process, such as the one in limbs, may have been facilitated by the existence of a gene cluster already at work in a different context. From the perspective of the embryo, any mechanism that will ultimately generate the necessary distribution of *Hox* gene products, within a given axial structure at the right time, should be acceptable. Thus, searching for a single and universal mechanistic explanation for colinearity may be a futile quest. Is this why we have experienced such difficulties in accommodating the large body of experimental data within a well-defined conceptual framework? It might be that the rational explanation of colinearity will not appear as aesthetically pleasing as the process itself. Instead, it may merely reflect a tinkering without any particular underlying logic other than that of the intended result.

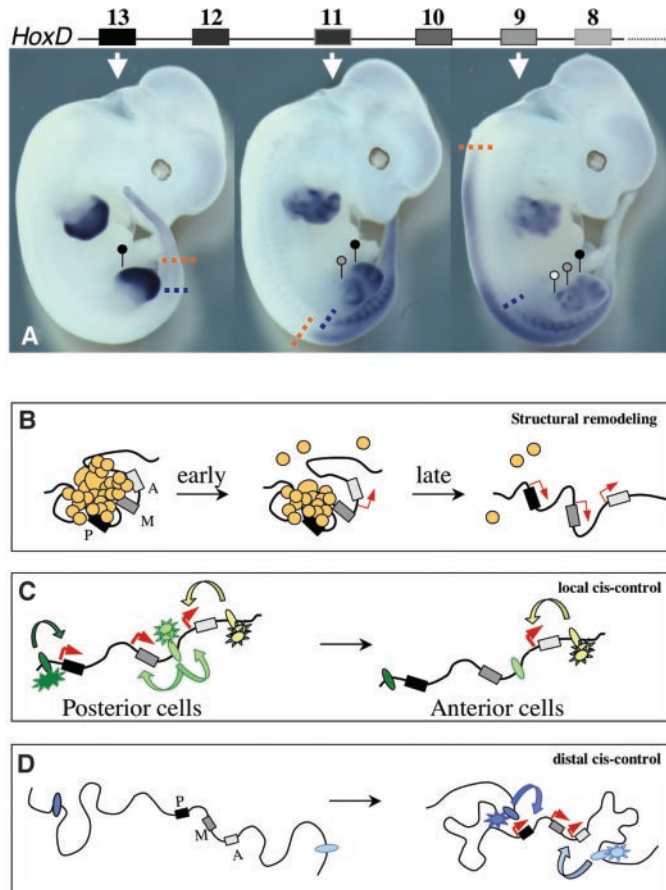


Fig. 2. Models of colinear transcriptional activation along a *Hox* cluster. (A) Expression of three *Hox* genes in developing mouse embryos. Nested domains are visible in both the trunk (dotted lines) and limbs (flags). The anterior (proximal) limits of expression are colinear with gene order on the chromosome. The positions of the transcript domains in the trunk can vary between the neural tube (orange) and paraxial mesoderm (blue), even to a large extent (*Hoxd9*), highlighting tissue-specific mechanisms for colinear gene regulation. [(B) to (D)] Molecular mechanisms proposed for colinear activation. A time-based activation may derive from the gradual accessibility of genes to the transcription machinery (B), through progressive chromatin remodeling. In this model, genes are initially silenced to become progressively accessible. Alternatively, interspersed locally acting cis-regulatory sequences could implement colinear expression (C), for instance by displaying graded affinities (green or yellow arrows) or various specificities for upstream regulators [stars in (C)]. The sharing of these local enhancers between neighboring genes could also participate in directional gene activation. In model (D), global enhancer sequences located outside the cluster regulate the various genes differently, as a result of a distance effect, promoter competition, or sequence-specific recognition. This latter model is implemented during vertebrate limb development. Dark blue and light blue arrows depict regulatory inputs of 5' and 3' remote cis-acting elements, respectively. These models are not mutually exclusive. Instead, it is likely that colinear activation in various structures relies upon distinct mixtures of these mechanisms. Red arrows in (B), (C) and (D) indicate genes that are active transcriptionally.

primarily different goals. Alternatively, the former might be necessary to implement the latter. In a given cellular population, an initial time-based activation could set the stage for a subsequent wave of activation that would depend on cis-acting mechanisms, including auto-regulation and cross-regulation. This second phase could account for the colinear expression of randomly integrated transgenes, which would correctly execute the second phase of activation, the endogenous genes having prepared the required molecular environment.

The mechanisms described here are not exclusive. Instead, they may work in combination with each other, which might explain why a clear picture has not yet fully emerged. It is possible that these various processes reflect different phylogenetic histories. Perhaps an increase in both the number and complexity of structures to be patterned, as well as the design or improvement of original developmental strategies (such as for segmentation or appendages), triggered a novel colinear process and superimposed it on an

References and Notes

1. E. B. Lewis, *Nature* **276**, 565 (1978).
2. S. J. Gaunt et al., *Development* **104**, 169 (1988).
3. A. Graham et al., *Cell* **57**, 367 (1989).
4. D. Duboule, P. Dollé, *EMBO J.* **8**, 1497 (1989).
5. M. Kessel, P. Gruss, *Cell* **67**, 89 (1991).
6. W. McGinnis, R. Krumlauf, *Cell* **68**, 283 (1992).
7. R. S. Mann, M. Affolter, *Curr. Opin. Genet. Dev.* **4**, 423 (1998).
8. G. Vachon et al., *Cell* **71**, 437 (1992).
9. V. Brodu, P. Elstob, A. P. Gould, *Development* **129**, 2957 (2002).
10. I. Lohmann et al., *Cell* **110**, 457 (2002).
11. M. P. Scott, *Cell* **71**, 551 (1992).
12. J. M. Greer et al., *Nature* **403**, 661 (2000).
13. F. Schock et al., *Nature* **405**, 351 (2000).
14. D. Duboule, *Curr. Opin. Genet. Dev.* **1**, 211 (1991).
15. D. Duboule, G. Morata, *Trends Genet.* **10**, 358 (1994).
16. G. S. Horan et al., *Genes Dev.* **9**, 1667 (1995).
17. E. van den Akker et al., *Development* **128**, 1911 (2001).
18. O. Pourquié, *Science*, **301**, 328 (2003).
19. J. Deschamps et al., *Int. J. Dev. Biol.* **43**, 635 (1999).
20. C. T. Kwan et al., *Dev. Biol.* **232**, 176 (2001).
21. F. Zhang et al., *Mech. Dev.* **67**, 49 (1997).
22. H. Marshall et al., *Nature* **370**, 567 (1994).
23. F. van der Hoeven et al., *Cell* **85**, 1025 (1996).
24. T. Kondo, D. Duboule, *Cell* **97**, 407 (1999).
25. M. Kmita et al., *Genes Dev.* **14**, 198 (2000).
26. M. Barna et al., *Dev. Cell* **3**, 499 (2002).
27. A. Gould, *Curr. Opin. Genet. Dev.* **7**, 488 (1997).
28. B. A. Roelen et al., *Mech. Dev.* **119**, 81 (2002).
29. J. P. Liu, E. Laufer, T. M. Jessel, *Neuron* **32**, 997 (2001).
30. S. Bel-Viard, N. Itasaki, R. Krumlauf, *Development* **129**, 5103 (2002).
31. J. Sharpe et al., *EMBO J.* **17**, 1788 (1998).
32. J. Zákány et al., *Cell* **106**, 207 (2001).
33. J. Dubrulle, M. J. McGrew, O. Pourquié, *Cell* **106**, 219 (2001).
34. F. Spitz et al., *Cell* **113**, 405 (2003).
35. M. Kmita et al., *Nature* **420**, 145 (2002).
36. A. C. Burke et al., *Development* **121**, 3163 (1995).
37. S. J. Gaunt, *Int. J. Dev. Biol.* **38**, 549 (1994).
38. We thank S. Cohen, M. McDonald, and J. Cobb for critical reading of the manuscript.