



Article
scientifique

Revue de la
littérature

2013

Published
version

Open
Access

This is the published version of the publication, made available in accordance with the publisher's policy.

Chromatin organization and global regulation of *Hox* gene clusters

Montavon, Thomas; Duboule, Denis

How to cite

MONTAVON, Thomas, DUBOULE, Denis. Chromatin organization and global regulation of *Hox* gene clusters. In: Philosophical transactions - Royal Society. Biological sciences, 2013, vol. 368, n° 1620, p. 20120367. doi: 10.1098/rstb.2012.0367

This publication URL: <https://archive-ouverte.unige.ch/unige:176635>

Publication DOI: [10.1098/rstb.2012.0367](https://doi.org/10.1098/rstb.2012.0367)

Review



Cite this article: Montavon T, Duboule D.
2013 Chromatin organization and global
regulation of *Hox* gene clusters. Phil
Trans R Soc B 368: 20120367.
<http://dx.doi.org/10.1098/rstb.2012.0367>

One contribution of 12 to a Discussion Meeting
Issue 'Regulation from a distance: long-range
regulation of gene expression'.

Subject Areas:

developmental biology, genetics,
molecular biology

Keywords:

embryonic patterning, collinearity, chromatin
architecture, long-range regulation, nuclear
organization

Author for correspondence:

Denis Duboule
e-mail: denis.duboule@unige.ch, denis.duboule@epfl.ch

[†]Present address: Department of Epigenetics,
Max-Planck Institute of Immunobiology and
Epigenetics, 79108 Freiburg, Germany.

Chromatin organization and global regulation of *Hox* gene clusters

Thomas Montavon^{1,†} and Denis Duboule^{1,2}

¹National Research Centre 'Frontiers in Genetics', School of Life Sciences, Ecole Polytechnique Fédérale,
Lausanne, Switzerland

²National Research Centre 'Frontiers in Genetics', Department of Genetics and Evolution, University of Geneva,
Geneva, Switzerland

During development, a properly coordinated expression of *Hox* genes, within their different genomic clusters is critical for patterning the body plans of many animals with a bilateral symmetry. The fascinating correspondence between the topological organization of *Hox* clusters and their transcriptional activation in space and time has served as a paradigm for understanding the relationships between genome structure and function. Here, we review some recent observations, which revealed highly dynamic changes in the structure of chromatin at *Hox* clusters, in parallel with their activation during embryonic development. We discuss the relevance of these findings for our understanding of large-scale gene regulation.

1. Introduction

Understanding the relationship between the structural organization of genomes, on the one hand, and their transcription during development, ageing and pathogenesis, on the other hand, is a major challenge of the post genomic-sequencing era. In recent years, considerable efforts have been devoted to the identification of functional elements within non-coding genomic intervals, as well as to the mapping of chromatin structure and spatial organization of chromosomes in the nuclear space [1–3]. These large-scale approaches tentatively define global relationships between the chromatin structure observed at a particular genomic locus, the three-dimensional organization of this locus and the associated gene activity. For instance, both local chromatin post-translational modifications as well as long-range physical interactions are used to identify candidate transcriptional regulatory elements and their target genes, genome-wide.

The emerging picture suggests that genes playing particularly important roles during embryonic development generally display a highly intricate regulatory organization locally, which involves a large number of control elements. Such elements are usually dispersed within large flanking regions which can be several hundreds of kilobases long and they establish complex patterns of contacts with their target genes. In a broader genomic context, active and inactive gene loci seem to occupy distinct nuclear spaces, which may relate to specialized nuclear compartments such as transcription factories [4–6]. These novel approaches are nevertheless largely descriptive and the future integration of these large datasets within a functional framework, for example by combining them with genetic approaches, will be critical to firmly establish the physiological relevance of these structural parameters.

The coordinated transcription of *Hox* genes within their respective genomic clusters has been a paradigm to study these questions, ever since high-throughput technologies have been available. This is mostly due to the structure of these loci, which display one of the highest concentrations of genes in the genome (in average ten genes within 120 kb) and their enigmatic transcriptional regulation, in response to both local and long-range strategies. Because these genes have a tightly constrained topological organization, they often have been (and still are) used as examples to illustrate the relationships between genome

structure and function. Here, we review recent progress in deciphering the potential mechanisms, which may link chromatin organization to the transcriptional control of *Hox* genes in embryo.

2. Structural and functional organization of *Hox* clusters

Hox genes encode homeodomain transcription factors critical for the proper establishment of regional identities along the main body axis of bilaterian animals. In many species, they are grouped into genomic clusters, which share a conserved structural organization, indicating that a clustered set of genes was probably present in the ancestor of all extant bilateria. This ancestral cluster was split in the *Drosophila* lineage, where eight *Hox* genes are distributed in the *Antennapedia* (*Ant-C*) and *Bithorax* (*BX-C*) complexes; in contrast, the two rounds of whole-genome duplications that accompanied the emergence of vertebrates led to the presence of four paralogous *Hox* clusters in this group, termed *HoxA* to *HoxD*, where a total of 39 genes can be scored in mammals (see [7] for references).

This peculiar genomic organization is closely associated with a regulatory process referred to as ‘collinearity’, i.e. tight correspondence that exists between the order of *Hox* genes within each cluster, on the one hand, and the succession of their expression territories along the anterior–posterior embryonic axis, on the other hand. This property, referred to as spatial collinearity, was originally proposed by studying the genetics of the *Drosophila* BX-C [8]. It was subsequently extended to vertebrates [9–11], which indicated that animals apparently displaying highly divergent morphologies nevertheless rely on the same genetic systems to pattern their body plans. Part of this spatial collinear distribution of *Hox* expression domain is caused by a temporal sequence in their transcriptional activation, which reflects their genomic order, along with the extension of the embryonic axis (‘temporal collinearity’ [12]). A tight control of *Hox* genes transcription, in both space and time, is thus critical for proper development, because variations in the expected HOX proteins combinations at any given anterior–posterior level in the embryo usually lead to alterations in axial patterning [13,14].

Because of this capacity to initiate a coordinated transcriptional response of this series of genes involved in patterning processes, these collinear properties were co-opted several times in the course of vertebrate evolution, along with the emergence of structures displaying some kind of axial specification, such as the gut or the appendages. In such cases, specific *Hox* gene clusters were recruited, usually via the evolution of global regulation involving remote control elements, leading to the global transcription of several *Hox* genes at once ([15]; see below). The mechanistic relationship(s) between *Hox* genes clustering and collinear regulation, however, is still elusive. For example, in some animal species, *Hox* genes display this peculiar expression patterns along the embryonic axis, even though the clustered organization has been lost ([16]; see [7] for a discussion of this issue). Furthermore, *Hox* genes often, yet not always, recapitulate some aspects of their genuine expression specificities in the developing trunk, when isolated from their endogenous cluster and integrated at random genomic locations as transgenes, indicating that transcription units carry some of the requested regulatory information [17,18].

Despite these observations, targeted modifications of *Hox* clusters have shown that changing the relative position of a gene within its cluster has a critical impact on several aspects

of its transcriptional regulation, such as the timing of expression along the main axis, as well as transcript distribution along secondary axes, such as the growing limbs [18–22]. Therefore, global regulatory influences act over *Hox* clusters as a whole, on the top of more local, gene-specific proximal controls. A transition in chromatin structure, from an initially repressed state to a configuration progressively open for transcription, was proposed to accompany this coordinated activation early on [23]. We discuss below recent observations, which suggest that both local histone modifications and changes in the higher-order organization of chromatin are involved in the transcriptional control of *Hox* clusters.

3. Epigenetic control by Polycomb and Trithorax complexes

The epigenetic regulation of *Hox* gene clusters seems to rely, mostly, on the activities of protein complexes encoded by *Polycomb* (*PcG*) and *Trithorax* (*TrxG*) group genes. During early *Drosophila* embryogenesis, maternally supplied transcription factors define the spatial patterns of *Hox* genes’ activity [24]. These patterns are maintained at later developmental stages through the action of *PcG* genes, required for stable *Hox* gene repression. *TrxG* genes, on the other hand, counteract *PcG* silencing and keep *Hox* genes expressed in the appropriate domains [25].

PcG and *TrxG* gene products are found in multi-protein complexes, which mediate the post-translational modification of histone tails and thus affect chromatin structure. *PcG*-mediated silencing relies on the combined actions of the Polycomb repressive complexes 1 and 2 (PRC1 and PRC2). PRC2 tri-methylates lysine 27 of histone H3 tail (H3K27), a mark that is tightly associated with gene repression [26–29]. PRC1 is recruited to H3K27me3 [26,29–31] and contains the Ring1B E3 ubiquitin ligase that triggers the ubiquitylation of H2A at lysine 119 [32,33]. By contrast, *TrxG* complexes are responsible for the tri-methylation of histone H3 tail at lysine 4 (H3K4me3), a chromatin mark associated with transcriptional activation [34].

In mouse and human cultured cells, H3K27me3 and H3K4me3 predominantly decorate silent or active promoters, respectively [35–38]. Surprisingly, a significant number of silent loci, including *Hox* clusters, display both H3K4 and H3K27 tri-methylation in pluripotent embryonic stem (ES) cells, a chromatin signature referred to as ‘bivalent domains’ [39,40]. Differentiation of ES cells leads to a resolution of these bivalency into either active or repressive chromatin marks, suggesting that such bivalent domains might label genes that are kept silent, but ‘poised’ for rapid activation upon lineage commitment. Such plasticity is at odds with the classical view of *PcG* and *TrxG* function as mediators of a stable memory of epigenetic states [41], but is supported by the identification of H3K27 lysine de-methylases recruited to gene promoters upon transcriptional activation [42–45].

Similar to their *Drosophila* orthologues, *PcG* and *trxG* genes are required for vertebrate embryonic development, and mutations in these genes lead to a deregulation of *Hox* genes, amongst other defects (reviewed in [46,47]). Recently, a progressive loss of H3K27me3, concomitantly to a gain of H3K4me3, was observed together with the sequential activation of *Hoxd* genes during the extension of the main body axis in the mouse *in vivo* ([48]; figure 1a). Transcriptional activation of *Hoxd* genes thus occurs within the region of transition between these

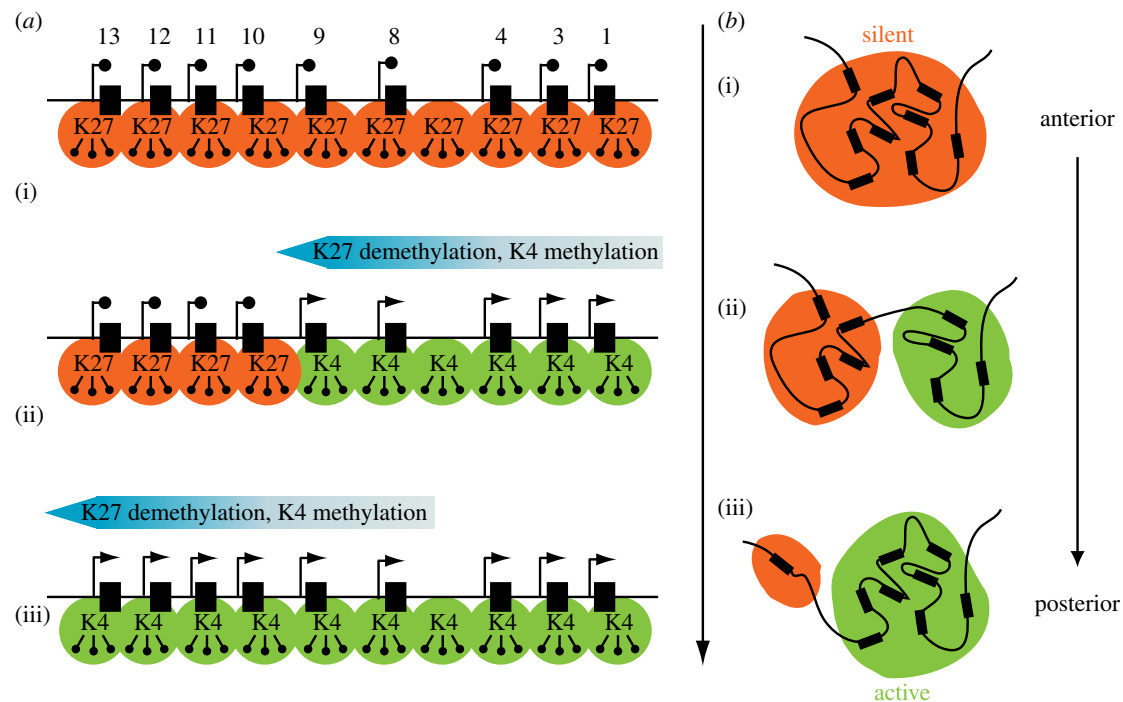


Figure 1. Collinearity during trunk extension and chromatin dynamics at *Hox* clusters. Expression of *Hox* gene along the anterior-to-posterior (AP) embryonic axis is collinear with gene order within the cluster. (a) During axial extension, the sequential onset of *Hox* gene transcriptional activation is accompanied by a transition in histone modifications over the gene cluster. In ES cells (i), the whole cluster is labelled with H3K27me3 (orange), a mark associated with Polycomb-mediated silencing. In the developing embryo, this mark is progressively erased and replaced by H3K4me3 (green), concomitantly with gene activation. (b) Active and silent *Hox* loci segregate into distinct spatial compartments along the AP axis. In embryonic tissues where the whole cluster is repressed, such as the forebrain (i), *Hox* clusters form a compact three-dimensional structure. In regions where subsets of *Hox* genes are expressed (anterior trunk, ii), active and silent genes segregate in distinct compartments, labelled with either H3K27me3 (silent compartment) or H3K4me3 (active compartment). In posterior embryonic regions (iii), most genes are transcribed and participate in the active compartment.

two epigenetic states, in a window that shifts along with time, from one extremity of the cluster to the other. The analysis of embryos carrying an engineered split *HoxD* cluster further revealed that clustering is not necessary for the initial deposition of H3K27me3, yet it is required for a fully coordinated transition in histone modifications [48].

The mechanisms recruiting PcG proteins to their target loci are not fully understood. In *Drosophila*, relatively short sequences termed Polycomb response elements (PREs) seem to be both necessary and sufficient for PcG recruitment and gene silencing, although these sites cannot be defined by a consensus DNA sequence [25]. The situation is more complex in vertebrates, where classical PREs—as defined in flies—remain to be identified. For instance, an element from the human *HOXD* cluster was recently reported to bind PcG proteins and to induce repression of a reporter gene [49]. However, the deletion of the murine orthologous sequence did not cause any dramatic change in *Hoxd* gene regulation [50], suggesting that it is not critical for PcG recruitment or, alternatively, that it is part of a robust mechanism with high compensatory capacities. Interestingly, CG-rich sequences, which are particularly abundant within *Hox* clusters, were proposed as a hallmark of PcG target promoters, at least in ES cells [51,52]. Silencing at these particular loci might thus involve the cooperative activity of multiple DNA elements.

4. Higher-order chromatin organization

The coordinated control of these dynamic epigenetic states might be facilitated by the spatial compartmentalization of

Hox clusters. Changes in higher-order chromatin organization of *Hox* clusters were first observed by microscopy approaches such as fluorescent *in situ* hybridization (FISH), which allow the visualization of specific loci within the nucleus. Differentiation of ES cells after retinoic acid treatment leads to a decompaction of the *HoxB* cluster that parallels its global transcriptional activation [53]. Furthermore, *Hox* clusters are also less compact in those embryonic territories where *Hox* genes are probably active, than in silent regions [54]. However, the resolution of this approach did not allow the stepwise, collinear transition to be documented.

Also, owing to its limited resolution, FISH cannot currently provide a precise mapping of the three-dimensional organization of *Hox* loci, in either their active or inactive states. The development of the chromosome conformation capture (3C) technique, which provides an estimation of the average frequency of specific DNA–DNA ‘contacts’ (or proximity, see below) within a cell population, has greatly helped to overcome this limitation. While early 3C analyses only addressed interactions between a limited number of sequences, variant approaches such as 4C (circular 3C), 5C (3C carbon copy) or Hi-C generate rather unbiased datasets, where the mapping of all sequences contacting a locus of interest, or the analysis of mutual interactions between a large number of pre-determined sites—or even within an entire genome, can be produced [55]. Using these approaches, changes in the three-dimensional organization of *Hox* clusters were detected upon cell differentiation in culture or between cell lines derived from various regions of the body [56–58].

In dissected murine embryonic tissues, this configuration is tightly associated with the transcriptional activity of the gene

clusters. In tissues where *Hox* genes are silent, such as the developing forebrain, *Hox* gene clusters form distinct spatial structures, as defined by widespread interactions between the various gene loci, within each cluster. In contrast, in regions where distinct subsets of *Hox* genes are transcribed, active and repressed genes segregate into distinct compartments, labelled by different chromatin marks [59,60]. In anterior regions of the main embryonic axis, expressed *Hox* genes appear to cluster together within a H3K4me3 decorated domain, whereas silent genes form a distinct compartment marked with H3K27me3. In more posterior areas, most *Hox* genes are found in the active compartment, suggesting a dynamic reorganization of the three-dimensional micro-architecture of these clusters ([59]; figure 1b).

This spatial separation between the active and silent parts of *Hox* clusters might in part reflect the PcG-mediated compaction of repressed loci, as shown *in vitro* [61]. Interestingly, the loss of either *Eed* (a protein member of PRC2) or *Ring1B* (a protein member of PRC1) leads to the decompaction of *Hox* clusters in ES cells [62]. In *Drosophila*, Polycomb targets segregate into discrete nuclear foci termed Polycomb bodies [63]. Despite a genomic distance of nearly 10 megabases (Mb), the *Ant-C* and *BX-C* complexes contact each other in a PcG dependent manner, in embryonic regions where both loci are silent [64].

Whether these patterns of contacts participate in the mechanism controlling *Hox* gene collinearity or, instead, merely reflect a general tendency of different chromatin segments to segregate into distinct nuclear domains depending on their modifications, remains to be established. However, the presence of separate spatial domains, as seen by using 4C, makes a biochemical artefact induced by cross-linking unlikely. In such a case, chromatin domains labelled either by H3K27me3 or by H3K4me3 would indeed be expected to cross-link together. Instead, a physical separation between active and repressed *Hox* subsets might participate in a tighter control of the sequential activation of *Hox* clusters, by isolating posterior genes from early/anterior activating influences.

5. Long-range control and regulatory archipelagos

The type of collinear regulation described above follows a mechanism acting purely in *cis*, and is shared by all paralogous *Hox* clusters. Therefore, it may represent an ancestral mechanism at work in those animals that activate their *Hox* genes in a precise time sequence (see above). In addition, in the vertebrate lineage, specific *Hox* clusters have evolved additional patterning functions [23,65,66], which required the emergence of novel regulatory modalities. These novel regulatory specificities often rely upon enhancer elements located outside the gene clusters, at a distance, likely to prevent deleterious effects of evolving additional regulatory elements within the gene clusters themselves, where the *cis*-regulatory sequences necessary to implement the ancestral collinear regulation along the anterior–posterior axis are mostly located [67].

Such potent enhancers have indeed been described and can act over large distances. They sometimes also affect unrelated, bystander transcription units, thus defining the concept of genomic ‘regulatory landscapes’ [68]. The emergence of these long-range controls, acting on several neighbouring *Hox* genes in a coordinated manner may have provided a

selective pressure for the gene clusters to consolidate their structural organization in vertebrates [7].

The transcription of *Hoxd* genes in developing limbs is the clearest example of such kind of acquired distal regulations around *Hox* gene clusters, as it has been investigated in some detail using both genetic and biochemical approaches. ‘Posterior’ *Hoxd* genes (from *Hoxd9* to *Hoxd13*) are required for the patterning of both proximal (arm and forearm or leg and lower leg) and distal (hands and feet) limb segments [69]. Their expression in growing limb buds follows two independent phases [21,70], controlled by distinct regulatory elements located on either side of the gene cluster [71]. In a first phase, corresponding to the future proximal limb segment, *Hoxd* genes are activated in a time sequence, which generates a nested pattern of expression along the limb anterior to posterior axis. While regulatory elements controlling this early activation are not yet reported, some genomic rearrangements suggest that they may be located on the telomeric side of the gene cluster, on chromosome 2 ([68,71]; figure 2a). Multiple highly conserved non-coding DNA sequences can be found there and thus represent potential candidates [72].

Subsequently, the transcription of posterior *Hoxd* genes (*Hoxd13* to *Hoxd10*) is activated at the distal extremities of the growing limbs, in a territory corresponding to presumptive digits. This ‘late phase’ is controlled by elements located centromeric from *Hoxd13* [73]. The mapping of the three-dimensional organization of the cluster revealed that active genes establish numerous long-range interactions with sequences dispersed over an 800 kb interval overlapping a gene desert located centromeric to the cluster [60]; figure 2a). These sites are grouped into ‘islands’ that are labelled with histone marks usually associated with enhancer elements, such as the monomethylation of H3K4 (H3K4me1) and the acetylation of H3K27 (H3K27Ac) [74–76]. The various islands contact each other, as well as the active part of the *HoxD* cluster, suggesting that they form an active conformation in developing digits.

In contrast, silent genes display the opposite profile of interactions, mostly involving the telomeric gene desert, which is devoid of active histone marks [60]. Therefore, the compartmentalization of *Hox* clusters in active and silent domains, in growing limb cells, also involves differential association with surrounding sequences. In the developing forebrain, where the entire *HoxD* cluster is repressed, some of these long-range interactions with distant islands are observed too, yet in this case the islands display a distinct epigenetic signature, involving limited levels of H3K4me1 and no H3K27Ac [60]. This indicates that the transition between a ground-state or poised structure towards a fully active configuration involves only a partial reorganization of the locus micro-architecture, associated with chromatin modifications over the regulatory elements (figure 2b). Interestingly, such megabase-scale regions of chromatin interactions, or ‘topological domains’ are widespread in vertebrate genomes, and appear stable when comparing different cell lines [77], suggesting that ‘poised’ regulatory conformations are not restricted to *Hox* clusters.

However, when using chromosome conformation capture approaches, it is important to keep in mind a few methodological aspects. First, the data obtained give an average of the contacts established within the cellular population under scrutiny and hence may not reflect the real situation within one particular cell, at a particular time. While this approach does not tell about the dynamics of the interactions, the subsequent

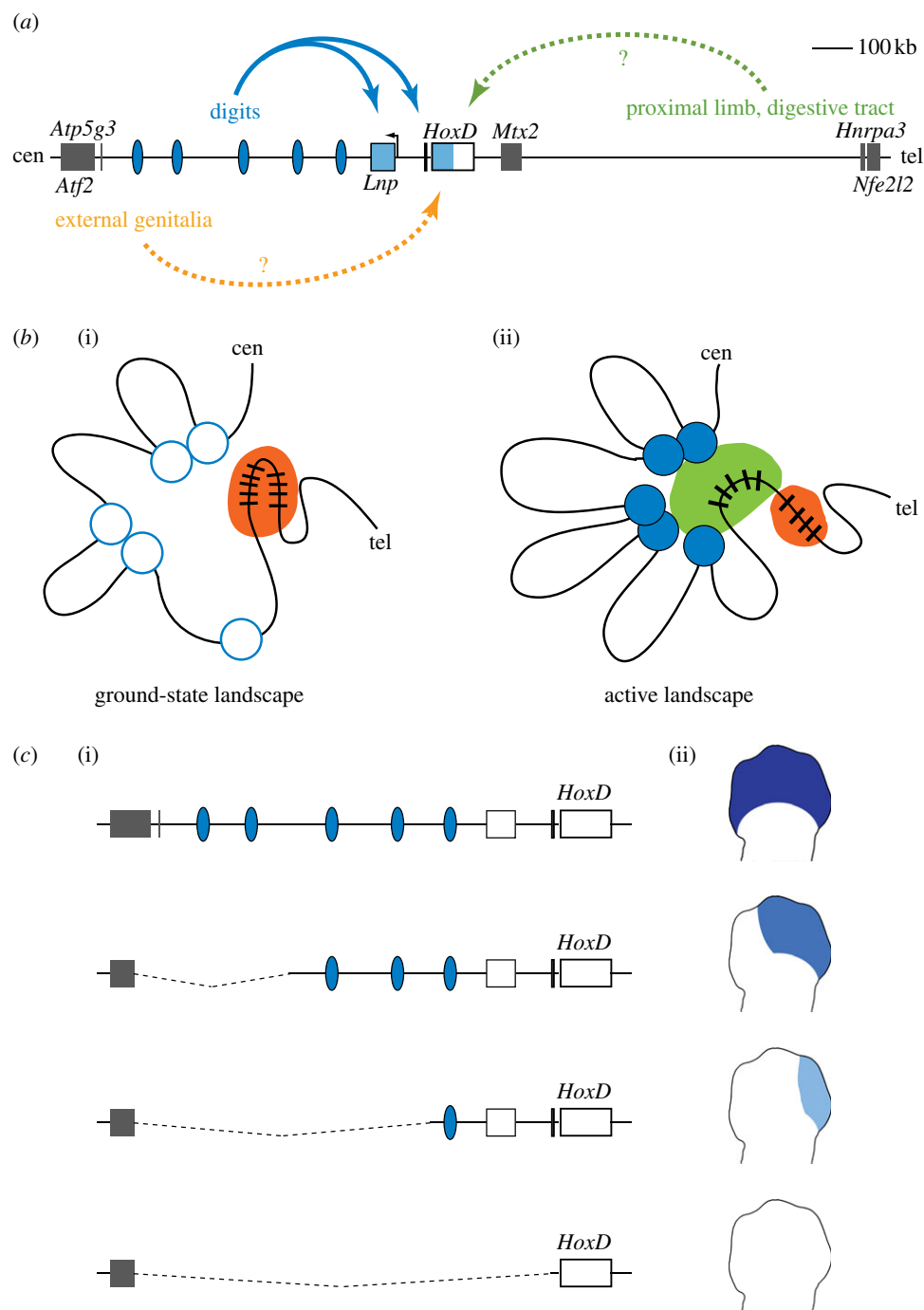


Figure 2. Long-range control and regulatory archipelagos. The coordinated transcription of *Hox* genes in different embryonic territories relies on remote regulatory elements located on either sides of the cluster. (a) Map of the *HoxD* locus and its flanking centromeric (cen) and telomeric (tel) conserved gene deserts. Multiple regulatory islands (blue ovals) participate in *Hoxd13–Hoxd10* regulation in developing digits (blue arrow). *Hoxd* gene activation in other embryonic structures also relies on long-range controls, yet the corresponding regulatory elements have not yet been identified (dashed arrows). (b) Spatial conformation of the locus. (i) In the silent state, a ground-state structure is formed, involving contacts between a subset of the regulatory elements and the *HoxD* cluster. (ii) In digits, additional contacts are formed, leading to a fully active conformation paralleled with histone modifications at the regulatory elements, and leading to *Hox* genes transcriptional activation. (c) Robustness of regulatory archipelagos. Different genetic configurations of the locus are shown in (i), with a scheme of the resulting *Hox* gene expression in developing limbs in (ii). The wild-type situation is depicted on top, and serial deletions within the archipelago are indicated below. Deleting subsets of the regulatory islands leads to partial downregulation of *Hoxd* genes in distal limbs, and only a full deletion of the regulatory interval fully abolishes transcription.

use of contact points as baits themselves, may at least give some hints about the complexity of the interactions (for example by documenting interactions amongst various enhancers sequences, rather than between enhancers and the target promoter). In any case, the cellular homogeneity of the sample and its physiological relevance must be carefully assessed beforehand to help interpret the interaction profiles.

Second, the cross-linking step will bring together DNA pieces that can be quite far from one another, provided

they are engaged in large DNA–protein complexes. In this respect, the notion of ‘contact’ can be understood both as a direct and close interaction between two pieces of DNA, for example between an enhancer sequence and its target promoter *via* looping, but also as a more diffuse proximity, where enhancers could act as platforms to recruit factors such as to increase their concentration around particular genomic loci. In this context, investigating these dynamic patterns of contacts at the single cell level and at different stages of

limb development should bring more insight into the establishment of this complex conformation. Interestingly, recent observations indicate that the frequencies of some of these interactions, as documented by FISH, display regional differences along the limb anterior–posterior axis [78].

Several of the regulatory islands, as defined by the interaction profile, could elicit a digit-specific transcriptional activation when isolated from the locus and tested in transgenic assays [60,68,79], further validating the 4C approach as a tool to isolate long-range acting enhancers. A genetic dissection of the 800 kb large genomic interval centromeric to *Hoxd13* indicated that these various elements contribute in a partially redundant fashion to the transcriptional activation of *Hoxd* genes ([60]; figure 2c). Such a dispersed enhancer system, somehow reminiscent of shadow enhancers described in *Drosophila* [80–82], was referred to as a regulatory archipelago, i.e. a collection of regulatory islands, located nearby one another and all devoted to a collective function though with slightly different contributions of the diverse elements. Such a complex regulatory system may both ensure a robust transcriptional response of the target genes, and confer some regulatory flexibility during the development and evolution of distal limbs. Genetic alterations within this interval could indeed impact on the growth and patterning of digits, and thus contribute to the diversity of digit morphologies amongst various tetrapods.

6. Concluding remarks and outlook

The correspondence between the structural organization of *Hox* clusters and their transcriptional control in space and time has fascinated biologists for decades. While the underlying mechanisms remain to be fully elucidated, recent progress in deciphering the epigenetic status of these loci in developing embryos, as well as their three-dimensional organization in the nuclear space, have provided substantial information.

Future investigations will involve the identification of upstream factors and signals controlling the recruitment of chromatin-modifying activities to *Hox* loci, as well as the definition of their spatial organization during development. Considerable interest has been focused recently on the involvement of non-coding RNA or structural proteins, such as CTCF, in these processes [56,58,83,84]. For instance, CTCF binding sites are enriched at the boundaries of topological domains within the *HoxA* cluster, as well as at other genomic loci [77]. Yet, a critical impact of these various factors on *Hox*

clusters regulation has not been confirmed by genetic analysis, so far [85,86]. It will also be critical to isolate the factors involved in the transcriptional activation of *Hox* genes, to understand the nature of the directionality of the ancestral process, during trunk extension.

Finally, the limb regulatory archipelago may be a paradigm to understand the emergence and evolution of such regulatory systems. Tetrapods indeed present a great variety of distal morphologies, in their limbs, and the analysis of the accompanying regulations will be informative in this context. Also, developing limbs are accessible structures, in the developing embryo, with rather well defined cell types and expression domains, unlike cultured cell systems, which may have little physiological relevance in this context, and where genetic approaches are not available.

Tracing back the emergence of novel regulatory landscapes in vertebrates may also bring some insights to our understanding of genome evolution. For instance, the acquisition of a late and distal phase of *Hoxd* gene activation in limbs was critical to the evolution of tetrapod digits, and this expression specificity does not seem to have a counterpart in fishes [87]. Likewise, the activation of *Hoxd* genes in developing external genitalia and digits follows highly similar patterns, which led to the hypothesis of a shared and ancestral regulation in both structures [22,88]. Addressing the regulatory potential of the syntenic regions of different species, as well as comparing the mechanisms controlling *Hox* clusters expression in different embryonic structures, should help reconstruct the evolutionary history of these regulations.

Similar concepts are likely to apply to other regulatory contexts. For instance, the involvement of gene deserts in long-range transcriptional control is not a specificity of *Hox* clusters, and regulatory elements were identified within gene deserts flanking other genes displaying intricate expression patterns [89–92]. The prevalence of regulatory archipelagos similar to that described at the *HoxD* locus nevertheless remains to be established and further efforts in mapping the spatial organization of genomes, combined with large-scale genetic approaches will be necessary.

We thank Natalia Soshnikova and colleagues from the Duboule laboratories for discussions. We acknowledge financial support from the Ecole Polytechnique Federale de Lausanne (EPFL), the University of Geneva, the Swiss National Research Council (SNF), the European Research Council (ERC) and the FP7 EU programme IDEAL.

References

1. Zhou VW, Goren A, Bernstein BE. 2011 Charting histone modifications and the functional organization of mammalian genomes. *Nat. Rev. Genet.* **12**, 7–18. (doi:10.1038/nrg2905)
2. Bernstein BE, Birney E, Dunham I, Green ED, Gunter C, Snyder M. 2012 An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74. (doi:10.1038/nature11247)
3. de Wit E, de Laat W. 2012 A decade of 3C technologies: insights into nuclear organization. *Genes Dev.* **26**, 11–24. (doi:10.1101/gad.179804.111)
4. Lieberman-Aiden E *et al.* 2009 Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293. (doi:10.1126/science.1181369)
5. Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Parrinello H, Tanay A, Cavalli G. 2012 Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148**, 458–472. (doi:10.1016/j.cell.2012.01.010)
6. Edelman LB, Fraser P. 2012 Transcription factories: genetic programming in three dimensions. *Curr. Opin. Genet. Dev.* **22**, 110–114. (doi:10.1016/j.gde.2012.01.010)
7. Duboule D. 2007 The rise and fall of Hox gene clusters. *Development* **134**, 2549–2560. (doi:10.1242/dev.001065)
8. Lewis EB. 1978 A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565–570. (doi:10.1038/276565a0)
9. Gaunt SJ, Sharpe PT, Duboule D. 1988 Spatially restricted domains of homeo-gene transcripts in mouse embryos: relation to a segmented body plan. *Development* **104**, 169–170.

10. Graham A, Papalopulu N, Krumlauf R. 1989 The murine and *Drosophila* homeobox gene complexes have common features of organization and expression. *Cell* **57**, 367–378. (doi:10.1016/0092-8674(89)90912-4)
11. Duboule D, Dolle P. 1989 The structural and functional organization of the murine HOX gene family resembles that of *Drosophila* homeotic genes. *EMBO J.* **8**, 1497–1505.
12. Izpisua Belmonte JC, Falkenstein H, Dollé P, Renucci A, Duboule D. 1991 Murine genes related to the *Drosophila* AbdB homeotic genes are sequentially expressed during development of the posterior part of the body. *EMBO J.* **10**, 2279–2289.
13. Mallo M, Wellik DM, Deschamps J. 2010 Hox genes and regional patterning of the vertebrate body plan. *Dev. Biol.* **344**, 7–15. (doi:10.1016/j.ydbio.2010.04.024)
14. Young T *et al.* 2009 Cdx and Hox genes differentially regulate posterior axial growth in mammalian embryos. *Dev. Cell* **17**, 516–526. (doi:10.1016/j.devcel.2009.08.010)
15. Tschopp P, Duboule D. 2011 A genetic approach to the transcriptional regulation of hox gene clusters. *Annu. Rev. Genet.* **45**, 145–166. (doi:10.1146/annurev-genet-102209-163429)
16. Seo H-C *et al.* 2004 Hox cluster disintegration with persistent anteroposterior order of expression in *Oikopleura dioica*. *Nature* **431**, 67–71. (doi:10.1038/nature02709)
17. Tschopp P, Christen AJ, Duboule D. 2012 Bimodal control of Hoxd gene transcription in the spinal cord defines two regulatory subclusters. *Development* **139**, 929–939. (doi:10.1242/dev.076794)
18. Tschopp P, Tarchini B, Spitz F, Zakany J, Duboule D. 2009 Uncoupling time and space in the collinear regulation of Hox genes. *PLoS Genet.* **5**, e1000398. (doi:10.1371/journal.pgen.1000398)
19. van der Hoeven F, Zakany J, Duboule D. 1996 Gene transpositions in the HoxD complex reveal a hierarchy of regulatory controls. *Cell* **85**, 1025–1035. (doi:10.1016/S0092-8674(00)81303-3)
20. Kmita M, Fraudeau N, Héroult Y, Duboule D. 2002 Serial deletions and duplications suggest a mechanism for the collinearity of Hoxd genes in limbs. *Nature* **420**, 145–150. (doi:10.1038/nature01189)
21. Tarchini B, Duboule D. 2006 Control of Hoxd genes' collinearity during early limb development. *Dev. Cell* **10**, 93–103. (doi:10.1016/j.devcel.2005.11.014)
22. Montavon T, Le Garrec J-F, Kerszberg M, Duboule D. 2008 Modeling Hox gene regulation in digits: reverse collinearity and the molecular origin of thumbness. *Genes Dev.* **22**, 346–359. (doi:10.1101/gad.1631708)
23. Dollé P, Izpisua Belmonte JC, Falkenstein H, Renucci A, Duboule D. 1989 Coordinate expression of the murine Hox-5 complex homoeobox-containing genes during limb pattern formation. *Nature* **342**, 767–772. (doi:10.1038/342767a0)
24. Maeda RK, Karch F. 2009 The bithorax complex of *Drosophila* an exceptional Hox cluster. *Curr. Top. Dev. Biol.* **88**, 1–33. (doi:10.1016/S0070-2153(09)88001-0)
25. Schwartz YB, Pirrotta V. 2007 Polycomb silencing mechanisms and the management of genomic programmes. *Nat. Rev. Genet.* **8**, 9–22. (doi:10.1038/nrg1981)
26. Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y. 2002 Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298**, 1039–1043. (doi:10.1126/science.1076997)
27. Czermin B, Melfi R, McCabe D, Seitz V, Imhof A, Pirrotta V. 2002 *Drosophila* enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* **111**, 185–196. (doi:10.1016/S0092-8674(02)00975-3)
28. Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P, Reinberg D. 2002 Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev.* **16**, 2893–2905. (doi:10.1101/gad.1035902)
29. Muller J *et al.* 2002 Histone methyltransferase activity of a *Drosophila* Polycomb group repressor complex. *Cell* **111**, 197–208. (doi:10.1016/S0092-8674(02)00976-5)
30. Fischle W, Wang Y, Jacobs SA, Kim Y, Allis CD, Khorasanizadeh S. 2003 Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes Dev.* **17**, 1870–1881. (doi:10.1101/gad.1110503)
31. Min J, Zhang Y, Xu RM. 2003 Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. *Genes Dev.* **17**, 1823–1828. (doi:10.1101/gad.269603)
32. de Napoles M *et al.* 2004 Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev. Cell* **7**, 663–676. (doi:10.1016/j.devcel.2004.10.005)
33. Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS, Zhang Y. 2004 Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431**, 873–878. (doi:10.1038/nature02985)
34. Ruthenburg AJ, Allis CD, Wysocka J. 2007 Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Mol. Cell* **25**, 15–30. (doi:10.1016/j.molcel.2006.12.014)
35. Bracken AP, Dietrich N, Pasini D, Hansen KH, Helin K. 2006 Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev.* **20**, 1123–1136. (doi:10.1101/gad.381706)
36. Lee TI *et al.* 2006 Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* **125**, 301–313. (doi:10.1016/j.cell.2006.02.043)
37. Boyer LA *et al.* 2006 Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**, 349–353. (doi:10.1038/nature04733)
38. Bernstein BE *et al.* 2005 Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120**, 169–181. (doi:10.1016/j.cell.2005.01.001)
39. Bernstein BE *et al.* 2006 A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315–326. (doi:10.1016/j.cell.2006.02.041)
40. Azuara V *et al.* 2006 Chromatin signatures of pluripotent cell lines. *Nat. Cell Biol.* **8**, 532–538. (doi:10.1038/ncb1403)
41. Ringrose L, Paro R. 2007 Polycomb/Trithorax response elements and epigenetic memory of cell identity. *Development* **134**, 223–232. (doi:10.1242/dev.02723)
42. Lan F *et al.* 2007 A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* **449**, 689–694. (doi:10.1038/nature06192)
43. De Santa F, Totaro MG, Prosperini E, Notarbartolo S, Testa G, Natoli G. 2007 The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of Polycomb-mediated gene silencing. *Cell* **130**, 1083–1094. (doi:10.1016/j.cell.2007.08.019)
44. Lee MG, Villa R, Trojer P, Norman J, Yan KP, Reinberg D, Di Croce L, Shiekhattar R. 2007 Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. *Science* **318**, 447–450. (doi:10.1126/science.1149042)
45. Mansour AA *et al.* 2012 The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. *Nature* **488**, 409–413. (doi:10.1038/nature11272)
46. Soshnikova N. 2011 Dynamics of Polycomb and Trithorax activities during development. *Birth Defects Res. A Clin. Mol. Teratol.* **91**, 781–787. (doi:10.1002/bdra.20774)
47. Surface LE, Thornton SR, Boyer LA. 2010 Polycomb group proteins set the stage for early lineage commitment. *Cell Stem Cell* **7**, 288–298. (doi:10.1016/j.stem.2010.08.004)
48. Soshnikova N, Duboule D. 2009 Epigenetic temporal control of mouse Hox genes *in vivo*. *Science* **324**, 1320–1323. (doi:10.1126/science.1171468)
49. Woo CJ, Kharchenko PV, Daheron L, Park PJ, Kingston RE. 2010 A region of the human HOXD cluster that confers Polycomb-group responsiveness. *Cell* **140**, 99–110. (doi:10.1016/j.cell.2009.12.022)
50. Beckers J, Duboule D. 1998 Genetic analysis of a conserved sequence in the HoxD complex: regulatory redundancy or limitations of the transgenic approach? *Dev. Dyn.* **213**, 1–11. (doi:10.1002/(SICI)1097-0177(199809)213:1<1::AID-AJA1>3.0.CO;2-L)
51. Ku M *et al.* 2008 Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet.* **4**, e1000242. (doi:10.1371/journal.pgen.1000242)
52. Mendenhall EM, Koche RP, Truong T, Zhou VW, Issac B, Chi AS, Ku M, Bernstein BE. 2010 GC-rich sequence elements recruit PRC2 in mammalian ES cells. *PLoS Genet.* **6**, e1001244. (doi:10.1371/journal.pgen.1001244)
53. Chambeyron S, Bickmore WA. 2004 Chromatin decondensation and nuclear reorganization of the

- HoxB locus upon induction of transcription. *Genes Dev.* **18**, 1119–1130. (doi:10.1101/gad.292104)
54. Chambeyron S, Da Silva NR, Lawson KA, Bickmore WA. 2005 Nuclear re-organisation of the Hoxb complex during mouse embryonic development. *Development* **132**, 2215–2223. (doi:10.1242/dev.01813)
 55. van Steensel B, Dekker J. 2010 Genomics tools for unraveling chromosome architecture. *Nat. Biotechnol.* **28**, 1089–1095. (doi:10.1038/nbt.1680)
 56. Ferraiuolo MA, Rousseau M, Miyamoto C, Shenker S, Wang XQ, Nadler M, Blanchette M, Dostie J. 2010 The three-dimensional architecture of Hox cluster silencing. *Nucleic Acids Res.* **38**, 7472–7484. (doi:10.1093/nar/gkq644)
 57. Fraser J, Rousseau M, Shenker S, Ferraiuolo MA, Hayashizaki Y, Blanchette M, Dostie J. 2009 Chromatin conformation signatures of cellular differentiation. *Genome Biol.* **10**, R37. (doi:10.1186/gb-2009-10-4-r37)
 58. Wang KC *et al.* 2011 A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **472**, 120–124. (doi:10.1038/nature09819)
 59. Noordermeer D, Leleu M, Splinter E, Rougemont J, De Laat W, Duboule D. 2011 The dynamic architecture of Hox gene clusters. *Science* **334**, 222–225. (doi:10.1126/science.1207194)
 60. Montavon T, Soshnikova N, Mascres B, Joye E, Thevenet L, Splinter E, de Laat W, Spitz F, Duboule D. 2011 A regulatory archipelago controls Hox genes transcription in digits. *Cell* **147**, 1132–1145. (doi:10.1016/j.cell.2011.10.023)
 61. Grau DJ, Chapman BA, Garlick JD, Borowsky M, Francis NJ, Kingston RE. 2011 Compaction of chromatin by diverse Polycomb group proteins requires localized regions of high charge. *Genes Dev.* **25**, 2210–2221. (doi:10.1101/gad.17288211)
 62. Eskeland R *et al.* 2010 Ring1B compacts chromatin structure and represses gene expression independent of histone ubiquitination. *Mol. Cell* **38**, 452–464. (doi:10.1016/j.molcel.2010.02.032)
 63. Lanzaolo C, Roue V, Dekker J, Bantignies F, Orlando V. 2007 Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. *Nat. Cell Biol.* **9**, 1167–1174. (doi:10.1038/ncb1637)
 64. Bantignies F, Roue V, Comet I, Leblanc B, Schuettengruber B, Bonnet J, Tixier V, Mas A, Cavalli G. 2011 Polycomb-dependent regulatory contacts between distant Hox loci in *Drosophila*. *Cell* **144**, 214–226. (doi:10.1016/j.cell.2010.12.026)
 65. Godwin AR, Capecchi MR. 1998 *Hoxc13* mutant mice lack external hair. *Genes Dev.* **12**, 11–20. (doi:10.1101/gad.12.1.11)
 66. Di-Poi N, Zakany J, Duboule D. 2007 Distinct roles and regulations for HoxD genes in metanephric kidney development. *PLoS Genet.* **3**, e232. (doi:10.1371/journal.pgen.0030232)
 67. Spitz F, Gonzalez F, Peichel C, Vogt TF, Duboule D, Zakany J. 2001 Large scale transgenic and cluster deletion analysis of the HoxD complex separate an ancestral regulatory module from evolutionary innovations. *Genes Dev.* **15**, 2209–2214. (doi:10.1101/gad.205701)
 68. Spitz F, Gonzalez F, Duboule D. 2003 A global control region defines a chromosomal regulatory landscape containing the HoxD cluster. *Cell* **113**, 405–417. (doi:10.1016/S0092-8674(03)00310-6)
 69. Zakany J, Duboule D. 2007 The role of Hox genes during vertebrate limb development. *Curr. Opin. Genet. Dev.* **17**, 359–366. (doi:10.1016/j.gde.2007.05.011)
 70. Nelson CE *et al.* 1996 Analysis of Hox gene expression in the chick limb bud. *Development* **122**, 1449–1466.
 71. Spitz F, Herkenne C, Morris MA, Duboule D. 2005 Inversion-induced disruption of the Hoxd cluster leads to the partition of regulatory landscapes. *Nat. Genet.* **37**, 889–893. (doi:10.1038/ng1597)
 72. Lee AP, Koh EG, Tay A, Brenner S, Venkatesh B. 2006 Highly conserved syntenic blocks at the vertebrate Hox loci and conserved regulatory elements within and outside Hox gene clusters. *Proc. Natl Acad. Sci. USA* **103**, 6994–6999. (doi:10.1073/pnas.0601492103)
 73. Tschopp P, Duboule D. 2011 A regulatory ‘landscape effect’ over the HoxD cluster. *Dev. Biol.* **351**, 288–296. (doi:10.1016/j.ydbio.2010.12.034)
 74. Heintzman ND *et al.* 2007 Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* **39**, 311–318. (doi:10.1038/ng1966)
 75. Creighton MP *et al.* 2010 Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl Acad. Sci. USA* **107**, 21 931–21 936. (doi:10.1073/pnas.1016071107)
 76. Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. 2011 A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279–283. (doi:10.1038/nature09692)
 77. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. 2012 Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380. (doi:10.1038/nature11082)
 78. Williamson I, Eskeland R, Lettice LA, Hill AE, Boyle S, Grimes GR, Hill RE, Bickmore WA. 2012 Anterior-posterior differences in HoxD chromatin topology in limb development. *Development* **139**, 3157–3167. (doi:10.1242/dev.081174)
 79. Gonzalez F, Duboule D, Spitz F. 2007 Transgenic analysis of Hoxd gene regulation during digit development. *Dev. Biol.* **306**, 847–859. (doi:10.1016/j.ydbio.2007.03.020)
 80. Hong J-W, Hendrix DA, Levine MS. 2008 Shadow enhancers as a source of evolutionary novelty. *Science* **321**, 1314. (doi:10.1126/science.1160631)
 81. Frankel N, Davis GK, Vargas D, Wang S, Payre F, Stern DL. 2010 Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature* **466**, 490–493. (doi:10.1038/nature09158)
 82. Perry MW, Boettiger AN, Bothma JP, Levine M. 2010 Shadow enhancers foster robustness of *Drosophila* gastrulation. *Curr. Biol.* **20**, 1562–1567. (doi:10.1016/j.cub.2010.07.043)
 83. Rinn JL *et al.* 2007 Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129**, 1311–1323. (doi:10.1016/j.cell.2007.05.022)
 84. Kim YJ, Cecchini KR, Kim TH. 2011 Conserved, developmentally regulated mechanism couples chromosomal looping and heterochromatin barrier activity at the homeobox gene A locus. *Proc. Natl Acad. Sci. USA* **108**, 7391–7396. (doi:10.1073/pnas.1018279108)
 85. Soshnikova N, Montavon T, Leleu M, Galjart N, Duboule D. 2010 Functional analysis of CTCF during mammalian limb development. *Dev. Cell* **19**, 819–830. (doi:10.1016/j.devcel.2010.11.009)
 86. Schorderet P, Duboule D. 2011 Structural and functional differences in the long non-coding RNA *hotair* in mouse and human. *PLoS Genet.* **7**, e1002071. (doi:10.1371/journal.pgen.1002071)
 87. Woltering JM, Duboule D. 2010 The origin of digits: expression patterns versus regulatory mechanisms. *Dev. Cell* **18**, 526–532. (doi:10.1016/j.devcel.2010.04.002)
 88. Dolle P, Izpisua-Belmonte JC, Brown JM, Tickle C, Duboule D. 1991 HOX-4 genes and the morphogenesis of mammalian genitalia. *Genes Dev.* **5**, 1767–1767. (doi:10.1101/gad.5.10.1767)
 89. Nobrega MA, Ovcharenko I, Afzal V, Rubin EM. 2003 Scanning human gene deserts for long-range enhancers. *Science* **302**, 413. (doi:10.1126/science.1088328)
 90. Jeong Y, El-Jaick K, Roessler E, Muenke M, Epstein DJ. 2006 A functional screen for sonic hedgehog regulatory elements across a 1 Mb interval identifies long-range ventral forebrain enhancers. *Development* **133**, 761–772. (doi:10.1242/dev.02239)
 91. Sagai T, Amano T, Tamura M, Mizushima Y, Sumiyama K, Shiroishi T. 2009 A cluster of three long-range enhancers directs regional Shh expression in the epithelial linings. *Development* **136**, 1665–1674. (doi:10.1242/dev.032714)
 92. Tena JJ, Alonso ME, de la Calle-Mustienes E, Splinter E, de Laat W, Manzanares M, Gomez-Skarmeta JL. 2011 An evolutionarily conserved three-dimensional structure in the vertebrate *lrx* clusters facilitates enhancer sharing and coregulation. *Nat. Commun.* **2**, 310. (doi:10.1038/ncomms1301)