Chromatin topology and the timing of enhancer function at the HoxD locus

Eddie Rodríguez-Carballo*,1,2, Lucille Lopez-Delisle*, Andréa Willeming*, Leonardo Beccari*, Sandra Gitto*, Bénédict Mascrez*, and Denis Duboule*a,b,c,1

*Department of Genetics and Evolution, University of Geneva, 1211 Geneva, Switzerland; †School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland; and ‡Collège de France, 75005 Paris, France

Contributed by Denis Duboule, October 26, 2020 (sent for review July 17, 2020; reviewed by Amos Tanay and Steven A. Vokes)

The HoxD gene cluster is critical for proper limb formation in tetrapods. In the emerging limb buds, different subgroups of Hoxd genes respond first to a proximal regulatory signal, then to a distal signal that organizes digits. These two regulations are exclusive from one another and emanate from two distinct topologically associating domains (TADs) flanking HoxD, both containing a range of appropriate enhancer sequences. The telomeric TAD (T-DOM) contains several enhancers active in presumptive forearm cells and is divided into two sub-TADs separated by a CTCF-rich boundary, which defines two regulatory submodules. To understand the importance of this particular regulatory topology to control Hoxd gene transcription in time and space, we either deleted or inverted this sub-TAD boundary, eliminated the CTCF binding sites, or inverted the entire T-DOM to exchange the respective positions of the two sub-TADs. The effects of such perturbations on the transcriptional regulation of Hoxd genes illustrate the requirement of this regulatory topology for the precise timing of gene activation. However, the spatial distribution of transcripts was eventually resumed, showing that the presence of enhancer sequences, rather than either their exact topology or a particular chromatin architecture, is the key factor. We also show that the affinity of enhancers to find their natural target genes can overcome the presence of both a strong TAD border and an unfavorable orientation of CTCF sites.

Significance

Many genes important for vertebrate development are surrounded by multiple series of remote enhancer sequences. Such regulatory landscapes and their target genes are usually located within the same chromatin domains, which appears to constrain the action of these regulatory sequences and hence to facilitate enhancer–promoter recognition and gene expression. We used the HoxD locus to assess the impact of modifying the regulatory topology upon gene activation in space and time. A series of chromosomal rearrangements involving deletions and inversions reveals that the enhancer topology plays a role in the timing of gene activation. However, gene expression was often recovered subsequently, illustrating the intrinsic capacity of some enhancers to find their target promoters despite an apparently adverse chromatin topology.

Author contributions: E.R.-C. and D.D. designed research; E.R.-C., A.W., L.B., S.G., and B.M. performed research; L.B. contributed new reagents/analytic tools; E.R.-C., L.L.-D., A.W., and B.M. analyzed data; and E.R.-C., L.L.-D., and D.D. wrote the paper.

Reviews: A.T., Weizmann Institute; and S.A.V., The University of Texas at Austin. The authors declare no competing interest.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

*To whom correspondence may be addressed. Email: edgardo.rodriguez@unige.ch or denis.duboule@unige.ch.

†Present address: Department of Molecular Biology, University of Geneva, 1211 Geneva, Switzerland.

‡Present address: Institut Neuromyogène, CNRS UMR 5310-INSERM U1217, Université de Lyon, 69008 Lyon, France.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2015083117/-/DCSupplemental. First published November 23, 2020.

www.pnas.org/cgi/doi/10.1073/pnas.2015083117

PNAS | December 8, 2020 | vol. 117 | no. 49 | 31231–31241
In the early limb bud, T-DOM is activated at E9.0 (embryonic day 9), leading to the first wave of colinear transcription, coinciding with the establishment of chromatin interactions between the newly activated genes (Hoxd9 to Hoxd11) and part of T-DOM (21). At E12.5, cells transcribing these genes are found in the proximal part of the limb buds, which will generate the arm and the forearm. At this stage, the distribution of interactions with T-DOM shows a clear topological segregation, with 3′-located genes (Hoxd1 to Hoxd8) interacting mostly with the first sub-TAD, whereas 5′-located genes (Hoxd9 to Hoxd11) associate in priority with the more distant sub-TAD (21, 22), suggesting a functional compartmentalization of T-DOM. All limb-specific enhancers were thus far associated to the distant sub-TAD, starting at the sub-TAD boundary and extending up to Hnrnpa3, including the CS65 and CS93 enhancers (21, 28).

In this work, we set up to assess whether a correlation exists between the precise temporal and spatial transcriptional activation of Hoxd genes in proximal limbs on the one hand, and a fine topological organization of its regulatory landscape on the other hand, or whether the mere presence of series of enhancers within T-DOM is necessary for HoxD regulation, regardless of their intrinsic organization. We show that, while the overall chromatin architecture determines the correct timing of gene activation, enhancer–promoter communication can be successfully established along with limb bud development, even after the engineering of major topological modifications, including the positioning of a strong TAD border in between them.

**Results**

**Multiple Early Limb Enhancers in T-DOM.** Hi-C profiles from several cell types have previously revealed that the HoxD cluster is positioned at the boundary between two TADs. T-DOM (i.e., the TAD located telomeric to the gene cluster) is necessary for the transcription of Hoxd genes, both during limb budding and, subsequently, in the formation of the proximal segment of the prospective arm. Instead, the C-DOM controls Hoxd gene expression in developing digits, at later time points (Fig. L4). From E9.5 to E12.5, T-DOM shows specific activation and decommissioning dynamics (21), which correlates with its 3D conformation, as only the more distant T-DOM sub-TAD (Fig. L4,
sub-TAD2) remains active at late (E12.5) embryonic stages. Most limb enhancers described thus far are located within this chromatin domain, in particular the CS39, CS65, and CS93 sequences (21, 28, 29).

In order to characterize the onset of activation of T-DOM in the incipient limb bud, we microdissected E9 forelimb buds and pooled them into two groups corresponding to embryos either between 18 and 22 somites (or early E9), or between 24 and 29 somites (or late E9). Chromatin immunoprecipitation (ChIP) of H3K27ac, a histone mark associated with enhancer activity and gene expression, revealed that most of the acetylated regions were located in sub-TAD2, which seemed particularly active in 24- to 29-somite-old limb buds (Fig. 1B). Two H3K27ac-positive regions were nevertheless identified in sub-TAD1 in E9 limb buds, which were not present in E12.5 proximal limb cells. In contrast to CS39, CS65, and CS93, however, these two early limb control regions (ELCR2 and ELCR3) were not found fully conserved in chicken, albeit they are present in all mammals (SI Appendix, Fig. S1) (21, 28). Transgenic analysis of both ELCR2 and ELCR3 showed strong LacZ expression in E9 limb buds, which coincides with the expression of CS39 and CS65 transgenes (Fig. 1C, arrowheads), as well as in other mesoderm derivatives (Fig. 1C, arrows).

To evaluate potential changes in the global architecture of T-DOM along with developmental time, we looked at the binding profiles of both CTCF and the cohesin subunit RAD21. The ChIP profiles of these architectural proteins using limb buds from 20- to 28-somite embryos did not substantially differ from the profiles obtained in E12.5 proximal limb (Fig. 1B; and figure 2 A and B of ref. 22). Most of the CTCF binding sites had a convergent orientation in relation to the HoxD cluster, including the three bound CTCFs found within the CS38-40 region, the boundary region that divides T-DOM into its two sub-TADs (Fig. 1B, arrow).

Deletion of the T-DOM Sub-TAD Boundary. We asked whether such a partitioning of T-DOM into two subdomains was mandatory for this early limb bud regulation to be properly implemented. We tried to merge both domains by deleting the CS38-40 region (SI Appendix, Fig. S2A), which contains three CTCF binding sites as well as the CS39 limb enhancer and the transcription start site (TSS) of the Hog and Tog long noncoding RNA (lncRNAs) (22, 26), which coincides with a CpG island. We performed circular chromosome conformation capture sequencing (4C-seq) experiments in E9.5 mutant forelimb buds using several viewpoints distributed both along T-DOM and inside the HoxD cluster (Fig. 2 and SI Appendix, Fig. S2B). We then compiled all data from the different viewpoints into a virtual Hi-C matrix using the 4Cin software (30), which we adapted to plot the relative distances in a linear manner according to the real genomic coordinates. Because the 4Cin processing has an inherent variability that leads to the generation of different models, we assessed the correlation between 20 iterations and clustered them, thus displaying a merged average (Materials and Methods).

When applied to control limbs, the 4Cin approach generated a map of computed distances at T-DOM that resembled the expected contact distribution of a wild-type Hi-C matrix, including the subdivision of the domain in two sub-TADs, as well as specific contacts between CTCF-bound regions and enhancer–promoter interactions (Fig. 2A, arrows and arrowheads, respectively). Using the same viewpoints, we confirmed that the deletion of region CS38-40 affected the spatial organization of this regulatory domain. We observed a substantial increase in the interactions established between the two sub-TADs, with a 17% decrease in computed distances between the two subdomains \( P = 3e−49 \) leading to their fusion into a single domain (Fig. 2B). This was accompanied by an apparent loosening of the interactions inside each subdomain, seen as an increase in distances within sub-TAD1 (12%; \( P = 0.008 \)) and sub-TAD2 (13%; \( P = 2e−12 \)) (Fig. 2B, Inset). The increase in contacts between the two sub-TADs could be observed when comparing any of the derived cluster representations (SI Appendix, Fig. S3 A and B). Also, the interactions established by the HoxD cluster throughout the regulatory domain seemed to decrease (Fig. 2 A and B, arrowheads).

More specifically, we analyzed the interaction profile of the Hoxd11 gene, whose expression in the posterior part of the E9.5 developing limb bud is maintained until E12.5, in the proximal
limb. The 4C-seq data revealed that, in wild-type E9.5 limb buds, Hoxd11 strongly interacted with both region CS38-40 and the more distant sub-TAD2 (Fig. 2C). Upon deletion of the sub-TAD border, a modest increase in interactions was detected in the bound CTCF sites located 3′ to region CS38-40 and at the telomeric TAD border close to Hmmpa3 (Fig. 2C, arrows). However, the contacts did not increase substantially along the region initially corresponding to sub-TAD2. Instead, interactions of Hoxd11 were reduced with the CS93 and CS65 limb enhancers, representing 65% and 81%, respectively, when compared to wild type (Fig. 2C, open arrowheads and SI Appendix, Table S1). These changes were specific to Hoxd11 as neither Hoxd9 nor Hoxd4 showed the same trend regarding enhancer interactions (SI Appendix, Fig. S2B and Table S1).

We assessed whether these alterations in contact distribution translated into changes in gene expression pattern. Whole-mount RNA in situ hybridization (WISH) showed a slight but visible decrease in Hoxd11 expression at E9.5 (Fig. 2D). To verify this observation, we performed qRT-PCR on forelimb buds dissected from embryos aged between 20 and 28 somites and plotted their individual values (Fig. 2E). The dynamics of Hoxd11 expression in control forelimb buds followed a strong increase right after the 24-somite stage. On the contrary, this dynamics in Hoxd11 messenger RNA (mRNA) was not observed in the mutant limb buds where the increase was not as fast (Fig. 2E). This was further confirmed by RNA-seq experiments showing that Hoxd10 and Hoxd12 had a delayed onset of transcription while more anterior genes (i.e., Hoxd4) did not seem to be affected at all (SI Appendix, Fig. S3C), in agreement with the lack of modification in enhancer contacts for these genes (SI Appendix, Fig. S2B). Altogether, the specific down-regulation of Hoxd11 could either be a consequence of the distinct spatial reorganization of T-DOM or be due to the removal of the CS39 enhancer.

To explore these possibilities, we used a CRISPR/Cas9 approach to eliminate the binding of CTCF to the three motives positioned within region CS38-40. We initially deleted 26 base pairs (bp) of the CTCF binding site located in region CS38 [delCTCF(CS38)] (SI Appendix, Fig. S4A), preserving both the neighboring CpG island and the TSS of the Hox and Tog IncRNAs (SI Appendix, Fig. S24). On top of this first editing, we generated a 1.5-kilobase (kb) large deletion that removed the two binding sites located around CS40, without removing the H3K27ac-enriched region localized around CS39 [delCTCFs(CS38:CS40)] (SI Appendix, Figs. S24 and S4A). We confirmed by ChIP that CTCF binding was no longer detected at any of these locations or elsewhere in this short DNA interval (SI Appendix, Fig. S24). The deletion of the three CTCF sites led to a merge of the sub-TADs at E12.5 (SI Appendix, Figs. S4 B and C), thus confirming the importance of these bond proteins in the establishment of this specific topological structure. We analyzed gene expression and observed that Hoxd11 was briefly delayed in its activation, a lag that was rapidly resumed to generate a late pattern indistinguishable from wild type (SI Appendix, Fig. S4D). These results indicated that the merging of the two T-DOM sub-TADs moderately affected the onset of Hoxd gene expression in early limb buds, with a stronger effect observed in the absence of the CS39 enhancer.

Reinforcing a Sub-TAD Separation. We next engineered the opposite situation to produce a more robust separation between the two sub-TADs such as to isolate them from one another as bona fide TADs. This was achieved by generating an inversion of the region comprising the three CTCF binding sites [the inv(CS38-40) allele]. In this configuration, the three CTCF sites now converged toward the strong telomeric TAD border at the 3′ end of the domain. In this allele, the three CTCF sites were still occupied, as expected (SI Appendix, Fig. S5A). A virtual Hi-C pattern of limb bud cells dissected from this mutant stock expectedly revealed that the inversion of this region had strengthened the segregation of the two sub-TADs. Indeed, a 39% increase in distances was observed (P value = 2.45e−163) while the compaction within sub-TAD1 decreased by 27% (P value = 1e−7) (Fig. 3 A and B, Inset). Concomitantly, it reduced the general contacts of the HoxD cluster with region CS38-40, as well as between some Hoxd genes and the CS93 and CS65 enhancers (Fig. 3 A and B and SI Appendix, Fig. S5B and Table S1). This was particularly well illustrated by using Hoxd11 as a 4C-seq viewpoint, showing a reduction of interactions over regions CS38-40, CS93, and CS65 (Fig. 3C), similar to what had been noted in the del(CS38-40) allele (Fig. 2C). These topological changes also correlated with a delay of Hoxd11 expression (Fig. 3 D and E), which was stronger than in the CTCF mutant

Fig. 3. Inversion of region CS38-40. (A and B) Virtual Hi-C maps of wild-type (A) and the inv(CS38-40) allele (B) from E9 forelimb buds, as reconstituted from several 4C-seq viewpoints (gray diamonds). The triangles showing CTCF orientation of region CS38-40 are inverted in the mutant. Arrowheads and arrows mark the position of some enhancers and CTCF binding sites, respectively. The Inset in B shows the percentage of the changes (mutant vs. wt) in distances inside and between sub-TADs. (C) The 4C-seq profiles of Hoxd11 for wt and mutant E9 forelimb buds (blue and red lines, respectively). The inverted region is shown as a dashed box around CS39. Open arrowheads show areas of decreased interaction in the inv(CS38-40) allele. Arrows indicate the position of some CTCF binding sites. (D) WISH analysis of Hoxd11 in E9 (18 to 22 somites) forelimb buds (earlier time point on the Left) (8x magnification). (E) Comparison of qRT-PCR values (wt n = 16 and mutant n = 17) of Hoxd11 in E9 forelimb buds at different somite stages. An exponential fit is represented out of the real qRT-PCR values.
alleles. This delay was nevertheless not pervasive for it did not affect all Hoxd mRNAs equally. For example, Hoxd9 did not show a clear transcriptional decrease, even at early stages (SI Appendix, Fig. S5), whereas more “posterior” genes (like Hoxd11) seemed to be more affected. All these delays, however, were subsequently resumed, and, in E12.5 forelimb buds, changes in expression patterns could hardly be scored when comparing the inv(CS58-40) allele to wild-type littermates (SI Appendix, Fig. SSD).

Inversion of T-DOM. Altogether, these genomic alterations did not produce long-lasting effects upon the transcription of Hoxd genes in limb buds. One possible explanation is that T-DOM contains several other CTCF binding sites, most of them displaying an orientation convergent to those numerous sites present in the telomeric part of the HoxD cluster itself. In this context, it is possible that such CTCF sites within T-DOM may assist remote enhancers reaching targets, regardless of small rearrangements occurring at their vicinity. We thus inverted the entire T-DOM and produced two novel configurations, one containing a strong TAD border between the inverted T-DOM and the HoxD cluster, and the other one lacking this TAD border (Figs. 4 and 5, Bd).

The inversion of T-DOM was induced by targeting CRISPR guides at both sides of this regulatory domain. To be as inclusive as possible, the break points of this inversion were selected close to the 3′ end of the Hoxd1 gene and 5 kb upstream of the TSS of the Hm transactional control element (TCE) gene, respectively (Fig. 4A and B, dashed lines). Due to the position of the latter break point, the Hmdependence3 TAD border (Fig. 4A, Bd) was inverted along with T-DOM and placed just between the HoxD cluster and the inverted T-DOM. Upon inversion, a substantial loss of contacts was scored by using 4C-seq profiles (Fig. 4B), likely due to the weak interactions with T-DOM enhancers. This strong variation in the enhancer tropism over chromatin topology.

To see whether the introduction of a strong new TAD border between the gene cluster and T-DOM, we further deleted the boundary region on top of the inverted allele to produce the inv(T-DOM)del(Bd) mutant line (Fig. 5). The deleted 20-kb-large boundary region, which

![Fig. 4. Inversion of T-DOM in the presence of a TAD border.](Image)

(A and B) Virtual Hi-C maps of wild-type (A) and inv(T-DOM) mutant mice (B) from E9 forelimb buds, as reconstituted from several 4C-seq viewpoints (gray diamonds). Dashed lines indicate the inverted region. The distant boundary is marked as an empty box (Bd) at the end of T-DOM in the wild type and close to the HoxD cluster in the mutant allele. The CTCF orientations in the inv(T-DOM) allele are inverted accordingly. In the mutant allele, open arrowheads represent lost interactions that can be scored in the wt (filled arrowheads). Arrows indicate the position of some CTCF binding sites, which are located close to the HoxD cluster after the inversion of T-DOM. (C) qRT-PCR values (wt n = 17 and mutant allele n = 21) of Hoxd11 in E9 forelimb buds at different somite stages. An exponential fit is represented out of the real qRT-PCR values. (D) WISH images of forelimb buds at E9.5 (Top, ~20 somites, 6.3x magnification) and E10.5 (Bottom, 4x magnification) of wt and the inv(T-DOM) mutant allele. The loss of expression in the mutant is indicated as open arrowheads.
Fig. 5. Inversion of T-DOM in the absence of a TAD border. (A) Scheme of the rearranged HoxD locus after inversion of T-DOM. The orientation of CTCF sites is represented. WISH images of Hoxd11 and Hoxd9 expression pattern in E12.5 forelimbs are shown for both wt and inv(T-DOM) mutant embryos. An open arrowhead demarcates the loss of expression in the anterior part of the proximal limb domain for both Hoxd9 and Hoxd11 (3.2× magnification). (B) Scheme showing the deletion of the TAD boundary (Bd) on top of the inv(T-DOM) allele. WISH of Hoxd11 and Hoxd9 using E12.5 wt and inv(T-DOM)del(Bd)-inv(T)del(Bd) mutant forelimbs shows the same patterns with the reappearance of the anterior domain missing in the inv(T-DOM) allele (arrowhead) (3.2× magnification). (C) Virtual Hi-C maps of E12 proximal limb buds using either inv(T-DOM), or inv(T-DOM)del(Bd) mutant limb bud cells. Both matrices were mapped on the artificial inv(T-DOM) genome. Dashed vertical lines demarcate the extension of the inverted region in both mutant lines. Filled arrowheads indicate the increase in contacts with enhancers CS65, CS93, and CS39 [indicated with open arrowheads in inv(T-DOM)] after the deletion of the TAD boundary (Bd).
normally tightly isolates T-DOM from its more telomeric TAD, contained three CTCF binding sites. Fetuses carrying this additional deletion fully recovered wild-type expression patterns for both Hoxd9 and Hoxd11, with expression domains in the limb buds undistinguishable from their wild-type counterparts (Fig. 5B). In particular, the proximal-anterior transcript domain lacking in inv(T-DOM) embryos (Fig. 5A, open arrowheads) was fully rescued after deletion of the ectopic TAD border (Fig. 5B, filled arrowheads).

This recovery in expression was concomitant to a clear increase in interactions between the HoxD cluster and various T-DOM limb enhancers when looking both at the virtual Hi-C matrices (Fig. 5C and SI Appendix, Fig. S7A) and to the Hoxd gene 4C-seq profiles (SI Appendix, Fig. S7B and Table S1). The reestablishment in the spatial deployment of transcripts at day E12.5 was, however, not observed at the earliest stages analyzed, which still showed an important time lag in target gene activation, even though the ectopic TAD border had been removed (SI Appendix, Fig. S7C). These results showed that the mere inversion of the regulatory domain had an impact upon the onset of Hoxd expression. The observed delay could nevertheless be caught up in a few days, a recovery that was not completely possible when the telomeric TAD border was present between the enhancers and the target Hoxd genes.

**Discussion**

The fine-tuned regulation of genes involved in developmental processes is often achieved by complex regulatory landscapes, which can extend up to megabases around the target gene(s). Such regulatory landscapes generally match the extents of TADs and contain all of the enhancers necessary for the various expression specificities. Even though a clear causal relationship is difficult to establish, the prevalent model is that TADs somehow restrict the sphere of operation for such regulations by providing a spatial unit where genes can be properly controlled, in isolation from their neighbors. The action of enhancers is thought to depend on their 3D spatial proximity to the target promoters they regulate, a hypothesis supported by several lines of evidence (reviewed in refs. 1 and 2). Recent reports, however, have challenged this view, showing that transcriptional activity does not always correlate with a direct promoter–enhancer physical interaction (31–33). In this work, we used the HoxD locus and one of its two flanking regulatory landscapes as a paradigm to look at the effect of modifying the regulatory topology upon the precisely orchestrated transcription of this series of genes. We engineered several rearrangements within the regulatory domain to determine the impact of both the distribution of enhancer sequences, the presence and orientation of CTCF binding sites, and the ectopic introduction of a TAD border between the promoters and the corresponding enhancers. We conclude that, while the global TAD architecture may serve to properly implement the regulatory modalities in time, major rearrangements do not critically modify the regulatory outcome at a later stage, making enhancer–promoter contacts very resilient and somewhat poorly dependent from the architectural context.

**A Split Regulatory Landscape.** T-DOM is normally divided into two sub-TADs at the level of region CS38-40 (21), a region that contains three CTCF sites with an orientation convergent to that of numerous sites within the HoxD cluster (22). The deletion of this border region expectedly led to the fusion of the two sub-TADs. However, rather than reenforcing contacts between enhancers and promoters in the de novo created single TAD, enhancer–promoter contacts tended to decrease. Therefore, the presence of these two subdomains within T-DOM favors maximal efficiency in the regulatory outcome (see summary scheme in Fig. 6). One potential explanation is that it is not the global structure itself that is important but, instead, the presence of three CTCF binding sites that may trigger part of the necessary interactions, in particular due to their shared orientation toward the HoxD cluster. In this deleted allele, Hoxd genes were expressed rather normally, but with a clear delay in their activation.

A more precise deletion strategy removing these three CTCF sites led to a similar fusion between the subdomains. However, the effect upon Hoxd gene transcription was even milder than in the deletion of the boundary, likely because the full boundary deletion also included the CS39 enhancer, which was left in place in the CTCF deletion allele. In the latter case, mRNA accumulation was also delayed, but even less than in the first allele. Therefore, it seems that the presence of these CTCF sites, rather than the global structure that they help to organize, as well as the full collection of limb enhancers are the key elements to properly activate the target genes in time.

The importance of CTCF sites and/or of their orientation for chromatin interactions was previously predicted in silico and illustrated experimentally at a variety of specific loci (20, 34, 35). In developing tissues, the presence of bound CTCF in a specific...
locus favored the ectopic action of some clustered enhancers when placed in a different TAD (36). Here, we show that, when the entire boundary region containing the three sites was inver-
ted, the isolation of the two sub-TADs became much stronger
due to the convergence of these inverted sites with the natural
telomeric 3′ TAD border, giving rise to two qualified TADs.
Despite this accentuated split, which further isolated many limb
enhancers from the target Hoxd genes, the transcriptional impact
was once again restricted to the onset of their expression, similar
to the effect of deleting the sub-TAD boundary. These results
confirmed that the partition of T-DOM into sub-TADs may not
respond to any particular regulatory necessity, at least in limb
cells. Instead, it may be a consequence of CTCF being engaged
in facilitating enhancer–promoter contacts.

Enhancer Topology and Regulatory Heterochrony. It is often argued
that chromatin architecture is instrumental to ensure the proper
temporal dynamics of gene activation (see, e.g., ref. 5). However,
in the case where regulatory landscapes contain multiple en-
hancers with either identical (37) or related (7, 8) specificities, it
is less clear whether the respective positions of these enhancers
and their distances to one another are critical factors for target
gene activation. T-DOM contains multiple limb bud enhancers
over 800 kb, which tend to be distributed far from the gene
cluster, interspersed with CTCF sites. The large engineered in-
version of T-DOM lacking the TAD border gave us a rather
clear answer to this question, at least regarding this particular
locus. While Hoxd genes were importantly delayed in their ac-
activation at E9.5, their expression patterns at E12.5 were indis-
tinguishable from control limb buds. This once more points to
the separation between two distinct regulatory aspects. On the
one hand, the full series of enhancers will end up delivering their
integrated information, regardless of their global organization
within the landscape. On the other hand, an appropriate order
and/or chromatin organization will help to properly orchestrate
this process.

While similar chromosomal rearrangements have been engi-
neered at other developmental loci, it is difficult to propose a
synthetic view of the results for several parameters are usually
involved and mixed with one another, such as the presence or
absence of TAD borders and/or CTCF sites, as well as the
presence of enhancers, their relative distribution, or their dis-
placement related to their target genes. Deletions and dupli-
cations at the Ihh locus disrupted the communication with multiple
enhancers, leading to limb malformations (10). Also, rear-
rangements of the TAD containing Shh and its enhancers led to
deleterious effects on gene activation and concurrent phenotypes
(38). Yet, at this specific locus, moderate topological modifica-
tions did not elicit any severe limb defects, suggesting that
enhancer–promoter communication may not rely only on a sus-
tained 3D structure and that sporadic interactions may be suf-
cient (31–33). Expectedly, stronger phenotypes were obtained
by deleting the ZRS enhancer region as this sequence is the only
known limb enhancer in this landscape (39–41). The transitory
effects observed when modifying T-DOM could either reflect a
normal enhancer deployment delayed by changes in the 3D con-
text or, alternatively, a novel organization in enhancer–
promoter interaction due to the known cooperative capacity that
neighboring enhancers can display during embryonic develop-
ment (6, 8, 42). Similarly, the deletion of the TAD border and
inversion of the regulatory domain at the Sox9/Kcnj2 locus only
had mild effects on gene expression (43).

The Resilience of Enhancer–Promoter Interactions. In its initial form,
the inversion of T-DOM introduced a strong ectopic TAD bor-
der between the enhancers and their target genes, in addition to
the reorientation of all CTCF binding sites. In this allele (Fig. 6),
accordingly, the access of Hoxd genes to their cognate limb
enhancers was dramatically reduced. While a severe delay was
scored in transcriptional activation, some interactions surpris-
ingly remained between the gene cluster and the regulatory do-
main, despite the latter being clearly in a distinct TAD. These
contacts could even resume the expression of Hoxd genes in
proximal limb cells, although with a truncated spatial distribu-
tion. This observation is slightly at odds with the view of TAD
borders restricting the access to neighboring enhancers and
delimiting regulatory interactions (4, 22, 44–46). Here, despite
the presence of a strong TAD border and the inversion of CTCF
sites, which clearly led to the formation of a new TAD excluding
the Hoxd cluster, some enhancer–promoter interactions could
still occur at a sufficient level to eventually produce detectable
mRNAs in the expected proximal domain, thereby indicating
that such contacts have intrinsic driving forces and do not en-
tirely depend upon an instructive 3D context.

Colinear Regulation and Phenotypic Effects. During early limb bud
development, Hoxd genes are activated in a time sequence that
follows their respective positions along the gene cluster (24, 47).

The mechanism underlying this temporal colinearity process has
been studied by intensive chromosome engineering whereby the
order and/or presence of genes was modified, as well as their
physical relationships with the adjacent regulatory landscapes
(see references in ref. 48). However, a potential involvement of
the regulatory topology, rather than the target end, remained to
be assessed. In this study, by inverting the entire T-DOM, we
could rule out the possibility that the physical order of various
distant enhancers could play a major role in this mechanism,
other than introducing a transcriptional delay, particularly visible
in late-expressed Hoxd genes.

Finally, it is legitimate to wonder whether such moderate
differences in the timing of gene activation could be detri-
mental to the development of the limb, considering that a
close-to-normal expression pattern was resumed in E12.5 limb
buds, except for the inversion of T-DOM containing the ectopic
boundary where the anterior part of the domain remained absent
even at later stages. This particular question was not addressed
in this paper since the detection of any loss of function pheno-
type would likely be hampered by the cooperative function of
both the HoxA and HoxD clusters in developing limbs. Indeed,
while their combined deletion led to very severe limb truncua-
tions, their deletion in isolation triggered much milder pheno-
types (49). While this functional complementation between these
two gene clusters makes phenotypic analyses very complex in the
mouse (it obliges one to systematically remove the other gene
cluster), it has allowed one to study the underlying regulatory
mechanisms in some detail due to the persistence of a rather
normal structure even after drastic chromosomal interventions.

In the above-mentioned alleles, it is thus difficult to anticipate
whether or not any phenotype would be observable in the ab-
sence of the HoxA cluster. In the case of T-DOM inversion in-
cluding the TAD border, it is, however, clear that the lack of
transcripts at the anterior margin of the proximal expression
domain would lead to an abnormal formation of the intermedi-
ate part of the limbs, as was shown in mice carrying a double
inactivation of Hoxa11 and Hoxd11, which displayed severely ill-
formed forelimbs (50).

Materials and Methods

Mouse Strains. The HoxDdel(CTCFs;CS38;CS40) or del(CTCF;CS38;40) allele was described in ref.
27. The HoxDdel(CTCFs;CS38;CS40) or del(CTCFs), HoxDdel(CTCFs;CS38;40) or inv(CTCF;CS38),
HoxDdel(CTCFs;CS38;40) or inv(T-DOM), and the HoxDdel(CTCFs;CS38;40) or inv(T-DOM)
deb allele were generated through CRISPR/Cas9 editing technology using electroporation of mouse zygotes. The del(CTCFs) allele was derived from the HoxDdel(CTCFs;CS38;40) or del(CTCFs;CS38) allele and was also generated for
this study using a guide RNA (gRNA) designed against the consensus CTCF
binding site located in region CS38, which generated a 26-bp large deletion.
Subsequently, two gRNAs flanking region CS40 were designed to produce a

Rodriguez-Carballo et al. 31328 | www.pnas.org/cgi/doi/10.1073/pnas.2015083117
1,533-bp-large deletion encompassing both CTCF binding sites at region CS40. For the inv(CS38-40) allele, two gRNAs were designed flanking the region CS38-40. Mice were genotyped either for a deletion or for an inversion of the region. Out of 48 specimens, only one mouse had an inversion, which we used as founder of the mutant line. For the inv(T-DOM) allele, two different gRNAs were simultaneously directed at each end of the T-DOM regulatory domain. Out of 43 mice, only one had a full inversion, which was subsequently used to establish the mutant line. The break points for the inversion were located 3,433 bp downstream of the Hoxd7 gene and 2,557 bp upstream of the Hmrgap3 gene, inducing an inversion of 888,111 bp. This line was subsequently used to generate the inv(T-DOM) allele, two gRNAs were designed flanking the boundary region now relocated in inv(T-DOM) close to Hoxd1, deleting a 17,303-bp-large region. The del(3Bd) allele, two different gRNAs were generated after cloning the gRNAs into the pX330-hSpCas9 (Addgene ID 42230) vector and DNA injection into pronuclei. All other alleles were generated by electroporation of one-cell embryos with transcribed RNAs. The gRNAs and genotyping primers are listed in SI Appendix, Table S2. All break points were validated through Sanger sequencing, and this information was used to generate the artificial mutant genomes, which can be found in https://zenodo.org/record/3826913#.X6QqXS9h2L4 and a diagram in SI Appendix, Fig. S8.

4C-Seq. The 4C-seq experiments were carried out as described in ref. 22. Samples were microdissected from E12.5 or E9.5 forelimbs and placed in 10% fetal bovine serum (FBS)/phosphate-buffered saline (PBS) and incubated with collagenase at 37 °C for 40 or 15 min, respectively. Cell suspension was filtered and fixed in 2% formaldehyde (FBS/PBS) for 10 min. For E12.5 experiments, between 10 and 12 pairs of distal or proximal forelimbs were used while between 90 and 150 pairs of forelimbs were dissected for the E9.5 experiments. All E9.5 and the E12.5 del(T(CTFs)) 4C-seqs were conducted in embryos obtained from homozygous crosses while all others were obtained from heterozygous crosses. All 4C-seq primers used in this study are listed in SI Appendix, Table S3. The fastq from 4C-seq were demultiplexed, mapped, and analyzed using a local version of the pipeline that was present in HTSTATION (51) on the wild-type mm10 mouse genome. For E12.5 experiments, between 10 and 12 pairs of distal or proximal forelimbs were used while between 90 and 150 pairs of forelimbs were dissected for the E9.5 experiments. All E9.5 and the E12.5 del(T(CTFs)) 4C-seqs were conducted in embryos obtained from homozygous crosses while all others were obtained from heterozygous crosses. All 4C-seq primers used in this study are listed in SI Appendix, Table S3. The fastq from 4C-seq were demultiplexed, mapped, and analyzed using a local version of the pipeline that was present in HTSTATION (51) on the wild-type mm10 mouse genome. For E12.5 experiments, between 10 and 12 pairs of distal or proximal forelimbs were used while between 90 and 150 pairs of forelimbs were dissected for the E9.5 experiments. All E9.5 and the E12.5 del(T(CTFs)) 4C-seqs were conducted in embryos obtained from homozygous crosses while all others were obtained from heterozygous crosses. All 4C-seq primers used in this study are listed in SI Appendix, Table S3. The fastq from 4C-seq were demultiplexed, mapped, and analyzed using a local version of the pipeline that was present in HTSTATION (51) on the wild-type mm10 mouse genome. For E12.5 experiments, between 10 and 12 pairs of distal or proximal forelimbs were used while between 90 and 150 pairs of forelimbs were dissected for the E9.5 experiments. All E9.5 and the E12.5 del(T(CTFs)) 4C-seqs were conducted in embryos obtained from homozygous crosses while all others were obtained from heterozygous crosses.
Whole-Mount In Situ Hybridization and LacZ Transgenes. WISH was performed as in ref. 59 with the following modifications. E10.5 and E12.5 embryos were bleached in a 3% H2O2/PBS solution. After rehydration, embryos were digested in a proteinase K solution (20 μg/mL, 10 to 12 min, for E12.5 embryos; 10 μg/mL, 5 min, for E10.5 embryos; and 5 μg/mL, 4 to 5 min, for E9.5 embryos). Digestion of E9.5 embryos was arrested by three quick washes in a 2-mg/mL glycine solution while E10.5 and E12.5 proteinase K digestions were stopped by a 10-min incubation in an acetic anhydride/trimethanolamine solution. After mRNA probe hybridization and anti-DIG incubation, E9.5 embryos were washed several times overnight in maleic acid buffer while E10.5 and E12.5 embryos were washed for 1 day.

The ELCR2 and ELCR3 regions were amplified and cloned into the beta-Globin reporter plasmid as in ref. 58 to generate the corresponding LacZ transgenes. These enhancer regions were amplified using the primer sequences: 5′-AACATCATTGTCTTTCTAGAGTGGAGTTTCTCT-3′ and 5′-GCTGGTGTGACACGCCCAATGGAC-3′ for ELCR2; 5′-GGTTTTTTGTTTCTACTTTGTTTTTTCACTGGTTCTACAC-3′ and 5′-GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA...
39. L. A. Lettice, P. Devenney, C. De Angelis, R. E. Hill, The conserved sonic hedgehog limb enhancer consists of discrete functional elements that regulate precise spatial expression. *Cell Rep.* **20**, 1396–1408 (2017).

40. C. Paliou et al., Preformed chromatin topology assists transcriptional robustness of Shh during limb development. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 12390–12399 (2019).

41. I. Williamson et al., Developmentally regulated Shh expression is robust to TAD perturbations. *Development* **146**, dev.179523 (2019).

42. C. C. Bolt, D. Duboule, The regulatory landscapes of developmental genes. *Development* **147**, dev171736 (2020).

43. A. Despang et al., Functional dissection of the Sox9-Kcnj2 locus identifies nonessential and instructive roles of TAD architecture. *Nat. Genet.* **51**, 1263–1271 (2019).

44. J. G. van Bemmel et al., The bipartite TAD organization of the X-inactivation center ensures opposing developmental regulation of Tsix and Xist. *Nat. Genet.* **51**, 1024–1034 (2019).

45. M. Franke et al., Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature* **538**, 265–269 (2016).

46. D. G. Lupiáñez et al., Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**, 1012–1025 (2015).

47. P. Dollé, J.-C. Izpisúa-Belmonte, H. Falkenstein, A. Renucci, D. Duboule, Coordinate expression of the murine Hox-5 complex homeobox-containing genes during limb pattern formation. *Nature* **342**, 767–772 (1989).

48. P. Tschopp, D. Duboule, A genetic approach to the transcriptional regulation of Hox gene clusters. *Annu. Rev. Genet.* **45**, 145–166 (2011).

49. M. Kmita et al., Early developmental arrest of mammalian limbs lacking HoxA/HoxD gene function. *Nature* **435**, 1113–1116 (2005).

50. A. P. Davis, D. P. Witte, H. M. Hsieh-Li, S. S. Potter, M. R. Capecchi, Absence of radius and ulna in mice lacking hoxa-11 and hoxd-11. *Nature* **375**, 791–795 (1995).

51. A. A. David et al., HTStation: A web application and open-access libraries for high-throughput sequencing data analysis. *PLoS One* **9**, e85879 (2014).

52. L. Lopez-Delisle et al., pyGenomeTracks: Reproducible plots for multivariate genomic data sets. *Bioinformatics*, 10.1093/bioinformatics/btaa692 (2020).

53. G. Geeven, H. Teunissen, W. de Laat, E. de Wit, peakC: A flexible, non-parametric peak calling package for 4C and Capture-C data. *Nucleic Acids Res.* **46**, e91 (2018).

54. A. Dobin et al., STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

55. T. Montavon, J. F. Le Garrec, M. Kerszberg, D. Duboule, Modeling Hox gene regulation in digits: Reverse collinearity and the molecular origin of thumbness. *Genes Dev.* **22**, 346–359 (2008).

56. Enis Afgan et al., The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* **46**, W537–W544 (2018).

57. C. Schmidl, A. F. Rendeiro, N. C. Sheffield, C. Bock, ChiPmentation: Fast, robust, low-input ChiP-seq for histones and transcription factors. *Nat. Methods* **12**, 963–965 (2015).

58. E. Rodríguez-Carballo, L. Lopez-Delisle, N. Yakushiji-Kaminatsui, A. Ullate-Agote, D. Duboule, Impact of genome architecture on the functional activation and repression of Hox regulatory landscapes. *BMC Biol.* **17**, 55 (2019).

59. J. M. Woltering et al., Axial patterning in snakes and caecilians: Evidence for an alternative interpretation of the Hox code. *Dev. Biol.* **332**, 82–89 (2009).

60. I. Guerreiro et al., Reorganisation of Hoxd regulatory landscapes during the evolution of a snake-like body plan. *eLife* **5**, 1–23 (2016).