

The *HoxD* cluster is a dynamic and resilient TAD boundary controlling the segregation of antagonistic regulatory landscapes

Eddie Rodríguez-Carballo,¹ Lucille Lopez-Delisle,² Ye Zhan,³ Pierre J. Fabre,^{2,4} Leonardo Beccari,¹ Imane El-Idrissi,¹ Thi Hanh Nguyen Huynh,¹ Hakan Ozadam,³ Job Dekker,³ and Denis Duboule^{1,2}

¹Department of Genetics and Evolution, University of Geneva, 1205 Geneva, Switzerland; ²School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland; ³Program in Systems Biology, Department of Biochemistry and Molecular Pharmacology, Howard Hughes Medical Institute, University of Massachusetts Medical School, Worcester, Massachusetts 01605, USA

The mammalian *HoxD* cluster lies between two topologically associating domains (TADs) matching distinct enhancer-rich regulatory landscapes. During limb development, the telomeric TAD controls the early transcription of *Hoxd* genes in forearm cells, whereas the centromeric TAD subsequently regulates more posterior *Hoxd* genes in digit cells. Therefore, the TAD boundary prevents the terminal *Hoxd13* gene from responding to forearm enhancers, thereby allowing proper limb patterning. To assess the nature and function of this CTCF-rich DNA region in embryos, we compared chromatin interaction profiles between proximal and distal limb bud cells isolated from mutant stocks where various parts of this boundary region were removed. The resulting progressive release in boundary effect triggered inter-TAD contacts, favored by the activity of the newly accessed enhancers. However, the boundary was highly resilient, and only a 400-kb deletion, including the whole-gene cluster, was eventually able to merge the neighboring TADs into a single structure. In this unified TAD, both proximal and distal limb enhancers nevertheless continued to work independently over a targeted transgenic reporter construct. We propose that the whole *HoxD* cluster is a dynamic TAD border and that the exact boundary position varies depending on both the transcriptional status and the developmental context.

[Keywords: TAD; CTCF; *Hox*; limb; genome architecture; Hi-C]

Supplemental material is available for this article.

Received September 30, 2017; revised version accepted November 21, 2017.

In mammals, 39 *Hox* genes play critical roles in the organization and patterning of structures during development. They are found clustered at four distinct loci—*HoxA* to *HoxD*—with a high level of structural organization. While all four gene clusters are activated early on during embryogenesis (Deschamps and van Nes 2005), both *HoxA* and *HoxD* clusters are subsequently reactivated during the development of the appendicular skeleton, where they also participate in the building of the limbs (Dolle et al. 1989; Zakany and Duboule 2007). In the latter case, *Hoxa* and *Hoxd* genes are controlled by large regulatory landscapes flanking the gene clusters and harboring multiple enhancers (Montavon et al. 2011; Andrey et al. 2013;

Berlivet et al. 2013). These regulatory landscapes were subsequently found to coincide with topologically associating domains (TADs) (Dixon et al. 2012; Nora et al. 2012), which are defined as genome regions in which chromatin interactions occur more frequently. Such domains tend to be constitutive (Dixon et al. 2012) and hence are mostly conserved between tissues and among various vertebrate species (e.g., Woltering et al. 2014). In addition, TADs correlate with lamina-associated domains (LADs) and DNA replication domains and may thus be considered as units of chromosome organization (see Gonzalez-Sandoval and Gasser 2016).

The *HoxD* gene cluster lies at the border between two such chromatin domains, and various subsets of *Hoxd* genes respond to either limb regulatory landscape.

⁴Present address: Department of Basic Neurosciences, University of Geneva, 1211 Geneva, Switzerland.

Corresponding author: denis.duboule@unige.ch or denis.duboule@epfl.ch
Article published online ahead of print. Article and publication date are online at <http://www.genesdev.org/cgi/doi/10.1101/gad.307769.117>. Freely available online through the *Genes & Development* Open Access option.

© 2017 Rodríguez-Carballo et al. This article, published in *Genes & Development*, is available under a Creative Commons License (Attribution 4.0 International), as described at <http://creativecommons.org/licenses/by/4.0/>.

Initially, the telomeric TAD (T-DOM), located in 3' of the gene cluster, is active and controls the transcription of *Hoxd3* to *Hoxd11* into the most proximal part of the future limb: the arm and the forearm. Subsequently, in distal limb bud cells, the T-DOM is switched off, while the opposite 5'-located TAD (C-DOM) becomes active to control the expression of *Hoxd13* to *Hoxd9* into presumptive digit cells (Andrey et al. 2013; Beccari et al. 2016). Therefore, two successive waves of transcription occur, triggered by distinct enhancer landscapes and in phase with the building of the two main pieces of the future limbs.

The existence of both this switch in regulations and a strong boundary effect introduces a discontinuity in the transcription of these genes, which allows the formation of a zone of low *Hoxd* expression, thus giving rise to the wrist or the ankle (Villavicencio-Lorini et al. 2010; Woltering and Duboule 2010). To produce these critical articulations, it is thus essential that enhancers located in either TAD do not regulate all *Hoxd* genes at once, which would lead to uninterrupted expression domains. Also, it was proposed that both the *Hoxd12* and *Hoxd13* products exert a dominant-negative effect over other HOX proteins (van der Hoeven et al. 1996; Zakany et al. 2004), referred to as "posterior prevalence" (see references in Duboule and Morata 1994; Yekta et al. 2008). This strong inter-TAD border may thus exist in response to the need for *Hoxd13* and *Hoxd12* to not respond to more "proximal" enhancers, since such an ectopic expression would lead to deleterious morphological effects (e.g., Herault et al. 1997), similar to other instances in which TAD boundaries were reported to prevent ectopic interactions that could potentially cause diseases (Lupianez et al. 2015; Fabre et al. 2017).

The exact nature of TAD borders as well as their causality are often difficult to establish. These DNA regions are enriched in bound CTCF and cohesin subunits, suggesting architectural constraints such as helping to either trigger or prevent interactions between promoters and enhancers (Kagey et al. 2010; Sofueva et al. 2013; Zuin et al. 2014). They were shown to function in the constitutive organization of TADs (Dixon et al. 2012; Phillips-Cremins et al. 2013; Rao et al. 2014), since removal of either CTCF or the cohesin complex affects TAD stability (Haarhuis et al. 2017; Nora et al. 2017; Rao et al. 2017; Schwarzer et al. 2017). In the case of *HoxD*, the TAD border can be mapped in the "posterior" part of the cluster, between *Hoxd11* and *Hoxd12*; i.e., in a genomic region showing one of the highest GC contents genome-wide and displaying nine bound CTCF sites within a 40-kb region (Soshnikova et al. 2010) as well as close to 10 active promoters. In this particular genomic context, a functional dissection of this TAD border would require multiple and separate genetic interventions in *cis* to disconnect promoter sequences from those involved in constitutive contacts and thus reveal whether enhancer-promoter contacts either impose a TAD structure or instead are constrained by such a chromatin domain, which would form independently from any transcriptional activity.

Here we address this conundrum by analyzing in embryos the structural and functional effects of a series of

nested deletions involving either part of the boundary region or larger pieces of the *HoxD* locus that include it. We used both proximal and distal microdissected limb bud cells (i.e., two highly related cell types) but where only one or the other of the two TADs is transcriptionally active. While small deletions elicited minor and mostly local effects, larger deletions triggered the rearrangement of interactions, leading to major chromatin reorganization. Altogether, the boundary activity for long-range contacts was surprisingly resilient, and only the absence of a 400-kb DNA region, including the *HoxD* cluster itself, generated a single large TAD made from the fusion between both the T-DOM and the C-DOM. We conclude that several elements in the *HoxD* locus cooperate to impose the requested segregation between the two opposite regulatory influences. The exact positioning of this boundary within the gene cluster as well as its strength in preventing ectopic interactions may have been powerful evolutionary cursors in the shaping of various tetrapod limb morphologies.

Results

A TAD border within the HoxD cluster

In order to gain insights into TAD organization around the *HoxD* locus during limb bud development, we performed Hi-C (chromosome capture followed by high-throughput sequencing) on microdissected distal and proximal limb bud cells isolated from embryonic day 12.5 (E12.5) embryos. At this stage, T-DOM enhancers regulate *Hoxd* gene expression in proximal cells and are silent in distal cells, whereas C-DOM enhancers control *Hoxd* gene targets in future digit cells and are silent in proximal cells. Therefore, the two TADs are either transcriptionally active or inactive in an exclusive manner in the two tissue samples (Fig. 1A,B, top schemes). In both cases, the Hi-C profiles positioned the *HoxD* cluster right between the TADs, similar to what was initially reported in either embryonic stem (ES) cells (Dixon et al. 2012) or CH12 lymphoblastic cells (Supplemental Fig. S1A–D; data extracted from Rao et al. 2014). Although the distribution of contacts was quite similar in the two cell populations, the internal organization of interactions within the TADs displayed few distinctive features at 40-kb resolution. Our analysis of the CH12 lymphoblast ENCODE data sets did not reveal any long-distance contact between the *HoxD* cluster and potential enhancer regions, as both gene deserts appeared globally devoid of H3K27ac marks (Supplemental Fig. S1C, bottom panel).

In distal cells, specific contacts were established between posterior *Hoxd* genes (*Hoxd13* to *Hoxd10*) and previously defined regulatory sequences within the C-DOM (island-1 to island-5, GCR, and Prox) (Fig. 1B; Gonzalez et al. 2007; Montavon et al. 2011). In proximal cells, however, some of these contacts were not detected as strongly (Fig. 1A), and the C-DOM showed lower contact intensities than in distal cells (P -value = 0.018), as revealed by performing a subtraction of both Hi-C data sets (Fig. 1D; Supplemental Fig. S1I). Altogether, however, the two

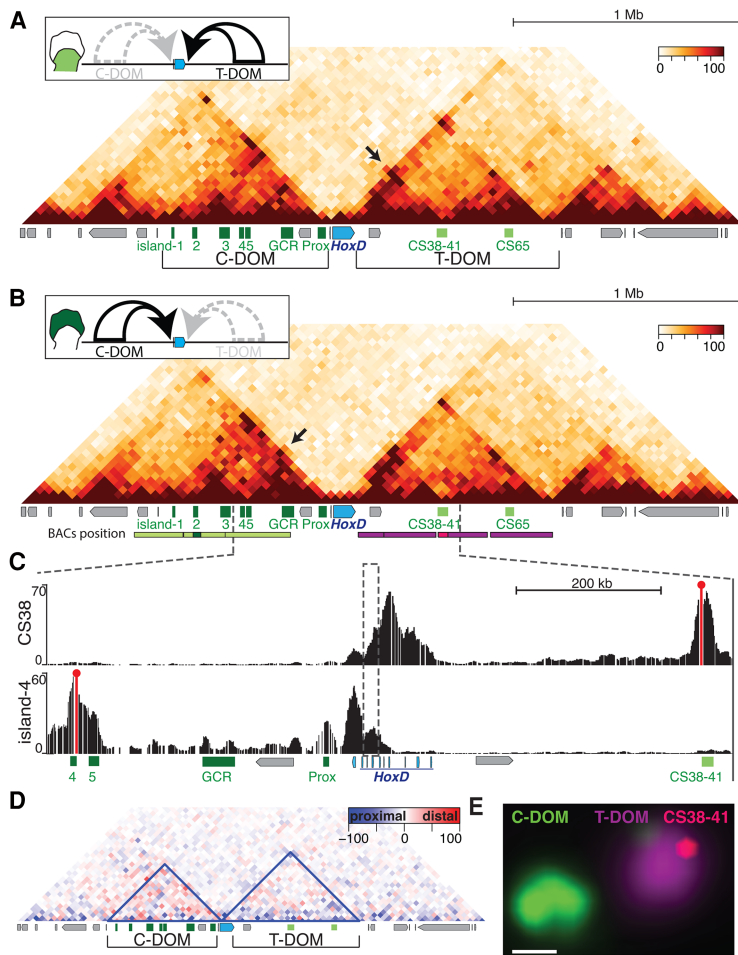


Figure 1. Three-dimensional organization of the *HoxD* locus in limb buds. (A) Hi-C heat map using proximal E12.5 murine limb bud cells. Three megabases of chromosome 2 (mm10, 73,320,000–76,480,000) are covered. The scheme at the top indicates that the T-DOM is active and the C-DOM is inactive. The positions of the *HoxD* cluster (blue) and surrounding enhancers (green) are shown below with CS38-41 and CS65 within the T-DOM, whereas Prox, GCR, and island-1 to island-5 are located within the C-DOM. Other surrounding genes are depicted as gray boxes. The arrow indicates contacts established between central *Hoxd* genes and CS38-41. (B) Hi-C heat map similar to A but using distal cells where the C-DOM is active. The arrow indicates contacts between 5'-located and central *Hoxd* genes and island-3. The red and green bars below illustrate the BAC clones used in DNA-FISH experiments. (C) 4C-seq (circularized chromosome conformation capture [4C] combined with sequencing) tracks showing contacts established by either CS38 (top; red line) or island-4 (bottom; red line) using distal limb bud cells. The dashed vertical rectangle marks the boundary region. (D) Subtraction of Hi-C matrices shown in A and B, with distal cells in red and proximal cells in blue. The blue line demarcates the extension of the identified TADs in distal cells. (E) DNA-FISH using distal limb bud cells and the series of BAC clones shown in B. Bar, 500 nm. The position of CS38-41 inside the T-DOM is shown in red by using a fosmid clone.

interaction maps were quite similar to one another. Likewise, the T-DOM displayed only a few changes in interactions when distal and proximal cells were compared (P -value = 0.87) (Fig. 1D; Supplemental Fig. S11).

Differences between distal and proximal limb cells were nevertheless observed around the so-called CS38-41 region, which also displayed bound CTCF molecules (see below) and appeared to be both a boundary between the two sub-TADs found within the T-DOM (Andrey et al. 2013) and a strong region of interaction with the *HoxD* cluster in the two cell populations. This region contains enhancers for limbs, the cecum, and mammary glands as well as a bidirectional transcription start site for the *hotdog* (*Hog*) and *twin of hotdog* (*Tog*) long noncoding RNAs (lncRNAs) (Delpretti et al. 2013; Schep et al. 2016). CS38-41 was specifically contacted by the central part of the *HoxD* cluster in proximal cells only (Fig. 1A, black arrow), since, in distal cells, the same region of *HoxD* interacted with the opposite C-DOM (Fig. 1B, black arrow). The Hi-C data also revealed a highly interacting region extending from the gene cluster up to CS38-41 in distal cells, where it was covered by H3K27me3 marks (Andrey et al. 2013). Altogether, however, no obvious interactions were detected between the two opposite regulatory landscapes. DNA-FISH analysis using independent

BACs labeling either the C-DOM, the T-DOM, or region CS38-41 (Fig. 1E, green, purple, and pink, respectively) confirmed the isolated spatial conformation of both TADs and their status as independent regulatory units (see Fabre et al. 2015).

While these Hi-C analyses illustrated the strict partitioning between the two TADs, their resolution (40 kb) made it difficult to precisely define the position of the TAD border within the *HoxD* cluster. To identify these limits, we applied various algorithms to our embryonic limb data sets based on isolation potential (Crane et al. 2015; Shin et al. 2016). This approach revealed a boundary with a dynamic position within an ~50-kb DNA interval, with a more centromeric position in proximal cells and a more telomeric position in distal cells (Supplemental Fig. S1E–H, red lines). When the TopDom algorithm was applied to either murine ES cell (Dixon et al. 2012) or CH12 cell (Rao et al. 2014) data sets, a shift in the TAD border along the *HoxD* cluster was also scored. In ES cells, a microdomain was detected involving most of the gene cluster (Supplemental Fig. S1B; Noordermeer et al. 2014; Kundu et al. 2017) as likely associated with the presence of H3K27me3 modifications found throughout *Hoxd* genes in these cells (Bernstein et al. 2005). In CH12, however, the algorithm placed the boundary at the position of

Hoxd9, engulfing the highly active *Hoxd4* gene into the T-DOM (Supplemental Fig. S1D). Therefore, the domain boundary was found located at different positions in the *Hox* cluster depending on the cell population and its set of transcribed *Hoxd* genes.

To more precisely define this TAD boundary in our experimental contexts, we used 4C-seq (circularized chromosome conformation capture [4C] combined with sequencing), an approach with a resolution <5 kb. For example, when the C-DOM island-4 was used as bait in distal limb cells, the strongest interactions were scored with the *Hoxd13* to *Evx2* region, with substantial contacts also observed over *Hoxd11* up to *Hoxd10* (Fig. 1C; Supplemental Fig. S1G). Likewise, when the T-DOM-located bait CS38 was used in the same cells, strong interactions were scored over *Hoxd8* and *Hoxd9*, with a striking decrease in contacts over the *Hoxd10* to *Hoxd11* region (Fig. 1C; Supplemental Fig. S1G), thus positioning a border at around *Hoxd10*, whereas this border was positioned over *Hoxd11* to *Hoxd12* when the CS38 bait was used in proximal cells (Andrey et al. 2013; Supplemental Fig. S1H). The use of these two opposite baits showed that the precise location of the boundary changed in relation to the on-off transcriptional activity of the TADs.

Different subgroups of transcribed *Hoxd* genes are bordered by bound CTCF and cohesin

TAD borders are often enriched in both CpG islands and sites bound by architectural proteins, which may be instrumental in either their formation or their maintenance (Guelen et al. 2008; Dixon et al. 2012). For instance, CTCF and the cohesin complex can form loops between distant regions and hence favor the segregation of chromatin interaction patterns (see Phillips-Cremins et al. 2013; Hnisz et al. 2016). The *HoxD* cluster displays a dense distribution of at least 21 identified CpG islands and contains >10 different promoters, including coding and noncoding genes (Fig. 2A).

In order to study the binding profiles of architectural proteins over the *HoxD* locus and associated TADs, we performed ChIP (chromatin immunoprecipitation) experiments to identify sites bound by either CTCF in distal and proximal limb cells or the cohesin RAD21 and SMC1 subunits in distal limb bud cells. It is noteworthy that the bound CTCF sites were mostly distributed within the centromeric half of the cluster, precisely where different blocks of genes were active in both limb cell populations, thus matching the genomic window where the boundary had been mapped (Fig. 2A). We first used MACS2 peak calling followed by consensus motif identification to classify the bound CTCF sites according to their orientations (<http://insulatordb.uthsc.edu>), given that CTCF sites with divergent orientations are present at many TAD borders (Rao et al. 2014; de Wit et al. 2015; Gómez-Marín et al. 2015; Guo et al. 2015; Tang et al. 2015; Vietri Rudan et al. 2015). Within *HoxD*, all four CTCF sites located at the centromeric side were oriented toward the C-DOM, whereas all but one located at more telomeric positions faced the T-DOM (Fig. 2A, col-

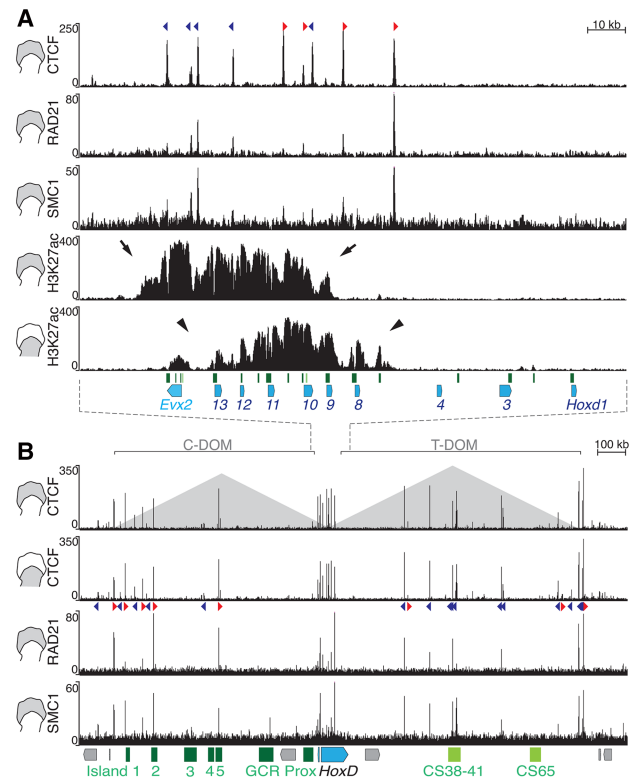


Figure 2. Subsets of *Hoxd* genes responding to C-DOM or T-DOM enhancers coincide with bound CTCF and the cohesin complex. (A) CTCF, RAD21, and SMC1 ChIP-seq (chromatin immunoprecipitation [ChIP] combined with high-throughput sequencing) profiles at and around the *HoxD* boundary region. (Top track) The CTCF profile in distal cells is identical to that in proximal cells, with peaks spanning the centromeric half of the gene cluster. CTCF motif orientation is shown with arrowheads. (Red arrowheads) Telomere-oriented CTCF; (blue arrowheads) motifs oriented toward the centromere. (Bottom) The profiles of RAD21 and SMC1 tend to label the extremities of the H3K27ac domains. These active domains are restricted within a large DNA interval where bound CTCF molecules are observed. (Arrows) Distal limb; (arrowheads) proximal limb. The green boxes below represent CpG islands. Diagrams at the left show whether distal or proximal cells were used. (B) CTCF, RAD21, and SMC1 profiles along both C-DOM and T-DOM TADs (schematized as pyramids). CTCF peaks are conserved in proximal and distal cells. CTCF motif orientation is as in A. Below is the *HoxD* cluster (blue) and various regulatory elements.

ored arrowheads), suggesting an inversion in orientations between *Hoxd12* and *Hoxd11*; i.e., on either side of the TAD border observed in proximal cells.

While the sites of bound cohesin subunits mostly coincided with sequences also bound by CTCF, these subunits were enriched on both sides of the series of bound CTCF; i.e., either between *Hoxd4* and *Hoxd8* or in the *Hoxd13* to *Evx2* intergenic region. Of note, the extension of H3K27ac domains, a histone modification associated with active gene transcription, identified the distinct subgroups of *Hoxd* genes actively transcribed in either proximal or distal limb cells. In both cases, CTCF and cohesin were

bound at—or in the vicinity of—both extremities of these domains (Fig. 2A), as if these proteins were used to somehow label those large target DNA regions successively accessible by first T-DOM enhancers and then C-DOM enhancers.

Bound CTCF and cohesin subunits were also scored within the C-DOM and T-DOM, in particular at important regulatory sequences such as the CS38-41 region as well as at island-1, island-2, and island-5, which were enriched for both CTCF and RAD21 (Fig. 2B). While most of these CTCF sites were orientated toward the *HoxD* cluster, their occupancy remained globally unchanged in the different limb cell populations (Fig. 2B), similar to the situation within the gene cluster, suggesting that CTCF alone may not bring any tissue specificity to these regulations (Fig. 2A,B).

Serial deletions of the TAD boundary or parts thereof

Our Hi-C and 4C-seq data sets thus located the TAD border region somewhere between *Hoxd8* and *Hoxd13*, with some variation depending on the cell type considered. To

try to assess the various components of this boundary, we used a set of deletion alleles where distinct portions of this DNA interval had been removed (Supplemental Fig. S2; see Tschopp and Duboule 2014). We used 4C-seq to document the interaction profiles generated by two opposite viewpoints located at each side of the TAD border (Fig. 3, orange bars). The *Evx2* bait lies immediately near *Hoxd13* on the centromeric side of the boundary, whereas *Hoxd4* is the first gene located clearly outside of this boundary interval on the telomeric side. Consequently, under wild-type conditions, *Hoxd4* is expressed only in proximal limb cells under the control of the T-DOM, while C-DOM enhancers control *Evx2* transcripts in distal cells exclusively.

We scored the interactions of these two baits in both deletion and control alleles. For each bait, we used cells where the operating TAD was on the other side of the border. In this way, we looked for ectopic gains of contacts crossing the boundary region toward a TAD containing enhancers functionally at work. We first analyzed the interactions of *Hoxd4* in distal cells; i.e., when the C-DOM is fully active and the T-DOM is switched off. In this

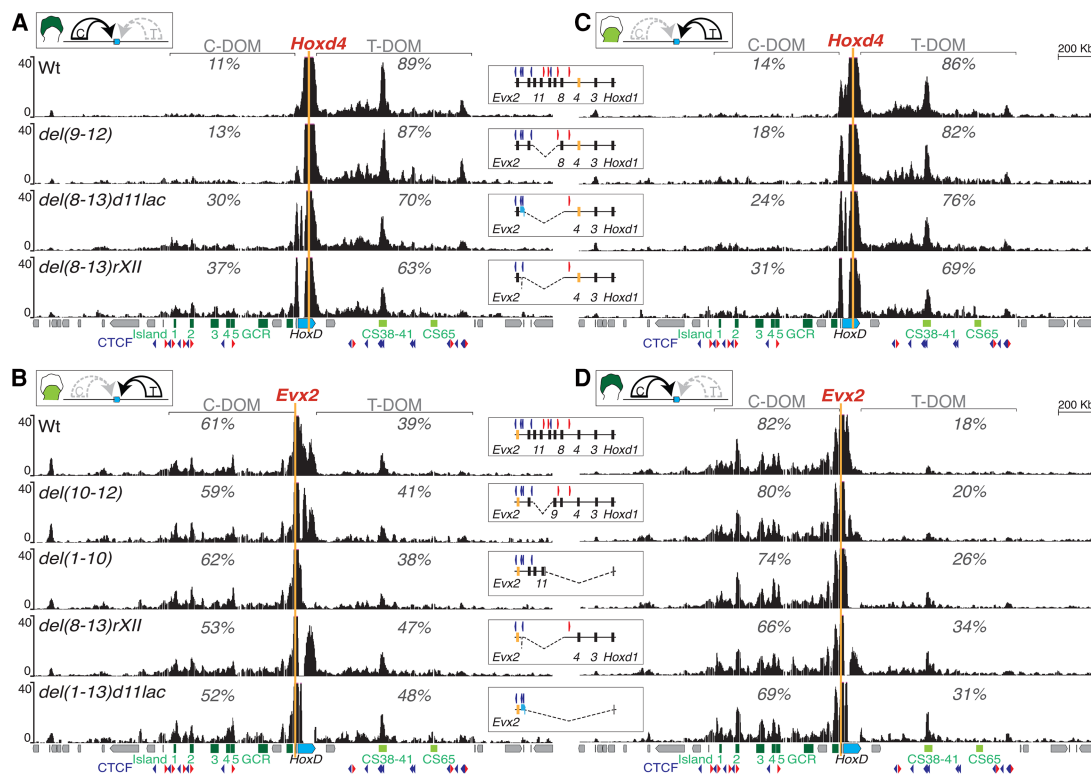


Figure 3. Partial deletions of the inter-TAD border. Interactions established by *Hoxd4* and *Evx2* in a set of deletion alleles, including part of the boundary region. (A) 4C-seq profiles by using *Hoxd4* as a viewpoint (orange line) in distal limb cells where the C-DOM is active (top scheme). The control [wild type [wt]], *HoxD^{del(9-12)}*, *HoxD^{del(8-13)d11lac}*, and *HoxD^{del(8-13)rXII}* are shown from top to bottom, with a schematic at the right indicating the deletion and the viewpoint (orange rectangle). The percentages reflect the ratios of contacts scored in either TAD after excluding reads mapping to the cluster itself. Below are the gene cluster (blue) and regulatory sequences (green). (B) 4C-seq profiles by using *Evx2* as a viewpoint (orange line) in proximal limb cells where the T-DOM is active (top scheme). The control (wild type), *HoxD^{del(10-12)}*, *HoxD^{del(1-10)}*, *HoxD^{del(8-13)rXII}*, and *HoxD^{del(1-13)d11lac}* are shown from top to bottom, with a schematic at the right indicating the deletion and the viewpoint (red rectangle). (C) 4C-seq as in A using *Hoxd4* as a viewpoint in proximal cells where the T-DOM is active. (D) 4C-seq profiles as in B using *Evx2* as a viewpoint in distal cells where the C-DOM is active. For all panels, percentages are as for A.

situation, the control allele revealed only 11% of contacts between *Hoxd4* and the C-DOM, while most of the contacts remained within the T-DOM, illustrating the robustness of the boundary. A fair part of the border interval was removed in the *HoxD^{del(9-12)}* allele, where the DNA region from *Hoxd9* to *Hoxd12* had been deleted. Nevertheless, very little effect, if any, was scored, and *Hoxd4* did not appear to have increased interactions with the active C-DOM (Fig. 3A). When the larger *HoxD^{del(8-13)d11lac}* deletion was used, where almost the full boundary region is removed and replaced by a *Hoxd11/lacZ* transgene, ectopic interactions between *Hoxd4* and the C-DOM started to significantly increase, from 11% to 30% of the contacts (Fig. 3A). Interactions with the C-DOM increased to almost 40% when both the *Hoxd11/lacZ* transgene and a small region containing a CTCF site between *Hoxd13* and *Evx2* (Supplemental Fig. S2) were further removed from this deletion [*HoxD^{del(8-13)rXII}*]. Even in this case, however, contacts established by the *Hoxd4* bait were still biased toward the T-DOM (Fig. 3A), indicating that some boundary activity was left, perhaps associated with the few CTCF- and cohesin-binding sites still present on either side of the latter deletion break points (Supplemental Fig. S3A,C).

The situation was comparable, yet slightly different, when *Evx2* was used as bait. In wild-type proximal limb cells where the T-DOM was active and the C-DOM was inactive, *Evx2* already established substantial interactions with sequences located in the opposite T-DOM (39%) (Fig. 3B). Small deletions such as *HoxD^{del(10-12)}* or larger deletions affecting mostly genes on the telomeric side of the cluster [for example, *HoxD^{del(1-10)}*] did not induce any significant increase of interactions with the T-DOM (Fig. 3B). *Evx2* nevertheless did increase its interactions with the T-DOM whenever the more centromeric *Hoxd* genes were removed; for instance, in the *HoxD^{del(8-13)rXII}* allele or when the entire gene cluster was substituted by a *LacZ* reporter transgene in either *HoxD^{del(1-13)d11lac}* or *HoxD^{del(1-13)d9lac}* (Fig. 3B). In these mutant alleles, the boundary appeared slightly more efficient in blocking telomeric enhancers from accessing C-DOM promoters than centromeric enhancers from leaking over T-DOM genes, an observation perhaps related to the orientation of the remaining CTCF sites (see the Discussion). Altogether, these results suggested that the boundary was a multipartite structure resilient to the deletion of its parts.

Ectopic interactions established by *Hoxd4* when parts of the boundary region were deleted were only marginally affected by the activity of the TADs. Indeed, in proximal cells where T-DOM was active, the increased interactions of *Hoxd4* toward the C-DOM in the various deletions were globally comparable with the situation in distal cells when the T-DOM was inactive (Fig. 3C). Likewise, when the same mutant alleles were compared, such as *HoxD^{del(8-13)rXII}*, the ectopic interactions established by *Evx2* in proximal cells were not drastically different from those scored in distal cells where C-DOM was active (Fig. 3B,D), in particular considering that the contacts between *Evx2* and the T-DOM were already higher in wild-type proximal cells than those between *Hoxd4* and the

C-DOM in distal cells (Fig. 3A,B). This again illustrated that a C-DOM-located promoter was more easily attracted by the opposite T-DOM than was a T-DOM-located gene by the activity of the C-DOM. This feature was also apparent when using *Hoxd13* as bait either in the wild-type chromosome or on a set of deletions. In these various cases, ectopic interactions toward the T-DOM were generally higher than with *Hoxd4* in the opposite situation. In addition, these interactions were increased whenever the T-DOM was transcriptionally active rather than inactive (Supplemental Fig. S4A,B).

Ectopic inter-TAD contacts are specific and productive

We next asked whether the reallocation of interactions observed when using some of these deletion alleles was merely structural or, alternatively, whether it could elicit a transcriptional outcome. We monitored the expression of both *Hoxd4* and *Evx2* in these various alleles and observed ectopic transcriptional activation concurrent with new interactions. For example, in the *HoxD^{del(8-13)rXII}* deletion, *Hoxd4* was strongly expressed in distal cells, and *Evx2* was strongly expressed in proximal cells, a situation never observed in control animals (Fig. 4A; Supplemental Fig. S5, arrowheads). Expectedly, *Evx2* transcripts were also gained in proximal cells after the deletion of the entire *HoxD* cluster (Supplemental Fig. S5B,E). Ectopic transcription precisely correlated with the reallocation of interactions with enhancers. The quantifications of these interactions on specific regions known to be required for transcription of *Hoxd* genes in distal cells (e.g., island-2) showed that the increases in contacts were significant only in those alleles where ectopic expression was scored (Fig. 4B; Supplemental Fig. S5F).

We confirmed these observations by analyzing the steady-state levels of *Hoxd8* mRNAs in various deletion alleles. As for *Hoxd4*, *Hoxd8* transcription remained unchanged in the *HoxD^{del(9-12)}* mutant limb buds, while a weak but significant ectopic expression was scored in distal cells of E12.5 embryos carrying the *HoxD^{del(11-13)rXII}* allele. Of note, *Hoxd8* expression was strongly gained in distal limbs of *HoxD^{del(10-13)rXII}* mutant embryos (Fig. 4C,D, arrowhead), suggesting that the sequential removal of gene promoters and/or CTCF-binding sites progressively weakened the TAD boundary (Narendra et al. 2015). However, we did not observe any ectopic expression of *Hoxd13* in proximal cells even when a large portion of the boundary region had been removed. It is possible that the deletion was not sufficient to induce the ectopic activation of *Hoxd13* even when interactions were gained along the T-DOM, such as in the *HoxD^{del(1-10)}* allele (Fig. 4C; Supplemental Figs. S4C, S5C; Zakany et al. 2004). Altogether, the ectopic interactions and the gains in transcription observed in our series of deletions could not be explained by the mere change in relative position of a given *Hoxd* target gene with respect to the appropriate enhancer sequences. Because of this lack of a simple correlation, we conclude that some specific regions inside this large boundary interval are stronger than others in exerting their isolation potential.

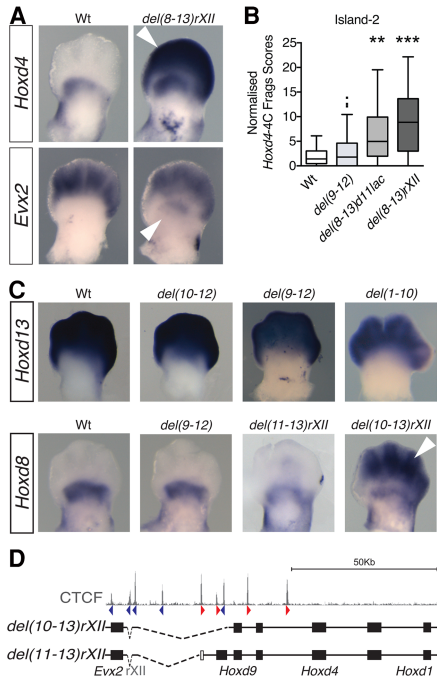


Figure 4. Ectopic contacts are transcriptionally productive. (A) Whole-mount in situ hybridization (WISH) using *Hoxd4* or *Evx2* probes in E12.5 wild-type and *HoxD*^{del(8-13)rXII} mutant forelimb buds. In this deletion allele, *Hoxd4* is massively gained in distal cells (*top*; arrowhead), and *Evx2* is significantly gained in proximal cells (*bottom*; arrowhead). (B) Quantification of 4C-seq contacts mapping to the digit-specific island-2 enhancer in the *HoxD*^{del(9-12)}, *HoxD*^{del(8-13)d11lac}, and *HoxD*^{del(8-13)rXII} mutant alleles. (**) $P < 0.01$; (***) $P < 0.001$, Kruskal-Wallis test. (C) WISH using probes for *Hoxd13* and *Hoxd8* in limb buds of control and several deletion alleles, as indicated at the *top*, illustrating the resilience of the inter-TADs' insulation effect and the gain of *Hoxd8* expression in the *HoxD*^{del(10-13)rXII} (arrowhead), which is much weaker in the shorter *HoxD*^{del(11-13)rXII}. (D) Schematic showing the two latter deletion alleles along with the profile of bound CTCF at the *top*.

Deletions of the TAD boundary

In these 4C-seq experiments, both the *Hoxd4* and *Evx2* baits are located close to the deletion break points and may thus be influenced by proximity effects. Consequently, while they illustrate the accessibility of target promoters to remote enhancers localized in the opposite TAD, they are not appropriate to assess the potential of the *HoxD* cluster to block inter-TAD contacts. In the latter case, ectopic interactions between enhancers located in one TAD and sequences located within the other would represent a major reorganization in local chromatin architecture. We thus performed 4C-seq using as viewpoints two regions with enhancer properties, which also seem to act as major interaction points between the *HoxD* cluster and each flanking TAD. Island-4 belongs to the C-DOM and is an enhancer region strongly contacted by *Hoxd* genes transcribed in distal cells. It is not contacted in brain cells where *Hoxd* genes are inactive (Montavon et al. 2011). In contrast, the CS38 bait belongs to the

CS38-41 region of the T-DOM, a conserved region with multiple enhancer activities in the intestinal cecum, limbs, and mammary buds (Delpretti et al. 2013; Beccari et al. 2016; Schep et al. 2016). Of note, this region contains three occupied CTCF sites, all oriented toward the cluster, and is also enriched in cohesin (Fig. 2B).

These remote viewpoints confirmed that the smallest deletions containing parts of the *HoxD* TAD border did not detectably affect its insulation potential. In the *HoxD*^{del(9-12)} allele, for example, CS38 and island-4 did not gain any substantial contact with the opposite TAD in either distal (Fig. 5A,B) or proximal limb bud cells (Fig. 5C,D). Moderate gains of inter-TAD interactions were nevertheless observed when larger deletions were used, such as the *HoxD*^{del(8-13)rXII}, *HoxD*^{del(1-10)}, or *HoxD*^{del(1-13)d9lac} alleles. With island-4, relative increases of up to 10% of interactions were scored on the opposite TAD when using small deletions, with only a weak effect associated with the on-off transcriptional status of the TAD. The gain in interactions detected between CS38 and C-DOM sequences was more significant in distal cells where the C-DOM was active than in proximal cells (Fig. 5A,C). To more precisely evaluate these effects, we generated in silico genomes corresponding to every deletion allele (Supplemental Fig. S2). In this way, we could analyze the cumulative signals along 3 Mb around the viewpoints (Fig. 5E,F) and cluster the results according to the Euclidean distance between the curves. We noticed a clear effect related to the size of the deletions, with small deletions clustering with the control allele, whereas larger deletions clustered together (Fig. 5G,H).

It is noteworthy that all of these moderate but significant gains in interactions observed with the larger deletion alleles involved contacts with active enhancer sequences. For instance, in the *HoxD*^{del(8-13)rXII}, *HoxD*^{del(1-10)}, or *HoxD*^{del(1-13)d9lac} alleles, CS38 contacts were gained with the island-1 to island-5 as well as with the Prox sequence in distal cells (Fig. 5A, arrowheads), whereas these contacts were not as specific in proximal cells where these enhancers are inactive (Fig. 5B). As for the *Hoxd4* bait (see Figs. 3, 4), we asked whether such ectopic interactions could be productive and trigger transcription of T-DOM sequences into distal limb cells, an expression specificity normally excluded from this TAD (Beccari et al. 2016). As a readout, we used the two lncRNAs *Hog* and *Tog*, which are transcribed in opposite directions starting from the CS38 region (Delpretti et al. 2013). As expected from their genomic localization within the T-DOM, both *Hog* and *Tog* were transcribed in control proximal limb bud cells (Supplemental Fig. S6A). In addition, both whole-mount in situ hybridization (WISH) and quantitative PCR (qPCR) revealed a gain of *Hog* and *Tog* transcripts in distal cells dissected from all mutant embryos carrying a deleted allele where ectopic contacts with the C-DOM digit islands were scored (Supplemental Fig. S6A,B). These gains in *Hog* and *Tog* transcripts in distal cells correlated with the quantification of CS38 interactions with known distal enhancers (Supplemental Fig. S6C). However, the newly established contacts between CS38 and C-DOM island-2 were not reflected by any

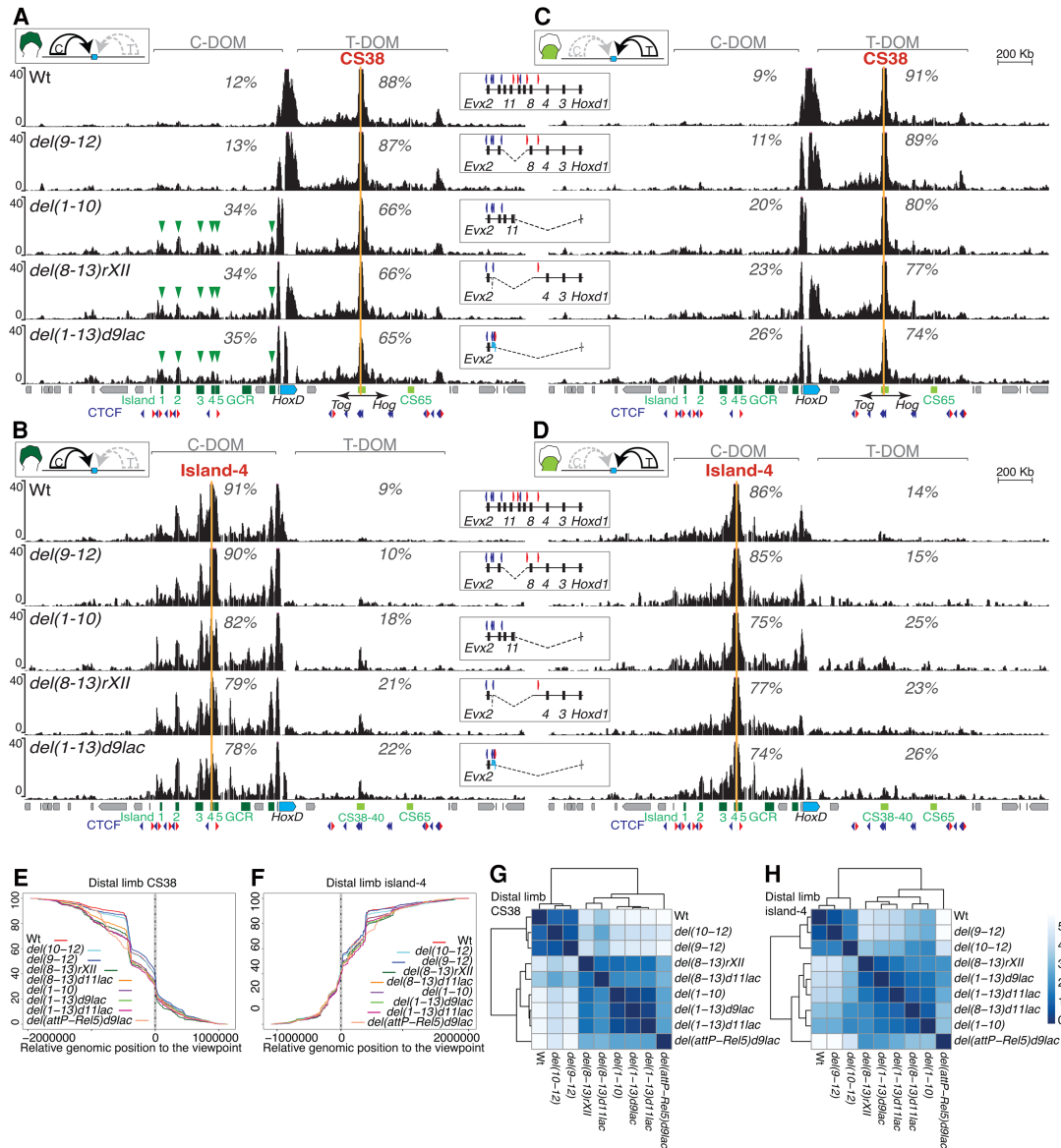


Figure 5. Inter-TAD contacts following partial boundary deletions. (A, from the top down) 4C-seq interaction profiles using region CS38 as a viewpoint (orange line) in distal limb bud cells (scheme at the top) of wild-type, *HoxD^{del(9-12)}*, *HoxD^{del(1-10)}*, *HoxD^{del(8-13)rXII}*, and *HoxD^{del(1-13)d9lac}* alleles. Quantification of contacts in the T-DOM and C-DOM, expressed as percentages, are as in Figure 3. Schematics of the deleted region alleles are shown at the right. The *HoxD* cluster (blue) and regulatory regions (green) are depicted below as well as the bidirectional transcription start site of the lncRNAs *Hog* and *Tog* (arrows) close to CS38. (B) 4C-seq interaction profiles using island-4 as a viewpoint (orange line) in distal limb bud cells (scheme at the top) using the same deleted alleles as in A. (C,D) 4C-seq profiles of CS38 and island-4 viewpoints as in A and B but in proximal cells. (E,F) Cumulative sums of 4C-seq reads relative to the distance to either CS38 (E) or island-4 (F), used as viewpoints, in E12.5 distal limb cells. Colors represent different mutant alleles, and the positions of viewpoints are shown by vertical dashed lines. (G,H) Heat map of Euclidean distances between each pair of curves obtained from E and F. A Ward clustering was performed on the resulting matrix.

substantial change in the spatial distance between these regions, as shown by DNA-FISH using the *HoxD^{del(8-13)rXII}* allele (Supplemental Fig. S6D,E).

The ectopic interactions observed between the C-DOM sequence island-4 and the T-DOM in the larger deletions were also slightly different depending on the activity status of each TAD. When the T-DOM was inactive in distal cells, most of the ectopic contacts in-

involved the CS38-41 region (Fig. 5B). When the T-DOM was active in proximal cells, ectopic interactions between island-4 and the T-DOM were more widespread, involving CS38-41 but also other surrounding sequences (Fig. 5D). The functional outcome, if any, of these ectopic contacts between island-4 and T-DOM sequences nevertheless could not be assessed due to the absence of any known transcription unit mapping to the C-DOM

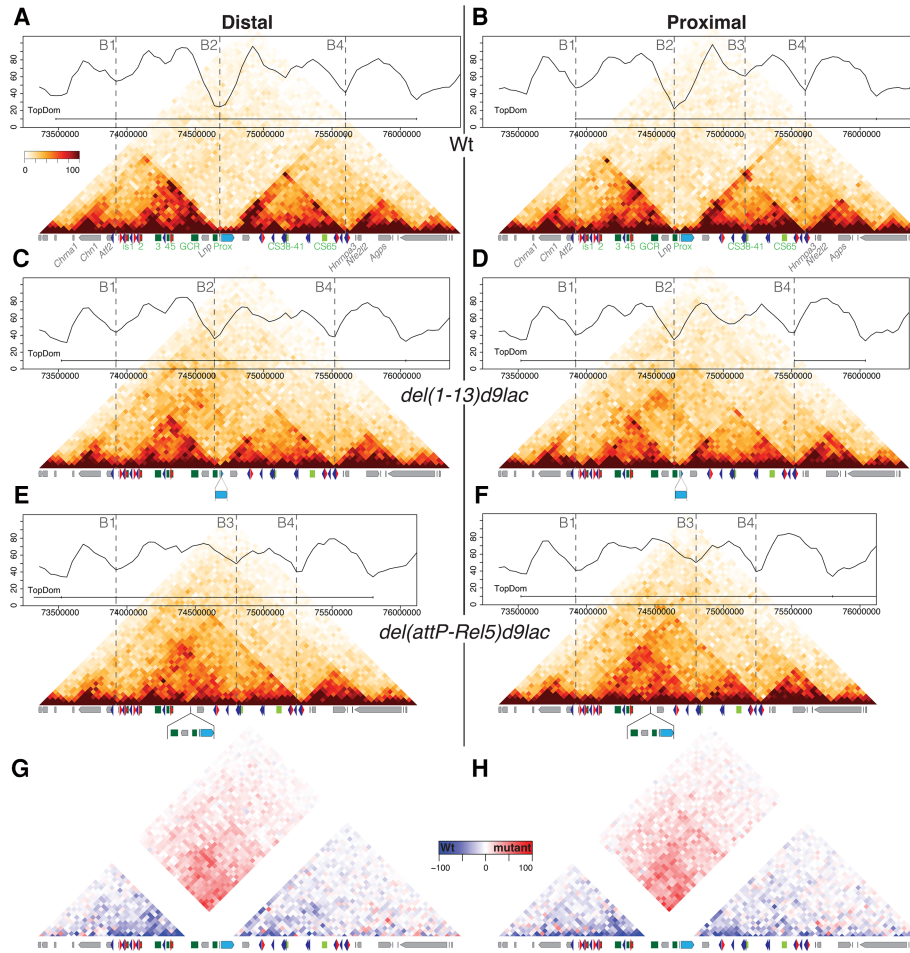


Figure 6. Reorganization of TADs after deletion of the *HoxD* cluster. (A–F) Hi-C profiles covering 3 Mb of mouse chromosome 2 and centered at the *HoxD* locus (blue rectangle) in limb bud cells. The control allele (wild type [wt]) is at the top (A,B), followed by the *HoxD*^{del(1–13)d9lac} (C,D) and *HoxD*^{del(attP–Rel5)d9lac} (E,F) deletion alleles. For each allele, distal cells are at the left, and proximal cells are at the right. Above the Hi-C profiles are graphs showing isolation potential based on the TopDom algorithm. A horizontal bar defines the “consensus” TADs, and vertical dashed lines label boundaries called by the algorithm. These boundaries are referred to as B1–B4 for better comparison between the various alleles. B2 is the TAD boundary at the *HoxD* locus (A,B), which weakens in the *HoxD*^{del(1–13)d9lac} allele (C,D) to disappear in the *HoxD*^{del(attP–Rel5)d9lac} allele (E,F). (E,F) In the latter allele, a new B3 boundary appears in distal cells. (G,H) Subtraction of Hi-C values between the *HoxD*^{del(attP–Rel5)d9lac} and wild-type alleles using the distal and proximal limb data sets. The mutant data sets were mapped on the wild-type genome prior to subtraction.

regulatory islands, which could have been used as a read-out similar to *Hog* and *Tog* for the T-DOM. Altogether, despite some substantial ectopic interactions observed with baits CS38 and island-4, a strong insulation between the two TADs was still observed even when the largest deletions were considered, again emphasizing the robustness of this border and its resistance to perturbations.

Reorganization of TADs

To better document the resilience of this TAD border after large deletions, we performed Hi-C with cells in which the entire gene cluster was deleted and replaced by a *Hoxd9/lacZ* reporter transgene. In our 4C analysis, this *HoxD*^{del(1–13)d9lac} allele still displayed a boundary effect even though significant inter-TAD contacts were de-

tected (Fig. 5; Supplemental Fig. S6). The Hi-C data highlighted this increase of inter-TAD interactions, in particular between the T-DOM sub-TAD flanking the deletion break point and the most centromeric region of the C-DOM (island-1 and island-2, close to the *Atf2* gene), when compared with control cells (Fig. 6A–D). Despite these de novo interactions, a *HoxD* TAD border was called even though it appeared less strong than in control animals, as judged by using the TopDom algorithm (Fig. 6A–D, dashed lines and profiles at the top; Shin et al. 2016). Interestingly, the *Hoxd9/lacZ* transgene contained two occupied CTCF sites with opposite orientations. While one of these sites was equally occupied at the wild-type *Hoxd9* locus, the second one was only very weakly bound in the wild-type condition but strongly re-enforced in the transgene present in this allele (Supplemental Figs. S2, S3B).

Besides this weakened boundary, some reorganizations in intra-TAD contacts were also detected. In distal cells, the C-DOM showed less heterogeneity in interactions in the mutant allele, likely due to a drastic reduction in either the number of target promoters or CTCF sites (Fig. 6C,D; Supplemental Fig. S3B). The same was true for the T-DOM, whose overall interaction density was also reduced in the mutant chromosome. In contrast, TADs located outside the C-DOM and T-DOM remained unchanged (Fig. 6A–D). In proximal cells, these changes were even more pronounced. For instance, the algorithm did not detect the boundary between the two sub-TADs of the T-DOM, at the position of CS38–41 (Fig. 6B, dashed line), which was routinely scored in the control allele, likely due to the reduced interactions between CS38–41 and the target promoters, leading to a lower discrimination between these two sub-TADs (Fig. 6D).

In this *HoxD*^{del(1–13)d9lac} allele, the *Evx2* and *Lunapark* (*Lnp*) promoters were retained as well as the *Hoxd9* promoter present on the *Hoxd9/LacZ* transgene containing two bound CTCF sites and two to three additional CTCF sites located over *Evx2*. The persistence of these CTCF sites may account for the weak but clear boundary effect remaining between the two TADs. To clarify this issue, we used a larger deletion removing both *Evx2* and *Lnp* in addition to the *HoxD* cluster. In this *HoxD*^{del(attP-Rel5)d9lac} allele, where an ~400-kb DNA segment is lacking, only the *Hoxd9/lacZ* reporter gene is left with its two bound CTCF sites with opposite orientations. In this case, despite the presence of the two divergent CTCF sites (Supplemental Fig. S3B), the boundary disappeared, and a new merged TAD formed (Fig. 6E,F).

However, the TAD formed de novo did not result from the fusion between the remains of the C-DOM and the T-DOM (from B1 to B4). Instead, it comprised the remains of the C-DOM (from B1), including island-1 to island-5 and the centromeric sub-TAD (to B3) of the T-DOM with a much weaker contribution of the telomeric sub-TAD of the T-DOM. This was materialized by a boundary call between the newly formed TAD and the telomeric sub-TAD (from B3 to B4) in both tissues using the same algorithm and threshold as before (Fig. 6E,F; Shin et al. 2016). In this case, the contacts established between region 38–41 in the former T-DOM and island-1 and island-2 in the former C-DOM to build the new TAD coincided with the presence of clusters of bound CTCF sites in convergent orientations (Fig. 2B), which normally interact with the series of bound CTCF and cohesin found around the target *Hoxd* genes on either side of the native *HoxD* boundary.

In both distal and proximal cells, the density of interactions within this newly formed TAD (from B1 to B3) was nevertheless below that observed in the control C-DOM (B1–B2) and T-DOM (B2–B4) (Fig. 6, cf. E,F and A,B), indicating that the global solidity of the TAD architecture was dependent on the presence of strong contact points at either side of the border. Presumably, this loss of strength in intrinsic interactions also translated into the establishment of contacts with the next telomeric boundary re-

gion, leading to the inclusion of this new TAD into a larger yet weaker structure delimited by the two original borders (B1 and B4). This marked the centromeric and telomeric extremities of the two TADs, containing all remote enhancers operating at the *HoxD* locus (Fig. 6E,F). These changes were clearly detected when a subtraction was performed between the mutant and the control data sets (Fig. 6G,H).

We used DNA-FISH to see whether such a fusion between the two TADs was accompanied by a reduction in the distance between two BACs covering the T-DOM and C-DOM (Fabre et al. 2015). In distal limb cells, the *HoxD*^{del(attP-Rel5)d9lac} allele indeed showed a significant reduction in inter-TAD distance when compared with control limb cells. However, this reduction was not scored when mutant proximal cells were used, further indicating that the transcriptional status of a given TAD may impact on some of its general properties (Supplemental Fig. S7A). This difference in inter-TAD distance between mutant distal and proximal cells was not anticipated from the Hi-C data set. This tendency nevertheless was supported by an extensive 4C analysis of this large deletion allele. For instance, when the CS38 sequence (in the T-DOM) was used as bait, cross-contacts, in particular with island-1 and island-2, were more noticeable in mutant distal cells than in proximal cells (Supplemental Fig. S7B,C), in agreement with the higher frequency of “short distances” observed in distal mutant cells in the DNA-FISH experiment.

In this large deletion allele, the global reorganization of the TAD architecture at the *HoxD* locus did not severely impact on the neighboring TADs. On the telomeric side, the small domain that includes the *Hnrnpa3*, *Nfe2l2*, and *Agps* genes was not affected at all (Fig. 6A–F). On the centromeric side, some contacts scored in control limbs between either *Hoxd13* or island-1 and island-2 and a sub-TAD containing the *Chn1* and *Chrna1* loci were no longer observed in the *HoxD*^{del(attP-Rel5)d9lac} allele. However, here again, the interaction profiles around these transcription units were not dramatically perturbed by the important modifications occurring in the neighboring C-DOM (Fig. 6E,F).

A recomposed enhancer landscape

In the *HoxD*^{del(attP-Rel5)d9lac} allele, both T-DOM- and C-DOM-specific enhancers are now located within the same TAD. This is in marked contrast to the normal situation, where a strict partitioning was observed between the C-DOM and T-DOM regulatory landscapes. The grouping of forearm enhancers in one TAD and of digit enhancers in the other TAD was considered as the basis of the collinear transcriptional mechanism driving *Hoxd* genes during limb development (Andrey et al. 2013). Therefore, we evaluated the impact of the fusion between TADs and the resulting promiscuity of both types of enhancers in the *HoxD*^{del(attP-Rel5)d9lac} allele by using the *Hoxd9/lacZ* transgene as a readout. In this configuration, a single *Hox* promoter-*lacZ* gene is left in the center of the newly produced TAD along with proximal enhancers located in

3' and digit regulatory island-1 to island-5, located in 5' (Fig. 7).

As a control, we used the exact same *Hoxd9/lacZ* transgene simply inserted at the *rel5* position (Spitz et al. 2003) without any deleted DNA (Fig. 7A). Because the *rel5* site is located within the C-DOM, *lacZ* staining was expectedly detected in distal limb buds as well as in a column of interneurons and some part of the developing brain specific for *Evx2* regulation (Fig. 7A,B; Kmita 2002). In the *HoxD^{del(attP-Rel5)d9lac}* allele, these expression specificities were all maintained. In addition, *LacZ* expression was scored in proximal limbs, in the whisker pads, and in a population of crest cells migrating toward the future mandibles and the axial mesoderm (Fig. 7B, white arrows), all of which are expression specificities controlled by enhancers located in the T-DOM (Spitz et al. 2001). Therefore, the physical separation of enhancers into two distinct

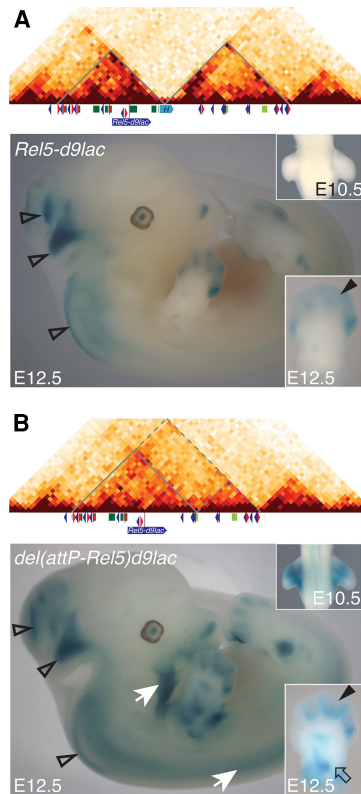


Figure 7. Impact of TAD merging on enhancer specificities. (A, B) β -Galactosidase staining of both *HoxD^{Rel5-d9lac}* (A) and *HoxD^{del(attP-Rel5)d9lac}* (B) mutant E12.5 embryos. The former allele was used as a nondeleted control. In addition to the staining observed in the central nervous system (open arrowheads) and in distal limbs in both cases (bold arrowheads in the *bottom right* boxes), the large deletion shows staining emanating from the T-DOM-based enhancers, as exemplified in proximal limbs (open arrow in the *bottom right* box) or crest cells and paraxial mesoderm (white arrows). The abnormal promiscuity between enhancers does not severely impair their modes of operation. (B) In the deleted allele, the transgene is expressed in limb buds as early as E10.5, whereas it is not detected in the *HoxD^{Rel5-d9lac}* allele (*top right* boxes); i.e., before deletion.

TADs may not be a prerequisite for various C-DOM and T-DOM enhancers to be properly operational in space and time.

Discussion

Alternating long-range regulations

During limb bud development, the T-DOM initially drives the early phase of *Hoxd* gene activation, whereas the C-DOM subsequently regulates the second wave of transcription. The boundary between these two TADs is dynamic and more or less well defined. In ES cells, in the absence of transcription, the entire *HoxD* gene cluster forms a dense domain, which is positioned at the border between the two TADs (Noordermeer et al. 2011; Dixon et al. 2012; Williamson et al. 2014; Fabre et al. 2015; Kundu et al. 2017). In the different developing tissues analyzed thus far, however, the position of the TAD border matches the transcriptionally active versus inactive transition in the gene cluster, reflecting the preferential interaction of transcribed genes with the active TAD (e.g., Andrey et al. 2013; Guerreiro et al. 2016). Therefore, the *HoxD* TAD boundary is initially established in the absence of transcription within an ~ 50 -kb window matching a large part of the *HoxD* cluster, likely in response to architectural proteins and/or other factors intrinsic to chromatin structure. Upon transcriptional activation, this border is refined and matches the transition between active and inactive *Hoxd* promoters. Consequently, the exact position of this boundary varies slightly along various cell types or tissues analyzed, in agreement with the proposal that insulation between TADs is favored by sharp transitions in both CTCF-binding sites and transcriptional activity (Zhan et al. 2017).

While the refinement of the boundary is associated with gene activity, the global positioning of the border at the *HoxD* cluster and its architecture may in turn cause a restriction in the subset of genes capable of responding to either TAD whenever they become activated. For instance, in both proximal limb bud cells and the intestinal cecum where the T-DOM is active, the border is established between *Hoxd11* (positive) and *Hoxd12* (negative) (Andrey et al. 2013; Delpretti et al. 2013). In the mammary gland, however, this boundary seems to form between *Hoxd9* (positive) and *Hoxd10* (negative) (Schep et al. 2016). In contrast, when the C-DOM is activated in either distal limb cells or the developing genitals, the boundary is found somewhere between *Hoxd10* (active) and *Hoxd9* (weakly active) (Lonfat et al. 2014). This partial overlap in the subsets of genes responding to either the T-DOM or the C-DOM may reflect structural constraints and thus participate in the functional exclusivity observed at this locus thus far because the two TADs are never activated concomitantly.

Active versus inactive TADs and loop extrusion

Upon TAD functional activation, specific changes were observed in the interaction profiles, reflecting several

states of configurations in chromatin architecture as reported earlier (Li et al. 2012; Sanyal et al. 2012; Berlivet et al. 2013; Rao et al. 2014; Dixon et al. 2015; Bonev et al. 2017). While some contacts were constitutive, others appeared only when the TAD enhancers were at work. For example, the 5'-located *Hoxd* genes are constitutively anchored to island-1 (which locates close to the next TAD border) as well as island-2 and island-5. In distal cells however, where the C-DOM shows high levels of H3K27ac, island-3 and other regions were also contacted and could thus be used as hallmarks of C-DOM transcriptional activity.

Our mapping of both CTCF sites and H3K27ac suggests that, once functionally active, enhancers within one TAD contact various subsets of target genes, depending on the cellular context. These distinct series of neighbor target genes are delimited by various combinations of bound CTCF sites, as if the presence of CTCF molecules would help define the different sets of target genes responding in any given regulatory context. While the dynamic role of CTCF in marking chromatin domains has been documented (e.g., Narendra et al. 2015), we suggest here that series of bound CTCF sites in close proximity in *cis* may allow for tissue-specific interactions between long-range enhancers and distinct contiguous groups of target *Hoxd* genes, perhaps through the selection of different CTCF sites in various contexts. However, the CTCF profiles analyzed in this study are invariable between distal and proximal limb cells, and TAD borders tend to be co-occupied by CTCF and cohesin complexes (see Ghirlando and Felsenfeld 2016), suggesting that other tissue-specific factors may be involved in the definition of subgroups of target *Hoxd* genes in combination with constitutive proteins. Our deletion analyses support this view, since the most notable effects on chromatin architecture were scored when the posterior part of the cluster was affected; i.e., the DNA interval where CTCF sites are concentrated. In such cases, deleting parts of the cluster would reconfigure the microarchitecture, thus leading to another set of possible target genes.

Within the *HoxD* cluster, the CTCF sites located centromeric to *Hoxd11* are orientated toward the C-DOM, whereas sites located telomeric point toward the T-DOM (Fig. 2A; Supplemental Fig. S2B). Also, the sites occupied by CTCF within either TAD and that correspond to the strongest interactions with *Hoxd* genes, including those at the two remote TAD boundaries, are mostly orientated toward the *HoxD* cluster. These observations support a loop extrusion model for the formation of these three-dimensional (3D) chromatin domains (Rao et al. 2014; Sanborn et al. 2015; Fudenberg et al. 2016). In this view, the multiple copies of CTCF sites in *cis* around *Hoxd* genes may offer different possibilities for determining the extent of loop extrusion and thus lead to distinct positions of the boundary in various contexts, perhaps due to slightly different stabilization of loop-extruding factors (for instance, cohesin) at neighboring but distinct sites.

However, while our mutant alleles can be generally reconciled with this interpretation, some alleles are more dif-

ficult to integrate into this model. The large majority of our deletion alleles indeed maintains at least one pair of CTCF sites with opposed orientations, which could thus account for the persistence of a *HoxD* boundary even with a much weaker insulation potential. For instance, in the *HoxD^{del(1-13)d9lac}* condition, while all native CTCF sites orientated toward the T-DOM are deleted, two opposite sites are brought by the *Hoxd9* transgene, which may account for the weak boundary still observed. In contrast, the *HoxD^{del(1-10)}* allele lost all CTCF sites orientated toward the T-DOM and kept only those sites pointing toward the C-DOM (Supplemental Figs. S2B, S3C). Despite this imbalance in site orientation, the interactions observed with the T-DOM-specific bait CS38 revealed a strong insulation effect virtually identical to that scored when the C-DOM-specific bait island-4 was used. This suggests that the series of CTCF sites orientated toward the T-DOM at the position of the boundary are not prerequired to the formation of the telomeric TAD. In this case, however, the centromeric TAD should not be affected (all appropriate CTCF sites remain), and this domain may prevent interactions with T-DOM region CS38 from occurring.

The *HoxD^{del(attP-Rel5)d9lac}* allele provided us with the minimal boundary elements potentially necessary for insulating the two TADs. A single *Hox* gene was left, with a transcriptionally active promoter fully capable of responding to both C-DOM and T-DOM enhancers. In addition, this transgene harbored two occupied CTCF sites with opposite orientations, each of them facing its neighboring regulatory landscape. However, no particular insulation effect was detected in this condition, and the transgene responded rather correctly to all surrounding enhancers now belonging to a large and unified TAD (see below). This result suggests that at this particular locus, the required border between TADs is built through an additive effect of many elements, which altogether provide the tightness necessary to prevent illegitimate enhancer-promoter interactions. This may also explain why this boundary is still clearly detected in the almost complete absence of CTCF (data in Nora et al. 2017).

Attracting landscapes, tolerated interactions, and border directionality

In several partial deletion alleles, ectopic interactions leaked over the border, leading to contacts between some *Hoxd* genes and the “wrong” TAD. These leakages were not passive but instead often coincided with the activity of the TAD involved, as if an active TAD could attract ectopic contacts more efficiently than when inactive. Shared transcription factors and RNA polymerase II occupy both active enhancers and the set of target promoters, likely stabilizing the interaction (Kieffer-Kwon et al. 2013; Mousavi et al. 2013) and thus making transboundary contacts easier to detect by chromosome conformation capture. The role of cohesin and mediator, which also seem to be enriched according to the differentiation status (Phillips-Cremins et al. 2013), could also be investigated in this context.

Two different types of ectopic contacts were observed in our deletion alleles. The first one involved contacts between regulatory sequences belonging to both TADs, such as, for example, increased interactions between the T-DOM CS38 sequence and regulatory islands located within the C-DOM. While these contacts were scored, their deleterious effects are unlikely, since they did not involve the misregulation of any important transcription unit. The second category of ectopic contacts involved the leakage of specific *Hoxd* target genes into another TAD, thus bringing them under the control of a distinct set of enhancers. For example, some deletions allowed *Hoxd4* to contact C-DOM and thus be expressed in distal limb cells, whereas some others activated *Evx2* into proximal limb cells due to its illegitimate interactions with the T-DOM. In this case, the misregulation of *Hoxd* genes could lead to potential alterations in morphological development. Accordingly, the tight and resilient boundary observed at the *HoxD* locus may have evolved primarily to prevent the precocious and ectopic expression of neighboring *Hoxd* genes during development rather than to avoid inter-TAD contacts.

This possibility is supported by the apparent directionality in the leakage potential of flanking genes in control and deletion alleles. The analyses of several deletions indeed pointed to a general tendency for 3'-located genes (*Hoxd4* and *Hoxd8*) to respond to C-DOM enhancers more readily than 5'-located genes (*Evx2* and *Hoxd13*) would respond to T-DOM enhancers, as if the boundary effect was more efficient in blocking proximal than distal regulations from reaching the opposite extremities of the gene cluster. This property could already be observed in control mice, with digit enhancers leaking up to *Hoxd9* even though only *Hoxd13* showed an unambiguous function during digit development. In contrast, proximal enhancers are readily blocked at the *Hoxd11* locus, one of the key genes for zeugopod development (Davis et al. 1995), and no leakage in contact is observed onto *Hoxd13*.

An adaptation to posterior prevalence

This directional property may be related to the rule of posterior prevalence, a functional property of posterior HOX proteins to often suppress the function of anterior ones when coexpressed, unlike in the opposite situation (Gonzalez-Reyes et al. 1990; Bachiller et al. 1994). As a consequence, the ectopic expression of group 13 *Hox* genes causes dramatic phenotypic alterations (Young et al. 2009; see Mallo et al. 2010). Therefore, while contacts between digit enhancers and more 3'-located anterior *Hoxd* genes may not have any functional consequences as long as *Hoxd13* is expressed there, the opposite situation (where *Hoxd13* would respond to various T-DOM enhancers) may readily elicit abnormal phenotypes. The sensitivity of this effect was observed previously when a subtle *Hoxd13* gain of function in proximal limb bud cells was enough to induce a light limb malformation (Tschopp and Duboule 2011). As a consequence, the *HoxD* TAD boundary must be very stringent in blocking proximal enhancers as a necessary adaptation to posterior prevalence,

whereas digit enhancers may have interactions with various target *Hoxd* genes without any particular effect. While the mechanistic basis of this directionality is unclear, it may rely on the complex distribution and various interaction strengths of architectural proteins at the boundary (see above), as reported in other cases (e.g., Symmons et al. 2014; Tang et al. 2015; Ji et al. 2016).

TAD ontology at the HoxD locus

HoxD lies between two TADs carrying distinct sets of regulatory sequences and operating one after the other in an exclusive manner. While the necessity of functionally separating two sets of target genes is discussed above, the question remains regarding whether groups of enhancer sequences with distinct specificities must segregate into different landscapes to properly work in space and time as suggested previously (Beccari et al. 2016). We show that in our largest deletion, a single TAD now forms containing at least five out of the seven digit regulatory elements as well as a strong proximal enhancer. Our targeted reporter transgene allowed us to conclude that most, if not all, enhancers could still exert their regulatory potential over this single promoter. This suggests that the two-TAD organization at the *HoxD* locus did not evolve to provide particular structural environments to series of holoenhancers, such as optimizing their regulatory inputs once they become functional. Instead, this partitioning in global regulations might be necessary to properly assign subsets of target *Hoxd* genes to their appropriate enhancers. This observation, added to a previous experiment showing that two enhancers located far from one another within the same TAD could work efficiently when associated with a unique small transgenic construct (Lonfat et al. 2014), supports a modular view of enhancer organization within TADs, whereby relative positions may not importantly impact on their functionalities.

Materials and methods

Animal experimentation and mouse mutant lines

All experiments were performed in agreement with the Swiss law on animal protection under license number GE 81/14 (to D. Duboule). All tissues were obtained from either E12.5 or E10.5 mouse embryos coming from the *HoxD^{del(10-12)}*, *HoxD^{del(9-12)}*, *HoxD^{del(8-13)rXII}*, *HoxD^{del(8-13)d11lac}*, *HoxD^{del(1-10)}*, *HoxD^{del(1-13)d9lac}*, *HoxD^{del(1-13)d11lac}*, *HoxD^{del(10-13)rXII}*, *HoxD^{del(11-13)rXII}*, and *HoxD^{Rel5-d9lac}* mutant stocks already reported by this laboratory. The *HoxD^{del(attP-Rel5)d9lac}* was generated by targeted meiotic recombination (TAMERE) (Herault et al. 1998) between the *HoxD^{attP}* (Andrey et al. 2013) and *HoxD^{Rel5-d9lac}* (Montavon et al. 2011) lines. To facilitate reading of the figures, the names of the alleles were reduced to the aforementioned superscripted annotations. All experiments were conducted using homozygous embryos derived from heterozygous crosses.

Mutant genomes in silico

For the *HoxD^{del(10-12)}*, *HoxD^{del(9-12)}*, *HoxD^{del(8-13)rXII}*, *HoxD^{del(8-13)d11lac}*, *HoxD^{del(1-10)}*, *HoxD^{del(1-13)d9lac}*,

HoxD^{del(1–13)d11lac}, and *HoxD*^{del(attP-Rel5)d9lac} deletion lines, a corresponding mutant genome was built in silico to allow for a precise mapping of reads without apparent gaps. Chromosome 2 of these mutant genomes was built using mm10 as a backbone and applying the insertions/deletions using the package *seqinr* (Charif and Lobry 2007) in R software (Supplemental Fig. S2; <http://www.R-project.org>).

Hi-C

Distal and proximal forelimb and hindlimb bud tissue from control, *HoxD*^{del(1–13)d9lac}, and *HoxD*^{del(attP-Rel5)d9lac} were microdissected and collected individually. Cells were dissociated in 10% FBS/PBS with collagenase XI (Sigma, C7657) to a final concentration of 0.4–0.6 µg/µL, and samples were incubated for 60 min at 37°C in agitation (650 rpm). The cell suspension was strained and fixed for 10 min in formaldehyde (2% final concentration in 10% FBS/PBS). Cells were then centrifuged to discard the supernatant and frozen at –80°C until subsequent use after genotyping. Hi-C libraries were generated using the HindIII enzyme as described in Belton et al. (2012). Hi-C libraries were sequenced on an Illumina HiSeq 4000 platform and 50-base paired-end reads were obtained. Reads were mapped, filtered, and bias-corrected as described before (Lajoie et al. 2015; Giorgetti et al. 2016). The Hi-C data sets generated in this work as well as the mouse ES cell (Dixon et al. 2012) and CH12 (Rao et al. 2014) available data sets were processed identically. Briefly, read pairs were mapped independently starting at 25 base pairs (bp) and iterated every 5 bp using Bowtie2 (version 2.2.4) (Langmead et al. 2009) as in Imaikaev et al. (2012) with the parameter `--very-sensitive` on either the mouse genome (mm10) or the mutant genomes generated in silico. Each read was assigned to a fragment using the 5' mapped position shifted 3 bp toward the 3' position to correctly assign the reads overlapping the cutting sites. The fragment assignment and mapping strand from R1 and R2 were combined and used to filter out the single-side mapped pairs, dangling end pairs, error pairs, and self-circle pairs. For each condition (same tissue and same genotype), two replicates were merged (15 replicates for CH12), and the interactions were filtered to discard duplicates. Each fragment was assigned to a bin (40 or 20 kb) based on the position of the middle of the fragment, and each valid interaction was assigned to a pair of bins (one bin for the R1 and one bin for the R2) and reported in the raw matrix. Prior to the ICE (iterative correction and eigenvector decomposition) normalization (Imaikaev et al. 2012), the rows and columns were masked if the sum of reads in this region was 10-fold less than the expectation with uniform coverage or if the number of fragments covered by at least two reads in this region was less than half the number of the fragments of this region. The normalized matrices were used for figures. In Figures 1D and 6, G and H, the difference between the two normalized matrices is plotted. All plots of matrices were generated with R software (<http://www.R-project.org>). The insulation index in Figure 6 was evaluated by TopDom (Shin et al. 2016) with a window size of six bins for the 40-kb matrices, corresponding to a –240-kb, +240-kb “diamond.” To call “consensus” TADs out of which the TAD borders were called, the TopDom algorithm was run with window sizes from three to 15 from the 40-kb binned matrices. Only the TADs present with the exact same coordinates in at least 40% of the window sizes were considered as “consensus.” In Supplemental Figure S1, consensus TADs were called from 20-kb resolution for ES cell and CH12 Hi-C data. To quantify the difference in contact intensities between the proximal and distal wild-type data sets, a Wilcoxon rank sum test with continuity correction was performed on the normalized values of every bin except the one on the diagonal in the C-DOM (chromosome 2: 73,960,000–74,680,000)

and in the T-DOM (chromosome 2: 74,720,000–75,600,000), as called in the distal data set. To be able to compare the contacts between the Hi-C data from *HoxD*^{del(attP-Rel5)d9lac} and the Hi-C data from wild type, the mutant Hi-C data were mapped on the wild-type mm10 genome. Before the ICE normalization, the contacts involving bins representing deleted regions in the mutant genome (chromosome 2: 74,400,000–74,760,000) were removed from both wild-type and mutant data sets. The computations were performed at the Vital-IT Center for High-Performance Computing of the Swiss Institute of Bioinformatics (<http://www.vital-it.ch>).

4C-seq

The distal and proximal parts of forelimb buds were dissected in cold PBS, placed in 250 µL of PBS/10% FBS, and digested in the presence of collagenase XI (Sigma, C7657) to a concentration of 0.4–0.6 µg/µL. Samples were incubated for 45 min at 37°C with agitation. The cell suspension was strained through a mesh (Falcon, 352235), fixed in 2% formaldehyde (in 10% FBS/PBS), lysed, and centrifuged in order to obtain free nucleus precipitate, which was frozen at –80°C and stored. After genotyping, 10–14 pairs of each tissue were pooled in 500 µL of 1.2× CutSmart buffer (New England Biolabs) and digested with NlaIII (New England Biolabs) as described in Noordermeer et al. (2011). After the first 4.5 h of ligation, samples were digested using DpnII (New England Biolabs) in the corresponding buffer overnight and ligated again for 4.5 h. Short fragments and nucleotides were discarded with the nucleotide removal kit (Qiagen), and libraries were prepared by means of 12–16 independent PCR reactions using 70–100 ng of DNA on each (Supplemental Table S1). PCR products were pooled and purified using the PCR purification kit (Qiagen). Up to 22 libraries were multiplexed by either combining different viewpoints or means of 4-bp barcodes added between the Illumina Solexa adapter sequences and the specific viewpoint the inverse forward primer and sequenced using 100-bp single reads on the Illumina HiSeq system. The obtained reads were demultiplexed, mapped, and analyzed using the pipeline present at Bioinformatics and Biostatistics Core Facility (BBCF) HTSstation (<http://htsstation.epfl.ch>) (David et al. 2014) on the ENSEMBL mouse assembly GRCm38 (mm10). The profiles were smoothed using a window size of 11 fragments. The numbers of replicates obtained for each experiment are listed in Supplemental Table S2.

4C-seq normalization and quantifications

All of the 4C-seq profiles were normalized to the distribution of reads along 5 Mb upstream of and downstream from each viewpoint region, except for the LacZ viewpoint. The quantification of contact distribution along the T-DOM (mm10, chromosome 2: 74,781,516–75,605,516) and C-DOM (mm10, chromosome 2: 73,914,154–74,636,454) was performed as read percentage of their read sum; e.g., T-DOM/(T-DOM + C-DOM) × 100. The reads from the stated regions were obtained through the post-processing operations offered at HTSstation (David et al. 2014; <http://htsstation.epfl.ch>). The quantification of contacts established at regulatory sequences was performed using the intersect BEDtools resource. The results show the distribution of 4C fragments at the given regions (mm10): island-1 (chromosome 2: 73,970,064–73,983,434), island-2 (chromosome 2: 74,060,473–74,082,287), island-3 (chromosome 2: 74,177,798–74,223,313), island-4 (chromosome 2: 74,263,814–74,284,643), island-5 (chromosome 2: 74,289,658–74,313,573), GCR (chromosome 2: 74,445,394–74,498,046), Prox (chromosome 2: 74,604,505–74,639,799), CS38-41 (chromosome 2: 75,120,051–75,165,771), and CS65 (chromosome 2: 75,413,472–

75,451,553). Graphs and statistical analysis were performed with GraphPad Prism 7.

4C-seq relative cumulative frequency

For the relative cumulative frequency, the 4C-seq data were mapped to their respective newly generated genome and processed using the pipeline present at the BCCF HTSstation. The output used for the relative cumulative frequency was the segtofrag file. In the plot and for each data set, the data were shifted in order to put the coordinates of the viewpoint at 0. Only the data between $-1,092,537$ and $2,006,380$ for island-4 and between $-1,957,157$ and $11,41,528$ for CS38 were used. These regions correspond to chromosome 2 (73,180,041–76,279,897) coordinates in the wild-type genome (mm10).

DNA-FISH

3D DNA-FISH was performed as in Morey et al. (2007) and Fabre et al. (2015). Fosmids were used for both CS38-41 (W11-2299-17, mm10, chromosome 2: 75,122,702–75,160,145) and island-2 (W11-109P4, mm10, chromosome 2: 74,064,904–74,104,783). Several BACs were used to cover the T-DOM and C-DOM: T-DOM-1 (RPC1-23-190O13, mm10, chromosome 2: 74,714,710–74,911,321), T-DOM-2 (CH29-519G12, mm10, chromosome 2: 74,893,841–75,119,533), T-DOM-3 (CH29-617N10, mm10, chromosome 2: 75,131,563–75,340,886), T-DOM-4 (CH29-6K11, mm10, chromosome 2: 75,354,051–7,5619,849), C-DOM-1 (RP23-146O7, mm10, chromosome 2: 73,821,548–74,029,145), C-DOM-2 (RP23-427C9, mm10, chromosome 2: 74,032,726–74,211,877), and C-DOM-3 (RP24-222J8, mm10, chromosome 2: 74,211,948–74,492,098). For the *HoxD^{del(8–13)rXII}* experiments, fosmid CS38-41 and island-2 were used as well as CS38-41, C-DOM-1, C-DOM-2, and C-DOM3. For *HoxD^{del(attP-Rel5)d9lac}* (Supplemental Fig. S7), C-DOM-1, C-DOM-2, T-DOM-1, T-DOM-2, T-DOM-3, and T-DOM-4 were used. Images were captured using an inverted Olympus IX81 microscope with a 60× plan-apo objective (numerical aperture of 1.42) and a black-and-white CCD Orca ER B7W Hamamatsu camera. Stacks with a 200-nm step were saved as TIFF stacks, reconstructed, and deconvoluted using FIJI (version 1.47q, National Institutes of Health) and Huygens remote manager (Scientific Volume Imaging). The distances between DNA-FISH signals were quantified using an automated spot and surface detection algorithm followed by visual verification and manual correction using IMARIS version 6.5 (Bitplane AG) and Matlab 7.5 (MathWorks SA). Statistical significance analyses of distances were performed using the Kruskal-Wallis test followed by Dunn's post-test. The displayed representative images (Supplemental Figs. S6, S7) were taken from distal forelimb samples.

ChIP-seq (ChIP combined with high-throughput sequencing)

All H3K27ac, SMC1, and RAD21 experiments were processed as ChIP-seq. All CTCF experiments were also processed as ChIP-seq with the exception of *HoxD^{del(8–3)rXII}*, where the ChIPmentation protocol was used (see below). Limb tissues were dissected and fixed in 1% formaldehyde/PBS for 10 min. Chromatin was sheared with either a tip point sonicator (BioBlock Vibra-cell) or a bath sonicator (Diagenode Bioruptor Pico) in order to obtain fragments ranging from 150 to 700 bp. Chromatin was precipitated with anti-CTCF (Active Motif, 61311), anti-RAD21 (Abcam, ab992), anti-SMC1 (Bethyl Laboratories, A300-055A), or anti-H3K27ac (Abcam, ab4729) using agarose beads and following the Active Motif protocol. Libraries were done with at least 4

ng of DNA following the Illumina protocol and sequenced to 50-bp single-end reads on an Illumina HiSeq.

ChIPmentation

Limb tissues were dissected, fixed, and sonicated as for ChIP-seq experiments. CTCF ChIP for *HoxD^{del(8–3)rXII}* was carried out using the ChIPmentation protocol of Schmidl et al. (2015). Chromatin was incubated overnight with antibodies, and magnetic beads were added for at least 3 h afterward. Washes were performed with TF-WBI, TF-WBIII, and 10 mM Tris-HC (pH 8). Next, 1 μ L of transposase was added for 1 min at 37°C, and washes were repeated with TF-WBI and TET. qPCR was carried out to determine the amount of cycles to be applied during library amplification. Libraries were done using Nextera custom adapter sequences and multiplexed for sequencing. All PCRs were done using the KAPA PCR system (KAPA Biosystems, KM2605) after heating the polymerase mix for 45 sec. Library purification was performed with AMPureXP beads. A beads to sample ratio of 0.7:1 was applied to remove long fragments, and the recovered supernatant was incubated in a beads to sample ratio of 2:1. Beads were then eluted using 25 μ L of 10 mM Tris. Libraries were sequenced to 50-bp single read on an Illumina HiSeq.

ChIP, ChIPmentation, and RNA sequencing (RNA-seq) analyses

The profiles of ChIP and ChIPmentation were obtained using the following process: Adapters and bad-quality bases were removed with Cutadapt version 1.8 (Martin 2011) options `-m 15 -q 30 -a GATCGGAAGAGCACACGTCTGAACTCCAGTCAC` for ChIP and `-a CTGTCTCTTATACATCTGACGCTGCCGACGA` for ChIPmentation. Reads were further mapped using Bowtie2 version 2.2.4 on the mm10 genome (Langmead and Salzberg 2012) using default parameters. BAM files were merged for replicates. The coverage was obtained as the output of MACS2 version 2.1.1.20160309 (Zhang et al. 2008) with command line `macs2 callpeak -t input.bam -call-summits -B`. By default, MACS2 kept only one tag at the same location (the same coordinates and the same strand), which would remove all potential contaminants from 4C experiments. A summary of the ChIP/ChIPmentation analyses is in Supplemental Table S3. Motif orientation was assessed using the resources of the CTCFBSDB 2.0 database (<http://insulatordb.uthsc.edu>) by focusing on the motifs identified as MIT_LM7 and their associated strands. For the CH12 lymphoblast H3K27ac ChIP-seq, ENCODE files ENCFF001KBR and ENCFF001KBQ were analyzed the same way. The BAM files of CH12 lymphoblast RNA-seq were downloaded from ENCODE ENCFF507RJZ and ENCFF469ZCH and merged. Only the uniquely mapped reads were kept for the coverage.

RNA extraction and qPCR

Total RNA was extracted following Qiagen's RNeasy minikit. RNA was retrotranscribed into cDNA using Promega GoScript reverse transcriptase. Custom SYBR probes were used for qPCR in a Bio-Rad CFX machine (96-well plates) or an ABIPrism machine (384-well plates). Fold inductions were assessed by the $\Delta\Delta$ -CT method being referred to *Tubb* expression levels. The primers used were those described in Montavon et al. (2008) and Delpretti et al. (2013). Graphs and statistical analysis were performed with GraphPad Prism 7.

β -Galactosidase staining and in situ hybridization

Embryos were fixed in 4% PFA/PBS for 30 min and washed three times for 10 min in PBS-T (0.1% Tween). Specimens were then

stained at 37°C in a solution containing 5 mM potassium hexacyanoferrate (III), 5 mM potassium hexacyanoferrate (II) trihydrate, 2 mM magnesium chloride, 0.01% sodium deoxycholate, 0.02% NP-40, and 1 mg/mL X-gal solution in PBS. After proper staining was achieved, the specimens were washed three times for 15 min in PBS-T, fixed again in 4% PFA/PBS for 30 min, and washed again. Images were taken with a Leica MZFLIII microscope. WISH was performed as described in Woltering et al. (2009). Images were taken with a Leica MZFLIII microscope.

Data repository

All original and reanalyzed sequencing data have been deposited in the Gene Expression Omnibus (GEO). The study superseries is under number GSE101717, which contains the subseries GSE101713 (4C-seq), GSE101714 (ChIP-seq), and GSE101715 (Hi-C).

Acknowledgments

We thank members of the Duboule laboratories for sharing material and discussions. We also thank Marion Leleu for help with the HTSstation platform; Anamaria Necsula for help with bioinformatics analyses; Christian Schmidl and Daan Noordermeer for help on ChIPmentation and 4C-seq experiments, respectively; Johan Gibcus and Jon-Matthew Belton for advice on Hi-C sampling and mapping; Bryan R. Lajoie for generating scripts for the analysis; and Mylène Docquier and Brice Petit from the genomics platform in Geneva. This work was supported by funds from the Ecole Polytechnique Fédérale (Lausanne), the University of Geneva, and the Swiss National Research Fund (no. 310030B_138662) and European Research Council grants SystemHox (no. 232790) and RegulHox (no 588029) (to D.D.) as well as National Institutes of Health grant HG003141 (to J.D.). J.D. is an investigator of the Howard Hughes Medical Institute. E.R.-C. planned the experiments, produced and analyzed the data sets, and wrote the manuscript. L.L.-D. performed bioinformatics analysis of the data sets and amended the manuscript. P.J.F. designed, produced, and analyzed the DNA-FISH data sets and amended the manuscript. L.B. produced some ChIP-seq data sets and amended the manuscript. I.E.-I. produced some WISH data sets. T.H.N.H. produced mice and managed the mouse strains. Y.Z. performed Hi-C experiments. J.D. planned the Hi-C analyses and cowrote the corresponding section of the manuscript. D.D. planned the experiments, analyzed data sets, and wrote the manuscript.

References

Andrey G, Montavon T, Mascrez B, Gonzalez F, Noordermeer D, Leleu M, Trono D, Spitz F, Duboule D. 2013. A switch between topological domains underlies HoxD genes collinearity in mouse limbs. *Science* **340**: 1234–1237.

Bachiller D, Macias A, Duboule D, Morata G. 1994. Conservation of a functional hierarchy between mammalian and insect Hox/HOM genes. *EMBO J* **13**: 1930–1941.

Beccari L, Yakushiji-Kaminatsui N, Woltering JM, Necsula A, Lonfat N, Rodriguez-Carballo E, Mascrez B, Yamamoto S, Kuroiwa A, Duboule D. 2016. A role for HOX13 proteins in the regulatory switch between TADs at the HoxD locus. *Genes Dev* **30**: 1172–1186.

Belton JM, McCord RP, Gibcus JH, Naumova N, Zhan Y, Dekker J. 2012. Hi-C: a comprehensive technique to capture the conformation of genomes. *Methods* **58**: 268–276.

Berlivet S, Paquette D, Dumouchel A, Langlais D, Dostie J, Kmita M. 2013. Clustering of tissue-specific sub-TADs accompanies the regulation of HoxA genes in developing limbs. *PLoS Genet* **9**: e1004018.

Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ, Gingeras TR, et al. 2005. Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120**: 169–181.

Bonev B, Mendelson Cohen N, Szabo Q, Fritsch L, Papadopoulos GL, Lubling Y, Xu X, Lv X, Hugnot J-P, Tanay A, et al. 2017. Multiscale 3D genome rewiring during mouse neural development. *Cell* **171**: 557–572.e24.

Charif D, Lobry JR. 2007. SeqinR 1.0-2: a contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. In *Structural approaches to sequence evolution. Biological and medical physics, biomedical engineering* (ed. Bastolla U, et al.), pp. 207–232. Springer, Berlin.

Crane E, Bian Q, McCord RP, Lajoie BR, Wheeler BS, Ralston EJ, Uzawa S, Dekker J, Meyer BJ. 2015. Condensin-driven remodeling of X chromosome topology during dosage compensation. *Nature* **523**: 240–244.

David FP, Delafontaine J, Carat S, Ross FJ, Lefebvre G, Jarosz Y, Sinclair L, Noordermeer D, Rougemont J, Leleu M. 2014. HTSstation: a Web application and open-access libraries for high-throughput sequencing data analysis. *PLoS One* **9**: e85879.

Davis AP, Witte DP, Hsieh-Li HM, Potter SS, Capecchi MR. 1995. Absence of radius and ulna in mice lacking *hoxa-11* and *hoxd-11*. *Nature* **375**: 791–795.

Delpretti S, Montavon T, Leleu M, Joye E, Tzika A, Milinkovitch M, Duboule D. 2013. Multiple enhancers regulate Hoxd genes and the Hotdog lncRNA during cecum budding. *Cell Rep* **5**: 137–150.

Deschamps J, van Nes J. 2005. Developmental regulation of the Hox genes during axial morphogenesis in the mouse. *Development* **132**: 2931–2942.

de Wit E, Vos ES, Holwerda SJ, Valdes-Quezada C, Verstegen MJ, Teunissen H, Splinter E, Wijchers PJ, Krijger PH, de Laat W. 2015. CTCF binding polarity determines chromatin looping. *Mol Cell* **60**: 676–684.

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.

Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, Ye Z, Kim A, Rajagopal N, Xie W, et al. 2015. Chromatin architecture reorganization during stem cell differentiation. *Nature* **518**: 331–336.

Dolle 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.

Duboule D, Morata G. 1994. Colinearity and functional hierarchy among genes of the homeotic complexes. *Trends Genet* **10**: 358–364.

Fabre PJ, Benke A, Joye E, Nguyen Huynh TH, Manley S, Duboule D. 2015. Nanoscale spatial organization of the HoxD gene cluster in distinct transcriptional states. *Proc Natl Acad Sci* **112**: 13964–13969.

Fabre PJ, Leleu M, Mormann BH, Lopez-Delisle L, Noordermeer D, Beccari L, Duboule D. 2017. Large scale genomic reorganization of topological domains at the HoxD locus. *Genome Biol* **18**: 149.

- Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N, Mirny LA. 2016. Formation of chromosomal domains by loop extrusion. *Cell Rep* **15**: 2038–2049.
- Ghirlando R, Felsenfeld G. 2016. CTCF: making the right connections. *Genes Dev* **30**: 881–891.
- Giorgetti L, Lajoie BR, Carter AC, Attia M, Zhan Y, Xu J, Chen CJ, Kaplan N, Chang HY, Heard E, et al. 2016. Structural organization of the inactive X chromosome in the mouse. *Nature* **535**: 575–579.
- Gómez-Marín C, Tena JJ, Acemel RD, López-Mayorga M, Naranzo S, de la Calle-Mustienes E, Maeso I, Beccari L, Aneas I, Vielmas E, et al. 2015. Evolutionary comparison reveals that diverging CTCF sites are signatures of ancestral topological associating domains borders. *Proc Natl Acad Sci* **112**: 7542–7547.
- Gonzalez F, Duboule D, Spitz F. 2007. Transgenic analysis of Hoxd gene regulation during digit development. *Dev Biol* **306**: 847–859.
- Gonzalez-Reyes A, Urquia N, Gehring WJ, Struhl G, Morata G. 1990. Are cross-regulatory interactions between homoeotic genes functionally significant? *Nature* **344**: 78–80.
- Gonzalez-Sandoval A, Gasser SM. 2016. On TADs and LADs: spatial control over gene expression. *Trends Genet* **32**: 485–495.
- Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W, et al. 2008. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453**: 948–951.
- Guerreiro I, Gitto S, Novoa A, Codourey J, Nguyen Huynh TH, Gonzalez F, Milinkovitch MC, Mallo M, Duboule D. 2016. Reorganisation of Hoxd regulatory landscapes during the evolution of a snake-like body plan. *Elife* **5**: e16087.
- Guo Y, Xu Q, Canzio D, Shou J, Li J, Gorkin DU, Jung I, Wu H, Zhai Y, Tang Y, et al. 2015. CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell* **162**: 900–910.
- Haarhuis JHI, van der Weide RH, Blomen VA, Yáñez-Cuna JO, Amendola M, van Ruiten MS, Krijger PHL, Teunissen H, Medema RH, van Steensel B, et al. 2017. The cohesin release factor WAPL restricts chromatin loop extension. *Cell* **169**: 693–707.e14.
- Herault Y, Fraudeau N, Zakany J, Duboule D. 1997. Ulnaless (Ul), a regulatory mutation inducing both loss-of-function and gain-of-function of posterior *Hoxd* genes. *Development* **124**: 3493–3500.
- Herault Y, Rassoulzadegan M, Cuzin F, Duboule D. 1998. Engineering chromosomes in mice through targeted meiotic recombination (TAMERE). *Nat Genet* **20**: 381–384.
- Hnisz D, Day DS, Young RA. 2016. Insulated neighborhoods: structural and functional units of mammalian gene control. *Cell* **167**: 1188–1200.
- Imakaev M, Fudenberg G, McCord RP, Naumova N, Goloborodko A, Lajoie BR, Dekker J, Mirny LA. 2012. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat Methods* **9**: 999–1003.
- Ji X, Dadon DB, Powell BE, Fan ZP, Borges-Rivera D, Shachar S, Weintraub AS, Hnisz D, Pegoraro G, Lee TI, et al. 2016. 3D chromosome regulatory landscape of human pluripotent cells. *Cell Stem Cell* **18**: 262–275.
- Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, et al. 2010. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**: 430–435.
- Kieffer-Kwon KR, Tang Z, Mathe E, Qian J, Sung MH, Li G, Resch W, Baek S, Pruetz N, Grontved L, et al. 2013. Interactome maps of mouse gene regulatory domains reveal basic principles of transcriptional regulation. *Cell* **155**: 1507–1520.
- Kmita M. 2002. Evolutionary conserved sequences are required for the insulation of the vertebrate *Hoxd* complex in neural cells. *Development* **129**: 5521–5528.
- Kundu S, Ji F, Sunwoo H, Jain G, Lee JT, Sadreyev RI, Dekker J, Kingston RE. 2017. Polycomb repressive complex 1 generates discrete compacted domains that change during differentiation. *Mol Cell* **65**: 432–446.e5.
- Lajoie BR, Dekker J, Kaplan N. 2015. The Hitchhiker's guide to Hi-C analysis: practical guidelines. *Methods* **72**: 65–75.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359.
- Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25.
- Li G, Ruan X, Auerbach RK, Sandhu KS, Zheng M, Wang P, Poh HM, Goh Y, Lim J, Zhang J, et al. 2012. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* **148**: 84–98.
- Lonfat N, Montavon T, Darbellay F, Gitto S, Duboule D. 2014. Convergent evolution of complex regulatory landscapes and pleiotropy at Hox loci. *Science* **346**: 1004–1006.
- Lupianez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, Horn D, Kayserili H, Opitz JM, Laxova R, et al. 2015. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**: 1012–1025.
- Mallo M, Wellik DM, Deschamps J. 2010. Hox genes and regional patterning of the vertebrate body plan. *Dev Biol* **344**: 7–15.
- Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.J* **17**: 10–12.
- 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.
- Montavon T, Soshnikova N, Mascrez 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.
- Morey C, Da Silva NR, Perry P, Bickmore WA. 2007. Nuclear reorganisation and chromatin decondensation are conserved, but distinct, mechanisms linked to Hox gene activation. *Development* **134**: 909–919.
- Mousavi K, Zare H, Dell'orso S, Grontved L, Gutierrez-Cruz G, Derfoul A, Hager GL, Sartorelli V. 2013. eRNAs promote transcription by establishing chromatin accessibility at defined genomic loci. *Mol Cell* **51**: 606–617.
- Narendra V, Rocha PP, An D, Raviram R, Skok JA, Mazzoni EO, Reinberg D. 2015. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science* **347**: 1017–1021.
- 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.
- Noordermeer D, Leleu M, Schorderet P, Joye E, Chabaud F, Duboule D. 2014. Temporal dynamics and developmental memory of 3D chromatin architecture at Hox gene loci. *Elife* **3**: e02557.
- Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Pilot T, van Berkum NL, Meisig J, Sedat J, et al. 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**: 381–385.
- Nora EP, Goloborodko A, Valton AL, Gibcus JH, Uebersohn A, Abdennur N, Dekker J, Mirny LA, Bruneau BG. 2017. Targeted degradation of CTCF decouples local insulation of

- chromosome domains from genomic compartmentalization. *Cell* **169**: 930–944.e22.
- Phillips-Cremins JE, Sauria ME, Sanyal A, Gerasimova TI, Lajoie BR, Bell JS, Ong CT, Hookway TA, Guo C, Sun Y, et al. 2013. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* **153**: 1281–1295.
- Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, et al. 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**: 1665–1680.
- Rao SSP, Huang S-C, Glenn St Hilaire B, Engreitz JM, Perez EM, Kieffer-Kwon K-R, Sanborn AL, Johnstone SE, Bascom GD, Bochkov ID, et al. 2017. Cohesin loss eliminates all loop domains. *Cell* **171**: 305–320.e24.
- Sanborn AL, Rao SSP, Huang S-C, Durand NC, Huntley MH, Jewett AI, Bochkov ID, Chinnappan D, Cutkosky A, Li J, et al. 2015. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc Natl Acad Sci* **112**: E6456–E6465.
- Sanyal A, Lajoie BR, Jain G, Dekker J. 2012. The long-range interaction landscape of gene promoters. *Nature* **489**: 109–113.
- Schep R, Necsulea A, Rodriguez-Carballo E, Guerreiro I, Andrey G, Nguyen Huynh TH, Marcet V, Zakany J, Duboule D, Beccheri L. 2016. Control of Hoxd gene transcription in the mammary bud by hijacking a preexisting regulatory landscape. *Proc Natl Acad Sci* **113**: E7720–E7729.
- Schmidl C, Rendeiro AF, Sheffield NC, Bock C. 2015. ChIPmentation: fast, robust, low-input ChIP-seq for histones and transcription factors. *Nat Methods* **12**: 963–965.
- Schwarzer W, Abdennur N, Goloborodko A, Pekowska A, Fudenberg G, Loe-Mie Y, Fonseca NA, Huber W, Haering C, Mirny L, et al. 2017. Two independent modes of chromatin organization revealed by cohesin removal. *Nature* **551**: 51–56.
- Shin H, Shi Y, Dai C, Tjong H, Gong K, Alber F, Zhou XJ. 2016. TopDom: an efficient and deterministic method for identifying topological domains in genomes. *Nucleic Acids Res* **44**: e70.
- Sofueva S, Yaffe E, Chan WC, Georgopoulou D, Vietri Rudan M, Mira-Bontenbal H, Pollard SM, Schroth GP, Tanay A, Hadjur S. 2013. Cohesin-mediated interactions organize chromosomal domain architecture. *EMBO J* **32**: 3119–3129.
- Soshnikova N, Montavon T, Leleu M, Galjart N, Duboule D. 2010. Functional analysis of CTCF during mammalian limb development. *Dev Cell* **19**: 819–830.
- 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.
- Spitz F, Gonzalez F, Duboule D. 2003. A global control region defines a chromosomal regulatory landscape containing the HoxD cluster. *Cell* **113**: 405–417.
- Symmons O, Uslu VV, Tsujimura T, Ruf S, Nassari S, Schwarzer W, Ettwiller L, Spitz F. 2014. Functional and topological characteristics of mammalian regulatory domains. *Genome Res* **24**: 390–400.
- Tang Z, Luo OJ, Li X, Zheng M, Zhu JJ, Szalaj P, Trzaskoma P, Magalska A, Wlodarczyk J, Rusczycki B, et al. 2015. CTCF-mediated human 3D genome architecture reveals chromatin topology for transcription. *Cell* **163**: 1611–1627.
- Tschopp P, Duboule D. 2011. A regulatory ‘landscape effect’ over the HoxD cluster. *Dev Biol* **351**: 288–96.
- Tschopp P, Duboule D. 2014. The genetics of murine Hox loci: TAMERE, STRING, and PANTHERE to engineer chromosome variants. *Methods Mol Biol* **1196**: 89–102.
- van der Hoeven F, Zákány J, Duboule D. 1996. Gene transpositions in the HoxD complex reveal a hierarchy of regulatory controls. *Cell* **85**: 1025–1035.
- Vietri Rudan M, Barrington C, Henderson S, Ernst C, Odom DT, Tanay A, Hadjur S. 2015. Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell Rep* **10**: 1297–1309.
- Villavicencio-Lorini P, Kuss P, Friedrich J, Haupt J, Farooq M, Turkmen S, Duboule D, Hecht J, Mundlos S. 2010. Homeobox genes d11–d13 and a13 control mouse autopod cortical bone and joint formation. *J Clin Invest* **120**: 1994–2004.
- Williamson I, Berlivet S, Eskeland R, Boyle S, Illingworth RS, Paquette D, Dostie J, Bickmore WA. 2014. Spatial genome organization: contrasting views from chromosome conformation capture and fluorescence in situ hybridization. *Genes Dev* **28**: 2778–2791.
- Woltering JM, Duboule D. 2010. The origin of digits: expression patterns versus regulatory mechanisms. *Dev Cell* **18**: 526–532.
- Woltering JM, Vonk FJ, Muller H, Bardine N, Tuduze IL, de Bakker MA, Knochel W, Sirbu IO, Durston AJ, Richardson MK. 2009. Axial patterning in snakes and caecilians: evidence for an alternative interpretation of the Hox code. *Dev Biol* **332**: 82–89.
- Woltering JM, Noordermeer D, Leleu M, Duboule D. 2014. Conservation and divergence of regulatory strategies at Hox Loci and the origin of tetrapod digits. *PLoS Biol* **12**: e1001773.
- Yekta S, Tabin CJ, Bartel DP. 2008. MicroRNAs in the Hox network: an apparent link to posterior prevalence. *Nat Rev Genet* **9**: 789–796.
- Young T, Rowland JE, van de Ven C, Bialecka M, Novoa A, Carapuco M, van Nes J, de Graaff W, Duluc I, Freund JN, et al. 2009. Cdx and Hox genes differentially regulate posterior axial growth in mammalian embryos. *Dev Cell* **17**: 516–526.
- Zakany J, Duboule D. 2007. The role of Hox genes during vertebrate limb development. *Curr Opin Genet Dev* **17**: 359–366.
- Zakany J, Kmita M, Duboule D. 2004. A dual role for Hox genes in limb anterior–posterior asymmetry. *Science* **304**: 1669–1672.
- Zhan Y, Mariani L, Barozzi I, Schulz EG, Bluthgen N, Stadler M, Tiana G, Giorgetti L. 2017. Reciprocal insulation analysis of Hi-C data shows that TADs represent a functionally but not structurally privileged scale in the hierarchical folding of chromosomes. *Genome Res* **27**: 479–490.
- Zhang Y, Liu T, Meyer CA, Eickhout J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**: R137.
- Zuin J, Dixon JR, van der Reijden MI, Ye Z, Kolovos P, Brouwer RW, van de Corput MP, van de Werken HJ, Knoch TA, van Ijcken WF, et al. 2014. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc Natl Acad Sci* **111**: 996–1001.