Convergent evolution of complex regulatory landscapes and pleiotropy at Hox loci

Nicolas Lonfat,1* Thomas Montavon,1† Fabrice Darbellay,1 Sandra Gitto,2 Denis Duboule1,2†

Hox genes are required during the morphogenesis of both vertebrate digits and external genitals. We investigated whether transcription in such distinct contexts involves a shared enhancer-containing landscape. We show that the same regulatory topology is used, yet with some tissue-specific enhancer-promoter interactions, suggesting the hijacking of a regulatory backbone from one context to the other. In addition, comparable organizations were observed at both HoxA and HoxD clusters, which separated through genome duplication in an ancestral invertebrate animal. We propose that this convergent regulatory evolution was triggered by the preexistence of some chromatin architecture, thus facilitating the subsequent recruitment of the appropriate transcription factors. Such regulatory topologies may have both favored and constrained the evolution of pleiotropic developmental loci in vertebrates.

T he evolution of vertebrates involved two rounds of genome duplications, which facilitated gene neofunctionalization (1) via novel regulatory sequences (2) and led to high pleiotropy for most known developmental genes, either coding for transcription factors or involved in signaling pathways (3, 4). Hox genes illustrate such functional cooptations, and particular global regulations are associated with specific Hox gene clusters. For example, in addition to organizing the major body axis, both HoxA and HoxD genes are required for the development of digits and external genitals (5–7), two structures sharing genetic and embryological similarities (8–10) (Fig. 1A). Accordingly, the combined inactivation of Hoxd13 and Hoxa13 led to both digits and external genitals agenesis (5–7).

During digit development, Hoxd13 is controlled by multiple enhancers located within an 800-kb flanking gene desert (Fig. 1B). This regulatory archipelago (11) matches a topologically associating domain (TAD), a chromatin structure in which both enhancer-promoter and structural interactions are enriched (12–14). Therefore, such a constitutive chromatin domain may be used as a template to support the necessary regulatory interactions between Hoxd13 and the various digit enhancers. Indeed, although some interactions were detected only when target genes were transcribed, others were established regardless of the transcriptional status, thus forming a preorganized chromatin backbone (11, 15). We assessed whether transcription of Hoxd13 during genitals outgrowth relies upon the same set of digit enhancers, within the same TAD, despite distinct developmental contexts. In transcription profiles from both fetal digits and genital tubercles (GT), Hoxd13 transcripts were most abundant, whereas Hoxd12 to Hoxd8 showed progressively lower expression levels (Fig. 1A). This similarity extended to both histone H3K27 acetylation and trimethylation profiles (Fig. 1A), confirming that the regulation of these genes during GT and digit development is similar (16).

We then analyzed the importance in GT of the digit regulatory landscape (11), using scanning deletions in vivo to compare Hoxd13 transcription in digits and GT cells. Deletion of the entire gene desert reduced Hoxd13 mRNAs in GT by 80% (Fig. 1, B and C), an effect close to that observed in digits (17). Using the Del(rel5-rel10) allele, which removes one-third of the regulatory landscape, including two digit enhancers (17), mRNAs were down by 40% in both digits and GT cells. When the adjacent 300 kb were removed [Del(rel5-SB)], a similar decrease was obtained in both cases. However, the 290-kb deletion, including the distal part of the landscape [Del(SB-Atf2)], had no effect on Hoxd13 in GT cells, unlike in digits

Fig. 1. Transcription of Hoxd genes in digits and genitals. (A) Hoxd13 is transcribed in both developing digits (top) and genitals (bottom). Transcripts (black), H3K27ac (blue), and H3K27me3 (red) profiles show similar patterns between digits and genitals, with Hoxd13 as the strongest expressed gene. Transcribed parts of the cluster are decorated with H3K27ac marks, and silent genes are H3K27me3 positive. (B) Scheme of the gene desert upstream from the HoxD cluster, including the seven digit regulatory islands (blue ovals) (11). The loxP sites used for deletions are shown in red along with the names of the alleles. (C) In vivo deletions within the HoxD regulatory landscape affect Hoxd13 transcription in digits (JJ) and genitals (GT), although with slight differences. Names of deletions refer to (B).
where transcription was down to 50% (Fig. 1C). Therefore, although the digit regulatory landscape also controls Hoxd13 transcription in external genitals, enhancers specific for either structure are organized with distinct topologies. We controlled the potential of this regulatory landscape in GT cells by using an inversion including a Hoxd8lacZ reporter transgene (fig. S1). In mutant embryos, both the digit enhancers and the reporter transgene are physically disconnected from their Hoxd targets. LacZ staining of mutant embryos revealed reporter staining into both GT and digit cells. Therefore, enhancers required for Hoxd transcription in GT and digit cells are located within the same regulatory landscape (fig. S1).

We compared the interaction profiles established by Hoxd13 in digits and GT cells by circular chromosome conformation capture (4C). In GT cells, Hoxd13 contacted mostly the centromeric gene desert, supporting the deletion approach. Interaction peaks extended over approximately 1 Mb and matched the TAD covering this region, as in digits (12, 15) (Fig. 2, A and B). However, although the profiles shared many interactions, they also displayed differences with contacts observed in digits but not in GT and vice versa (Fig. 2C). Hoxd13 interaction peaks were aligned with both evolutionary conserved noncoding sequences and the GT H3K27 acetylation profile (fig. S2). Potential regulatory sequences active either in GT only (GT1 and GT2), in digits only (Island II-1) (13), or in both (Prox) (17) were thus narrowed down and assayed in vivo using a LacZ reporter system (16) (Fig. 3). Accordingly, embryo transgenics for the GT2 sequence displayed strong activity in GT cells exclusively, whereas sequence II-1 triggered expression in digits but not in GT (Fig. 3). The staining elicited by GT1, however, was scored both in GT and in digits, perhaps reflecting the weak interaction peak between Hoxd13 and GT1 in digit cells (Fig. 2C). As a control, the Prox element showed strong interactions with Hoxd13 in both digit and GT cells, and staining was detected at both sites (Fig. 3). Islands I to IV did not trigger any staining in GT, even though all but island I were H3K27ac-positive in this tissue (fig. S2). Therefore, within this complex regulatory landscape, some intermingled enhancers are specific either for the genitals or for the digits, whereas others are used by both structures.

We asked whether the large spacing between these enhancers is necessary for their proper activities—for example, to isolate them from one another and prevent undesired regulatory interferences. We produced a reporter transgene containing 300 base pairs of the GT2-specific enhancer located within digit island II-1. This chimeric DNA fragment elicited a strong additive staining in both digits and GT cells (Fig. 3), showing that each sequence could work autonomously and independently from its immediate neighborhood. Therefore, in this particular case, the TAD may not be critical to properly organize internal enhancers into a particular functional architecture.

Hoxa13 is also involved in the morphogenesis of both digits and external genitals with specificity and function similar to those of Hoxd13 (5–7, 19) (Fig. 4A). Hoxa13 transcription in digits relies on several enhancers located within a regulatory landscape located telomeric to the Hoxa4 cluster and containing some unrelated yet coexpressed genes (20, 21). As for Hoxd13, this regulatory landscape matches a TAD—as defined by Hi-C, chromosome conformation capture carbon copy (5C), and 4C (12, 21, 22)—with a boundary located within the Hoxa4 cluster, as for Hoxd. To see whether in this case again, digit and genital enhancers were embedded into the same TAD, we generated and compared the Hoxa13 4C interaction profiles in developing genitals and digits. Both profiles extended over the same genomic distance, overlapping with the same TAD (Fig. 4, B and C, and fig. S1). Although some contacts involving known sequences were common to both cell types, others were clearly more prominent either in GT or in digit cells. The GT2 sequence strongly interacted with Hoxa13 in GT, whereas much weaker contacts were scored in digits. In contrast, the e66 DNA regulatory sequence...
[from (21)] displayed robust interactions with Hoxa13 in digits, whereas contacts were reduced in GT cells (Fig. 4C). In transgenic mice, GT e2 triggered preferential expression in developing genitals, whereas e16 was active mostly in digits (Fig. 4D) (20, 21). As a control, the GT e5 sequence contacted by Hoxa13 in both contexts gave some staining in both structures (Fig. 4D). Therefore, the regulation of Hoxa13 transcription in both digit and GT cells follows the same general principles as observed at the HoxD locus.

These similarities in size and organization of both enhancer landscapes are unexpected because the duplication of an ancestral Hox cluster occurred in a nonvertebrate animal, which had neither limbs nor external genitals (23, 24) (Fig. S4). Therefore, these related regulatory domains evolved independently, which may explain why only a few sequence homologies are found between HoxA and HoxD enhancers (21). Because digits and genitals were likely of high adaptive value for both locomotion and reproduction in a terrestrial environment, the parallel emergence of these morphological novelties may not be a surprise. However, the convergence between regulatory mechanisms is more difficult to explain. One possibility involves the presence in an ancestral Hox cluster of a constrained functionality or structure, which was used as a starting point to facilitate this convergence. In this view, a TAD might have been present before cluster duplications (fig. S4) and thus favored enhancer recruitment by providing the molecular and structural niches necessary to elicit productive interactions. Because group 13 HOX proteins were key to developing these appendices, any up-regulation of these genes during the evolution of digits and genitals—for example, via a response to appropriate signaling pathways (9, 25)—may have been selected and further reinforced. The presence of a comparable organization in teleost fishes (22) supports the idea that these domains existed before the emergence of digits and genitals. Alternatively, convergent enhancer evolution may have driven the emergence of comparable TADs at both loci. We consider this as less probable because it implies the independent construction of comparable regulatory architectures without any prior constraint.

Our results suggest that TADs may act as genomic niches where tissue-specific factors can hijack global transcriptional readout from one cell type to the other. Although the use of the same regulatory landscape in digits and genitals indeed suggests that it was coopted from one context to the other, it is admittedly difficult to establish whether digits hijacked this regulation from genitals or vice versa. In this case, however, this particular situation may illustrate why many large regulatory landscapes accumulated pleiotropic regulations (e.g., (26)). Accordingly, such well-defined topological domains may not be as clearly visible for genetic loci transcribed more constitutively or with a high cellular specialization. The case reported in this study might represent a permissive system where multiple and distinct inputs could all elicit a comparable transcriptional response. The latter situations, in contrast, would reflect a more instructive role of transcription factors in forming the spatial context necessary for transcribing target genes (27).

REFERENCES AND NOTES

ACKNOWLEDGMENTS
We thank D. Noordermeer, P. Fabre, G. Andrey, A. Necsulea, and all colleagues from the Duboule laboratories for their help and for reagents and discussions. We thank M. Docquier and the Geneva Genomics Platform, S. Verp and S. Offerer for help with transgenic animals, and M. Leleu from the Bioinformatics and Biostatistics Core Facility (BBCF) of the Ecole Polytechnique Fédérale (EPFL). Computations were performed at Vital-IT center for high-performance computing of the Swiss Institute of Bioinformatics, using tools developed by the BBCF. This work was supported by funds from the Federal Institute of Technology (EPFL), the University of Geneva, the Swiss National Research Foundation (no. 310030B_138662), and European Research Council grant SystemsHox.ch (no. 232790). Data sets presented in this study are available with accession no. GSE62509.

SUPPLEMENTARY MATERIALS
www.sciencemag.org/content/346/6212/1006/suppl/DC1
Materials and Methods: Figs. S1 to S4, Table S1, References (28–33)
16 June 2014; accepted 10 October 2014 10.1126/science.1257493