In many animal species with a bilateral symmetry, Hox genes are clustered either at one or at several genomic loci. This organization has a functional relevance, as the transcriptional control applied to each gene depends upon its relative position within the gene cluster. It was previously noted that vertebrate Hox clusters display a much higher level of genomic organization than their invertebrate counterparts. The former are always more compact than the latter, they are generally devoid of repeats and of interspersed genes, and all genes are transcribed by the same DNA strand, suggesting that particular factors constrained these clusters toward a tighter structure during the evolution of the vertebrate lineage. Here, we investigate the importance of uniform transcriptional orientation by engineering several alleles within the HoxD cluster, such as to invert one or several transcription units, with or without a neighboring CTCF site. We observe that the association between the tight structure of mammalian Hox clusters and their regulation makes inversions likely detrimental to the proper implementation of this complex genetic system. We propose that the consolidation of Hox clusters in vertebrates, including transcriptional polarity, evolved in conjunction with the emergence of global gene regulation via the flanking regulatory landscapes, to optimize a coordinated response of selected subsets of target genes in cis.

Hox genes are key players in the organization of the animal body plan. They encode transcription factors, the combination of which can instruct cells at different body levels as to their future morphological contributions. In addition to this ancestral function along the anterior-to-posterior (AP) axis, HOX proteins also participate in the organization of secondary axes or of a variety of organs or structures. In many animals displaying a bilateral symmetry, Hox genes are found clustered in the genome. This particular genomic topology has a functional relevance, as the succession of genes in cis within the cluster corresponds to the order of their expression domains along the various axes. This colinearity phenomenon was initially proposed by Lewis (1) in Drosophila and subsequently extended to vertebrates (2–5). In fact, Drosophila shows a breakpoint into the Hox cluster (6), which was thus split into 2 subclusters, a separation that occurred repetitively at different positions within drosophilids (7). Also, some species display a complete disaggregation of their ancestral cluster into a collection of single gene loci, such as in Urochordata (8).

The existence of an entire Hox gene cluster—that is, when all major paralogy groups are present and linked together in cis—was proposed to be always associated with animals whose segmental development occurs in a rostral-to-caudal time sequence. In such animals, the activation of Hox genes must occur following a precise timing, referred to as the Hox clock, which would be either regulated or coordinated by the clustered organization (see references within ref. 9). The availability of genome sequences for all major groups of animals has not yet proved this conjecture wrong. However, genome analyses have revealed an unexpected property for vertebrate Hox clusters, which differ from their invertebrate counterparts by a higher order in their structural organization. For instance, vertebrate Hox clusters are barely over 100 kb in size (with the exception of axolotl; see Discussion), whereas cephalochordate or echinoderm clusters are around 500 kb large, similar to all characterized single invertebrate clusters (SI Appendix, Fig. 1A). Therefore, the current situation is that not a single animal outside gnathostome (jawed) vertebrates has been reported to carry a complete Hox gene cluster of a size and compaction level close to that of vertebrates (SI Appendix, Fig. 1B).

In addition to their general reduction in size, vertebrate Hox clusters contain little if any repeats; they never include genes unrelated to the Hox family and they transcribe all their genes from the same DNA strand, with a direction of transcription opposite to the time sequence of gene activation. Another important difference between the vertebrate genomes and those of cephalochordates, echinoderms, and most invertebrates is that they contain multiple copies of their Hox clusters as a result of the 2 initial rounds of genome duplication (10–12). This exact correlation between the number of Hox gene clusters on the one hand and number of genomes on the other hand is often referred to as the “Hox clock.”

Significance

Vertebrate Hox genes are clustered. This organization has a functional relevance, as the transcription of each gene in time and space depends upon its relative position within the gene cluster. Hox clusters display a high organization, and all genes are transcribed from the same DNA strand. Here, we investigate the importance of this uniform transcriptional polarity by engineering alleles where one or several transcription units are inverted, with or without a CTCF site. We observe that inversions are likely detrimental to the proper implementation of this genetic system. We propose that the enhanced organization of Hox clusters in vertebrates evolved in conjunction with the emergence of global gene regulation to optimize a coordinated response of selected subsets of target genes.
hand, and their higher level of compaction and organization on the other hand, lead to the proposal that these 2 distinct aspects may have coevolved at the origin of the vertebrate lineage. A possible scenario was proposed whereby much in the same way that neofunctionalization can occur after horizontal or vertical gene duplication, global gene regulation achieved through the flanking regulatory landscapes may have been favored after duplication of the entire gene cluster (SI Appendix, Fig. S1C). In this view, the duplication of genomic Hox loci may have allowed emergence of multiple potent enhancer sequences located outside the clusters and controlling several Hox genes at once. The potential negative effects of such high-order regulations and structures—for example, on the ancestral colinear mechanism at work during axial extension—could have been compensated for by having several clusters implementing this colinear process at the same time (13).

On the other hand, the evolution of de novo global enhancers may have represented an interesting adaptive value, in particular to evolve redundancy, compensatory mechanism, or quantitative regulatory controls. As a result of this accumulation of enhancers in the regulatory landscapes flanking Hox clusters, the genes (or subgroups thereof) would have progressively maximized their responses to these enhancers, leading to a stepwise elimination of interfering repeat sequences, shortening of intergenic distances, and placing all genes in the same transcriptional orientation. The latter point is of importance, for in such a tightly organized group of genes, the inversion of a transcription unit could either interfere with the neighboring genes’ transcription or bring 2 promoters close to one another, leading to similar regulatory controls—a situation that would go against the general principle governing the evolution of this temporal mechanism in vertebrates. The presence of bound CTCF protein, a factor known to be involved in the insulation of chromatin domains (14, 15) between almost every gene of the cluster (16), supports the importance of this iterative genomic topology for a precise processing of gene activation.

In this study, we investigated the importance of transcriptional directionality in physiological conditions by producing and analyzing a set of targeted inversions within the Hox D cluster and looking at the induced effects over the neighboring genes in various developmental contexts. We report the impact of inversion of both the Hoxd11 and Hoxd12 loci, separately, without disturbing the distribution of intervening CTCF sites, as well as the effect of a combined inversion of the 2 loci together, along with the repositioning of an inverted CTCF site. While the former 2 inversions revealed regulatory disturbances, they led to rather minor effects, whose long-term impact on the animals was difficult to evaluate. The larger Hoxd11-Hoxd12 inversion, in contrast, elicited a dramatic up-regulation of the neighboring Hoxd13 gene. By using additional engineered alleles, we show that this up-regulation is likely due to the reorganization of chromatin microdomains, rather than the leakage of transcription on the opposite DNA strand, sent toward Hoxd13 by the inverted Hox genes. We show that such chromatin domains are separated by a critical CTCF site, the deletion of which also leads to a transitory up-regulation of Hoxd13 in the developing metanephric kidneys. Finally, we use a different allele to illustrate the deleterious effect of a stable gain of expression of Hoxd13 in developing metanephros.

Results
Inversion of the Hoxd11 Locus. We used CRISPR-Cas9 to produce a first inversion involving the Hoxd11 transcription unit. The 5′ part of the HoxD cluster was selected due to the rather late timing of activation of these genes, which makes their study easier in terms of developmental stages, amount of material to collect, and phenotypes to observe than their more 3′-located neighbors. This inversion was designed not to interfere with neighboring transcription start or termination sites and did not include any of the CTCF binding sites (Fig. 1A). Therefore, any effects not generated by the mere inversion of Hoxd11 transcription were reduced to a minimum.

A strain referred to as HoxDinv(11) was produced and is mentioned as inv(11) mice or allele throughout this work. Animals carrying this inversion were analyzed by whole-mount in situ hybridization (WISH) at embryonic day 12.5 (E12.5). In control animals, Hoxd11 expression was well established both along the primary body axis, where it labeled the transition between the limb and the sacral regions (Fig. 1B, Top Center, arrowhead), as well as the proximal and distal parts of developing limb buds (blue arrows). At the same stage, Hoxd12 was expressed more posteriorly, whereas Hoxd10 transcripts were found at more anterior positions, following the rule of colinearity (Fig. 1B, Top Left and Top Right, respectively), whereby the more 5′ located a gene is positioned within its Hox cluster, the later this gene will be activated and the more posterior its expression domain will be.

WISH using homozygous inv(11) mutant littersmates showed a clear decrease in the amount of detected Hoxd10 and Hoxd12 messenger RNAs (mRNAs) visible in both the developing trunk axis and limb buds (Fig. 1B, Bottom). Because WISH is poorly quantitative, we carried out RNA sequencing (RNA-seq) analyses by using the developing digits and metanephros, which were selected because they strongly express Hoxd11 under normal conditions and require its function for proper development. Also, these tissues express distinct combinations of other Hox genes, due to the topologies of their regulations. In presumptive digit cells, Hoxd11 is transcribed along with Hoxd10, Hoxd12, and Hoxd13 as well as the Evx2 gene positioned 8.8 kb upstream of Hoxd13. These genes respond to series of enhancers located in the centromeric regulatory landscape (C-DOM), which coincides with a topologically associating domain (TAD) (SI Appendix, Fig. S1C) (17). In contrast, future metanephric cells express Hoxd11 together with Hoxd10, Hoxd9, Hoxd8, and a moderate level of Hoxd12 due to enhancers positioned in the opposite telomeric regulatory landscape (or T-DOM; SI Appendix, Fig. S1C), whereas Hoxd13 mRNAs are not detected. While expression of Hoxd13 and Hoxd12 is necessary for digit development, these genes are detrimental to the development of metanephric kidneys due to a probable dominant-negative effect over other HOX proteins, referred to as posterior prevalence (18).

In digit cells, Hoxd13 is expressed the strongest, followed by Hoxd12 and Hoxd11 (Fig. 2 A and B). While transcription termination seems to occur faithfully for Hoxd13, the RNA-seq dataset obtained from Hoxd11pho(12-15)-control fetuses revealed transcriptional leakages of both Hoxd12 and Hoxd11. These 2 genes exhibit low levels of mRNAs extending toward their neighboring downstream Hox transcription units, Hoxd11 and Hoxd10, respectively (Fig. 2B, arrows). After inversion of Hoxd11 (Fig. 2C, shaded area), Hoxd11 mRNAs, now encoded by the other DNA strand, continued at high level toward the Hoxd12 transcription unit (Fig. 2C, Middle, anti-Hox profile, arrow). This transcription leakage was likely due to the fact that the major termination signal for Hoxd11 was not inverted along with the transcription unit. This abnormally elongated Hoxd11 transcript extended up to the termination site of the Hoxd12 mRNA, where it decreased abruptly, as shown by the superimposition of both DNA strands (Fig. 2C, Bottom profile, black arrowhead). However, some Hoxd11 transcripts continued to leak over the Hoxd12 locus up to the 3′ extremity of Hoxd13 (Fig. 2C, Bottom profile, gray arrowhead). In parallel, transcription of Hoxd12, from the other DNA strand, was greatly reduced (Fig. 2A, arrow), as if the abnormal level and extension of Hoxd11 transcription on the anti-Hox coding strand would interfere with the correct transcription process of Hoxd12 from the normal—noninverted—strand (Fig. 2C, Top and Bottom profiles).
In contrast to digit cells, developing metanephric cells do not express any *Hox*13 mRNAs and only moderate levels of *Hox12*, with *Hox11*, *Hox10*, and *Hox9* being the most expressed genes (Fig. 2D). As in digit cells, however, the inversion of the *Hox11* locus led to a decrease in the level of *Hox12* mRNAs (Fig. 2D, arrow). Again, the *Hox11* transcripts now encoded by the *Hox* opposite DNA strand did not terminate after exon 2, as in the normal case (Fig. 2E, arrow), but extended up to the termination site of the *Hox12* transcripts (Fig. 2F, black arrow and arrowhead), with some weak but clear signal continuing over the *Hox12* transcription unit. Of note, in the inverted configuration, *Hox10* transcripts were scored as in the wild-type chromosome, showing that under normal conditions, the transcription of this gene is not dependent upon the leakage of transcripts coming from *Hox11* and extending up to *Hox10* (Fig. 2E, arrow).

Therefore, in this particular case, the inversion of a *Hox* gene locus induced the down-regulation of transcripts coming from the gene located immediately in 5′, likely as a result of a collision effect of the transcriptional machineries associated with the respective genes. It is unclear as to why the opposite effect was not scored—that is, why the transcription of the inverted *Hox11* gene was not affected by *Hox12* transcripts—in particular in digit cells where *Hox12* is normally expressed at higher levels than *Hox11* and where the former gene seems to also send some transcripts toward the latter (Fig. 2B, arrows). It is nevertheless clear that *Hox11* transcripts have the capacity to negatively interfere with *Hox12* transcription, whereas the opposite was not observed.

**Inversion of the Hoxd12 Locus.** We next inverted the *Hoxd12* transcription unit—the piece of DNA just adjacent to the *Hoxd11* 5′ inversion breakpoint—leaving in place the CTCF site separating these 2 loci from the *Hoxd13* gene (SI Appendix, Fig. S24). As for the inversion of *Hoxd11*, we looked at the effect of this inversion on the transcription profiles of developing metanephros and digits. RNA quantification of a subset of *Hox* genes in E13.5 metanephros revealed opposite variations in the amount of *Hoxd12* and *Hoxd11* mRNA, with an increase of *Hoxd12* transcription and a slight decrease in *Hoxd11* transcription (SI Appendix, Fig. S2B, black arrows). The expression of other *Hox* genes remained unchanged. This increase was equally visible on the RNA-seq profiles obtained from *Hoxd10* (together with *Hoxd12*) and *Hoxd13* mutant metanephros (SI Appendix, Fig. S2 C and D, compare red and blue profiles). The inversion of *Hoxd12* also resulted in a limited transcriptional leakage from the anti-*Hox* DNA strand, extending toward *Hoxd13* (SI Appendix, Fig. S2C, black arrow). However, transcripts terminated at the same position as previously observed for the inv(11) allele (Fig. 2C and F and SI Appendix, Fig. S2C, gray arrowhead) and hence did not overlap with the *Hox13* gene. *Hoxd13* and *Evx2* remained transcriptionally silent.

During digit development, the amount of *Hoxd12* transcripts was reduced in inv(12), in contrast to the situation described in metanephros, whereas other *Hox* genes did not display any substantial variation in RNA content (SI Appendix, Fig. S2E). The RNA-seq profiles obtained from developing transheterozygote mutant metanephros (SI Appendix, Fig. S2F, compare red and blue profiles). The inversion of *Hoxd12* also resulted in a limited transcriptional leakage from the anti-*Hox* DNA strand, extending toward *Hoxd13* (SI Appendix, Fig. S2C, black arrow). However, transcripts terminated at the same position as previously observed for the inv(11) allele (Fig. 2C and F and SI Appendix, Fig. S2C, gray arrowhead) and hence did not overlap with the *Hox13* gene. *Hoxd13* and *Evx2* remained transcriptionally silent.

During digit development, the amount of *Hoxd12* transcripts was reduced in inv(12), in contrast to the situation described in metanephros, whereas other *Hox* genes did not display any substantial variation in RNA content (SI Appendix, Fig. S2E). The RNA-seq profiles obtained from developing transheterozygote mutant metanephros (SI Appendix, Fig. S2F, compare red and blue profiles). The inversion of *Hoxd12* also resulted in a limited transcriptional leakage from the anti-*Hox* DNA strand, extending toward *Hoxd13* (SI Appendix, Fig. S2C, black arrow). However, transcripts terminated at the same position as previously observed for the inv(11) allele (Fig. 2C and F and SI Appendix, Fig. S2C, gray arrowhead) and hence did not overlap with the *Hox13* gene. *Hoxd13* and *Evx2* remained transcriptionally silent.

**Inversion of Both Hoxd11 and Hoxd12.** The independent inversions of either *Hoxd11* or *Hoxd12* did not significantly disturb the general transcriptional activity within the *Hox* cluster. Importantly, both inverted DNA segments were contained between 2 bound CTCF sites, without perturbing either their locations or their orientations. Therefore, we examined an inversion of both genes in *Hoxd13* fetuses did not show any X-gal staining in developing metanephros (Fig. 3B, Left), in agreement with the absence of *Hoxd13* transcription normally observed in this organ located rostral to *Hoxd13* expression boundary. In contrast, a robust staining was scored in *Hoxd10* (together with *Hoxd12*) and *Hoxd13* mutant metanephros, in conjunction with the general anteriorization of *Hoxd13* expression (Fig. 3B, Right).
Fig. 2. (A) Normalized quantification (FPKM) of Hoxd RNAs in E12.5 digits in control HoxDdel(8-13)/+ (gray) and HoxDdel(8-13)/inv11(11) mutant (orange). Hoxd genes display quantitative colinearity, with Hoxd13 expressed the strongest. In the inv(11) allele, Hoxd12 RNAs are reduced (black arrow). (B) Strand-specific RNA signals from heterozygous control HoxDdel(8-13)/+ developing digits. Both Hoxd12 and Hoxd11 transcripts leak toward their 3′ neighbor gene (black arrows). (C) Strand-specific RNA signals from transheterozygous HoxDdel(8-13)/inv(11) developing digits. Inversion of Hoxd11 (shaded area) resulted in transcript leakage toward Hoxd12 (arrow), which stopped downstream of Hoxd12 (black arrowhead). Some transcripts extended toward Hoxd13, whose transcription remained unaffected until its 3′ end (gray arrowhead). (D) Normalized quantification (FPKM) of Hoxd RNAs in E13.5 metanephros. Hoxd12 expression in the inv(11) allele also decreases (black arrow), despite its low level in control metanephros due to a posteriorly restricted domain. (E) Strand-specific RNA signals from control heterozygous HoxDdel(8-13)/+ E13.5 metanephros. Hoxd8 to Hoxd11 are highly transcribed whereas Hoxd12 and Evx2 are silent. Some Hoxd11 transcripts leak over Hoxd10 (black arrow). (F) Strand-specific RNA signals from transheterozygous HoxDdel(8-13)/inv(11) mutant metanephros. Leakage of Hoxd11 transcripts on the anti-Hox DNA strand was maintained (black arrow). A fraction of these transcripts extended up to Hoxd13 3′ UTR termination site (gray arrowhead) and thus covered Hoxd12. Hoxd13 remained silent in this mutant allele. The inversion of Hoxd11 does not reduce the amount of Hoxd10 transcripts (green arrow). (B, C, E, and F) Signals were normalized by the number of million uniquely mapped reads. The mapped RNA signals are shown either in blue (Hox DNA strand) or in red (anti-Hox DNA strand). The inverted Hoxd11 locus is represented by a shaded area. The vertical gray lines around Hoxd11 in the control scheme (last track of B and E) indicate the breakpoints of the inverted allele.
RNA-seq profiles obtained from homozygous HoxDinv(11-12)d13lacZ E13.5 metanephros confirmed the quasi-absence of Hoxd13lacZ transcripts (Fig. 3C, black arrow), with the majority of mRNAs produced by the central Hoxd11 to Hoxd8 genes, all from the same Hox DNA strand, whereas Hoxd12 showed a low level of activity. An important leakage of Hoxd11 RNA extending toward Hoxd10 was detected (Fig. 3C, blue arrow). After inversion, several changes were observed. First, the transcription of Hoxd12, now at the relative genomic position of former Hoxd11, was increased. Second, Hoxd11 transcription was maintained at a rather high level, despite its relative position matching that of former Hoxd12, and its transcripts now leaked toward the Hoxd13lacZ gene, although on the opposite anti-Hox DNA strand (Fig. 3D, red arrow). Concomitantly, Hoxd13lacZ was significantly up-regulated in metanephros from inverted mutant fetuses, even though its relative position in the gene cluster had not changed (Fig. 3D, black arrow). These differences in RNA amounts were quantified by computing the fragments per kilobase per million reads (FPKM) values, confirming the gain of signal for both the Hoxd13lacZ and Hoxd12 genes, located on opposite DNA strands (SI Appendix, Fig. S4, black arrows).

This ectopic activation of Hoxd13 was accompanied by modifications in both the coverage in some histone marks and the micro–3-dimensional structure of the locus. In the control locus, PRC2-dependent H3K27me3 marks (Fig. 4A, orange profile), which label silent genes, were particularly enriched over both Hoxd13 and Evx2, the 2 genes completely inactive during metanephros development (Fig. 4A, bracketed orange arrows). Other Hoxd genes were also labeled by this mark, although to a lesser extent, due to the presence of mixed cellular populations in this developing organ, including a fair proportion of Hoxd-negative cells. The transition between high and low levels of H3K27me3 marks coincided with the presence of both a bound CTCF site and a RAD21 peak (Fig. 4A, large blue arrowhead). This CTCF site labels the TAD boundary at the Hoxd cluster (Fig. 4A, vertical dashed line and ref. 20) and is positioned exactly where clusters of CTCF sites change their orientations (Fig. 4A, blue and red arrowheads). Therefore, it seemed that domains of high versus low H3K27me3 coverage were separated by the first CTCF site with an orientation (blue) opposed to all CTCF sites but one found in the center of the cluster (red). This transition in histone marks was reinforced by the analysis of H3K4me3 modifications (Fig. 4A, green profile), which displayed a distribution complementary to those of H3K27me3 marks, with a robust coverage over Hoxd11 to Hoxd10, a weaker signal over Hoxd12, and only traces over Hoxd13lacZ.

In the inverted allele, the H3K27me3 and H3K4me3 profiles remained complementary (Fig. 4B, track 2, orange and green, respectively). However, the boundary was clearly displaced toward the centromeric end of the gene cluster. The amount of H3K27me3 over Hoxd13lacZ was severely reduced compared with the coverage over Evx2 (Fig. 4B, bracketed orange arrows), while at the same time, H3K4me3 marks robustly increased over the Hoxd13lacZ gene. The boundary between these 2 complementary epigenetic profiles was now positioned between Hoxd13lacZ and Evx2 (Fig. 4B, dashed line) and precisely matched the presence of a pair of CTCF sites oriented in the direction of the centromeric TAD (circled in Fig. 4B). The CTCF site that labeled this transition in the control allele (Fig. 4A, large blue arrowhead) no longer marked the transition in the inverted allele (Fig. 4B, large red arrowhead), even though this site was still occupied by CTCF and matched a peak of RAD21 in the mutant allele (Fig. 4B, arrow in track 3).

This shift in the position of the boundary between active and inactive domains of the Hoxd cluster in the inverted allele was challenged by a circularized chromosome conformation capture...
sequencing (4C-seq) approach, whereby the contacts established by the Hoxd13lacZ gene (Fig. 4 A and B, track 4) were assessed before and after the inversion. In the control allele, contacts were enriched toward the centromeric side (i.e., in agreement with the position of the boundary) (Fig. 4 A, track 4). In the inverted allele however, the bulk of contacts was now detected toward the telomeric side (i.e., following the change in the position of the boundary) (Fig. 4 B, track 4). Therefore, the inversion importantly modified the tropism in contacts of the Hoxd13lacZ gene: from contacts mainly established with a repressive chromatin structure before the inversion to a clear enrichment of contacts with a transcriptionally permisive chromatin domain after the inversion. In the latter case, interestingly, the gain of contacts appeared somehow restricted to the DNA interval between the new boundary and Hoxd9, as if a boundary effect was also observed at this position (Fig. 4 B, open arrowheads).

This effect was confirmed by using a viewpoint corresponding to Hoxd9, on the same material. In the control HoxDinv(11-12)d13lacZ allele, Hoxd9 established contacts with all genes up to the CTCF-labeled boundary, where interactions abruptly stopped (Fig. 4 A, track 4, Lower). In the mutant HoxDInv(11-12)d13lacZ allele, however, contacts extended up to the new boundary (Fig. 4 B, track 4, Lower) corresponding to the CTCF/RAD21 doublet located between Hoxd13lacZ and Evx2 (circled in Fig. 4 A). The new boundary effect triggered by the inverted CTCF site and observed with the Hoxd13lacZ viewpoint (Fig. 4 B, open arrowheads) was nevertheless also detected when using Hoxd9 as bait. Therefore, after inversion, while Hoxd9 interacted with Hoxd13lacZ, these contacts appeared to result from the interactions between 2 microsubdomains rather than from a single domain (Fig. 4 B, track 4, Lower).

**Change in Chromatin Topology or Promoter Cleaning?** These results suggested that local changes in chromatin topology, due to the modified position and orientation of a CTCF site, were responsible for the up-regulation of the Hoxd13lacZ gene in developing metanephros. Alternatively, the transcript emanating

---

**Fig. 4.** Chromatin marks, CTCF/RAD21, and 4C-seq profiles generated from either control homozygous HoxDd13lacZ (A) or homozygous HoxDinv(11-12)d13lacZ mutant (B) E13.5 metanephros. From top to bottom: scheme of the alleles; track 1, strand-specific RNA (adapted from Fig. 3 C and D); track 2, H3K27me3 (orange) and H3K4me3 (green) profiles; track 3, CTCF (black) and RAD21 (turquoise blue) profiles; and track 3, 4C-seq profiles using either a lacZ (Upper) or a Hoxd9 (Lower) viewpoint (red pins). The H3K4me3 and H3K27me3 profiles reveal 2 domains within the HoxD cluster, with silent genes in 5’ and active genes in 3’.

(A) In control metanephros, both Hoxd13lacZ and Evx2 show high levels of H3K27me3 marks (bracketed orange arrows) and little if any H3K4me3 marks, with a boundary matching a CTCF/RAD21 peak (black arrow and vertical dashed dark line), which display an orientation facing the C-DOM (large blue arrowhead). The 4C-seq profiles after inversion were in agreement with the position of this new boundary. The lacZ bait had more interaction on the telomeric side than on the centromeric side, and the Hoxd9 bait now extended its contacts up to include Hoxd13lacZ to stop at the new boundary. The inverted CTCF site induced a slight and local boundary effect (open arrowheads). The inverted DNA segment is shown with a shaded area. Green box indicates lacZ sequences.

(B) Comparable tracks from homozygous HoxDinv(11-12)d13lacZ E13.5 mutant metanephros. The inversion leads to a gain of Hoxd13lacZ expression along with a diminution of H3K27me3 over Hoxd13 (bracketed orange arrows) and a gain of H3K4me3, thus repositioning the boundary between Evx2 and Hoxd13 (vertical dashed line). This position coincided with a pair of CTCF sites (circled), which then labeled the new centromeric side of the boundary, while the initial CTCF site (large blue arrowhead in A) was inverted and displaced (large red arrowhead). The latter site still binds CTCF and RAD21 (black arrow). The 4C-seq profiles after inversion were in agreement with the position of this new boundary. The lacZ bait had more interaction on the telomeric side than on the centromeric side, while the Hoxd9 bait interacted with all genes up to the CTCF-labeled boundary, where interactions abruptly stopped (Fig. 4 A, track 4, Lower). In the mutant HoxDInv(11-12)d13lacZ allele, however, contacts extended up to the new boundary (Fig. 4 B, track 4, Lower) corresponding to the CTCF/RAD21 doublet located between Hoxd13lacZ and Evx2 (circled in Fig. 4 B). The new boundary effect triggered by the inverted CTCF site and observed with the Hoxd13lacZ viewpoint (Fig. 4 B, open arrowheads) was nevertheless also detected when using Hoxd9 as bait. Therefore, after inversion, while Hoxd9 interacted with Hoxd13lacZ, these contacts appeared to result from the interactions between 2 microsubdomains rather than from a single domain (Fig. 4 B, track 4, Lower).
from inverted Hoxd11 and overlapping with the Hoxd13lacZ gene on the opposite anti-Hox DNA strand could elicit a transcriptional response from the latter unit. Such “promoter cleaning” would allow subsequent control by the appropriate upstream factors, which would normally not access this promoter due to the coverage by H3K27me3 marks. To try to discriminate between these 2 explanations, we further deleted the Hox11 transcription units from the inverted allele such as to abrogate the RNA leakage over Hoxd13lacZ. In this HoxDinv(11-12)d13lacZ allele, only Hoxd11 was left inverted and positioned at the same distance from Hoxd13lacZ as inverted Hoxd11 in the HoxDinv(11-12)d13lacZ allele. This secondary mutation did not affect the distribution of CTCF content in any way (SI Appendix, Fig. S5A).

WISH using a lacZ probe revealed that the deletion of Hoxd11 from the inv(11-12)d13lacZ allele failed to suppress the anteriorization of Hoxd13lacZ along the AP axis of HoxDinv(11-12)d13lacZ fetuses (SI Appendix, Fig. S5B, black arrow). This indicated that the anteriorization was likely not due to Hox11 transcripts running through the Hoxd13lacZ promoter, thus permitting a lacZ pattern similar to that of inverted Hox11 (SI Appendix, Fig. S5 B, Middle and Right). Also, HoxDinv(11-12)d13lacZ mutant fetuses transcribed lacZ in their developing eucary (SI Appendix, Fig. S5B, blue arrow) in a manner similar to their aged-matched HoxDinv(11-12)d13lacZ transcripts observed in the Hoxd11 allele (see also refs. 19 and 21), suggesting again that this ectopic expression was not caused by a Hox11 transcript leakage. Likewise, X-gal staining was scored in E13.5 HoxDinv(11-12)d13lacZ metanephros as similar to or even stronger than the HoxDinv(11-12)d13lacZ situation (SI Appendix, Fig. S5C).

Strand-specific RNA-seq carried out using E13.5 homozygous HoxDinv(11-12)d13lacZ metanephros revealed the expected transcriptional activity of Hoxd11 on the opposite strand, as well as the strong gain of expression of the Hoxd13lacZ reporter gene (SI Appendix, Fig. SSD). A faint Hoxd11 transcript was detected extending up to the Hoxd13promoter, yet this RNA was in much lower amount than the leaking Hoxd11 RNA observed in the HoxDinv(11-12)d13lacZ allele (SI Appendix, Fig. SSD; compare with Fig. 3D, red arrow). We compared this decrease in amount of leaking RNAs observed between the HoxDinv(11-12)d13lacZ and the HoxDinv(11-12)d13lacZ alleles, with a quantification of the genes’ transcripts obtained by computing the FPKM described above. The steady-state levels of the various mRNA were globally comparable between the 2 alleles, yet with a clear increase in Hoxd11 mRNAs in the HoxDinv(11-12)d13lacZ allele (which essentially also lacked any Hoxd11 transcripts; SI Appendix, Fig. SSE). This latter observation suggested a negative correlation between the amount of Hoxd13lacZ mRNA produced and the importance of transcript leaking on the opposite DNA strand.

**Mutation of the CTCF Site.** These results favored a change in local topology over a mere transcriptional leakage as a cause for the up-regulation of the Hoxd13lacZ reporter gene. Since the CTCF site present in the inverted DNA seemed to play a particular function in the spatial organization of these chromatin domains, we mutated this site in a wild-type background. This HoxDinv(11-12)d13lacZ allele was generated by CRISPR-Cas9–induced mutagenesis using a guide RNA targeting the core of the CTCF binding motif (SI Appendix, Fig. S6A, pink sequence) positioned between Hoxd12 and Hoxd13. The selected deletion was 21 bp in size, including 13 bp of the core CTCF motif. CTCF Chipmentation, a modification of Chip-seq using the transposase Tn5 to integrate sequencing adapters, performed in transheterozygote HoxDdel(CTCF:d12d13)d13lacZ E13.5 forebrain confirmed the absence of CTCF binding (SI Appendix, Fig. S7A, black arrow).

Dissected metanephros of E13.5 HoxDdel(CTCF:d12d13)d13lacZ and control HoxDinv(8-13) fetuses were analyzed for their content in Hoxd mRNAs through normalized quantification (FKPM) computed from the RNA-seq datasets. In the absence of the CTCF site, a conspicuous gain of expression was scored for both Hoxd12 and Hoxd13—that is, those 2 Hox genes normally located on the other side of the mutated CTCF site—as if the boundary effect had entirely disappeared (SI Appendix, Fig. S6B). This effect was further controlled by in situ hybridization using Hoxd13 as a probe in E13.5 embryos, which revealed a clear gain of expression of Hoxd13 in the developing metanephros of homozygous HoxDdel(CTCF:d12d13) mutant embryos (SI Appendix, Fig. S7B). This gain-of-function effect observed after the mutation of a single CTCF site was rapidly compensated for, however, and ectopic Hoxd13 was already much reduced at E18.5 and no longer scored in adult metanephros of mutant specimens, whereas Hoxd8 and Hoxd9 transcripts were still scored at high levels in both control and mutant adult metanephros (SI Appendix, Fig. S8). This result indicated that while the CTCF mutation had frangible the boundary, the remaining CTCF sites as well as the transcription of other Hox genes might still exert a strong insulation effect, preventing HOXD13 protein to leak into the developing metanephros. Therefore, to address the potential effect of gained HOXD13 in an inversion similar to HoxDinv(11-12)d13lacZ, we used another allele where the gain of Hoxd13 in metanephros was slightly more robust and stable over time such that it was still scored in adult metanephros.

**Detrimental Effect of HOXD13 on Metanephric Kidney Development.** To address this point, we used the HoxDinv(Nsi-Ig6) allele, where an inversion was engineered directly upstream of Hoxd13 up to the Ig6 gene, 3 Mb away (22) (Fig. 5 A, Top). As a consequence, Hoxd13 loses some of its strong contact points in the centromeric neighborhood and reallocate contacts toward its telomeric side (17), where metanephros enhancers are located (23). Using RNA-seq, we looked at whether this partial change in interaction tropism may elicit a gain of function during metanephros development, and a substantial gain was observed for Hoxd13 transcripts, while at the same time, the levels of Hoxd9, Hoxd10, and Hoxd11 transcripts seemed to decrease, likely due to competition between promoters (SI Appendix, Fig. S9).

Mice carrying one copy of this allele were found to suffer from polydipsia; that is, they were consuming more than twice the amount of water than their control littersmates (Fig. 5B, blue arrow). A histological analysis of heterozygous mutant metanephros revealed serious malformations, in particular at the level of the medulla, suggesting that the polydipsia was indeed due to problems in metanephric kidney development (Fig. 5 C, Center). The (Nsi-Ig6) inversion that caused this phenotype is large and has a breakpoint in a gene-rich region around the Ig6 locus and, hence, the contribution of the gained HOXD13 protein in this kidney alteration remained to be clearly established.

Accordingly, we used a CRISPR-Cas9 approach to inactivate the function of the HOXD13 protein on top of the inversion. We induced a small deletion into the homeobox region of HOXD13 (the HoxDinv(Nsi-Ig6)ydel(13)d13 allele; SI Appendix, Fig. S10) and verified that the transcript profile displays similar features, in particular with respect to the gain of Hoxd13 transcripts in developing metanephros (SI Appendix, Fig. S1). Mice carrying this double-mutant rescue allele displayed a normal level of water consumption (Fig. 5B, black arrow), as well as metanephros with a normal morphology (Fig. 5 C, Right), indicating that an ecotopically expressed HOXD13 protein unable to bind DNA could revert the gain-of-function phenotype induced by the inversion. This experiment formally demonstrated that an anterior expansion of Hoxd13 transcripts, in physiological amounts, was detrimental to the development of structures located anterior to its normal expression level, thus providing a selective pressure for the evolution and maintenance of a strong boundary effect in the HoxD cluster.
genes interspersed, the unusual scarcity of repeated elements in these gene clusters, and their reduced sizes compared with invertebrate Hox clusters. In this study, we challenged the functional relevance of another hallmark of these loci; that is, the fact that all HOX proteins are encoded by the same DNA strand. Of note, the only known exception thus far is the aoxotol, where the HoxA cluster spans more than 1 Mb, mostly due to the high density of repeated elements (24). Interestingly, this neotenic animal develops at a speed globally 3 to 4 times slower than vertebrates, including other amphibians (25, 26).

**Inversions with Various Effects.** The various inversions analyzed had different impacts on the regulation of their neighbor Hox genes. While the inversion of Hoxd12 had little (if any) effect on either Hoxd13 or Hoxd11 regulation, the inversion of Hoxd11 had a robust negative outcome upon Hoxd12 transcription. In contrast, the inversion of both Hoxd11 and Hoxd12 in cis elicited a substantial gain of expression of Hoxd13 in various developing tissues. While the functional impact of a down-regulation of Hoxd12 has not yet been uncovered (27, 28), a gain of expression of Hoxd13 in places where it is normally not expressed leads to severe conditions, due to the dominant-negative effect of the HOXD13 protein over other HOX proteins. In limbs and metameres, for instance, ectopic expression of Hoxd13 was shown to correlate with severe anomalies (29, 30), so in metameres, phenocopied the combined loss of function of Hoxa11 and Hoxd11 (31). Therefore, the Hoxd11–Hoxd12 inversion would very likely be detrimental to the mouse. In the current study, however, the inactivation of Hoxd13 by a lacZ prevents these phenotypes from developing and thus allowed for a detailed analysis to be carried out.

The transcriptional effects observed in these inversion alleles can be interpreted within 2 distinct, yet not exclusive, explanatory frameworks. The first involves the transcription dynamics of these genes; that is, the necessity due to the high density of transcription units to have all elongations occurring with the same polarity. In the wild-type situation, some transcription units do not stop before reaching the 5′ untranslated region (UTR) of the gene located in 3′. For example, some Hoxd11 transcripts extend over Hox10 and, likewise, Hoxd4 and Hoxd3 exons are intermingled, thus giving the possibilities for hybrid transcripts to be generated. Upon inversion, these transcription units may either collide with transcripts coming from the other DNA strand or trigger transcription of a silent gene through a promoter-cleaning effect.

The second explanatory framework involves modifications in the topology of the locus, which may directly impact upon the regulations of the various genes. For example, an inversion may either bring a promoter closer to that of the 3′-located neighbor gene, thus inducing promoter sharing and hence the misregulation of at least one of the 2 units, or bring it closer to long-range-acting enhancers. Alternatively, the inversion can reposition structural elements such as CpG islands or CTCF sites, which might in turn modify the general architecture of the locus, inducing novel enhancer–promoter contacts.

**Transcriptional Interference or Topological Modification?** The decrease in the steady-state level of Hoxd12 transcripts upon Hoxd11 inversion may reflect a collision effect, with transcripts elongating from Hoxd11 and reaching the Hoxd12 3′ UTR, thus interfering with the correct transcription of this latter gene and its mRNA’s stability. In Saccharomyces cerevisiae, the main source of transcriptional interference is caused by the frontal collision between 2 transcriptional machineries moving on opposed DNA strands (32–34). These collisions result in mutual arrest of the 2 complexes, leading to their dissociation from the DNA through polyubiquitylation of the RNA polymerase II and the degradation of the aborted RNA transcripts (35–38). However, it is not known whether factors such as the relative...
elongation rates or some characteristics of the chromatin can selectively impact one RNA (32, 39). In the case of the Hoxd11–
Hoxd12 inversion, however, transcripts elongating from Hoxd11
did not collide with those of Hoxd13lacZ since the latter gene
was silent and covered by H3K27me3 marks. This suggested that
such Hoxd11 transcripts were able to displace these repressive
chromatin marks from the Hoxd13 promoter, thus allowing it to
be transcribed in the same domains where Hoxd11 was active.

Recent studies have highlighted the capacity of different
PRC2 components to bind nascent RNA transcripts, resulting (in
some cases) in the eviction of PRC2 out of the chromatin and to
the modulation of its EZH2-dependent methyl-transferase
activity (40–43). To demonstrate this possibility, we deleted the
Hoxd11 gene from this inverted allele so as to stop the mRNA
leakage over Hoxd13. This was successfully achieved, as the now
repositioned inverted Hoxd12 gene sent only a few transcripts
over the Hoxd13lacZ locus, in a much lower amount than inverted
Hoxd11 before its deletion. However, the up-regulation of
Hoxd13lacZ in metanephros was at least as strong as in the
presence of inverted Hoxd11, if not stronger, suggesting that such
a promoter-cleaning mechanism was likely not causal.

Modifications in Local Chromatin Topology. Instead, the Hoxd11–
Hoxd12 inversion also inverted a CTCF site, which seems to be
of particular importance because it is the first of a series of
CTCF sites located at the 5′ extremity of the gene cluster that all
display the same orientation, opposite to those sites found at
more central positions within the HoxD locus. This very region
where the polarity of CTCF sites is inverted coincides with the
strongest boundary effect that separates the 2 TADs flanking the
gene cluster, likely due to large loops established in either di-
rection (20). As a consequence of the Hoxd11–Hoxd12 inversion,
this site was relocated toward a more 3′ position, with an op-
posite orientation. Therefore, rather than separating Hoxd13
from the rest of the gene cluster, the new distribution of CTCF
sites predicted a shift of the boundary upstream Hoxd13 where 2
CTCF sites still have the requested orientations to induce a
boundary effect.

Several observations support this “topology-based” interpre-
tation. First, the Hoxd13lacZ reporter gene, which before the
inversion interacted with a H3K27me3-labeled negative domain
on the centromeric side of the TAD boundary (including Evx2),
contacted the telomeric side of the boundary after the inversion.
Also, in the inverted configuration, Evx2 remained robustly
covered by H3K27me3 marks, suggesting that the new TAD
boundary, at least in metanephros, had been moved to between
Evx2 and Hoxd13lacZ. Finally, the inverted CTCF site continued
to impose a boundary effect (as the CTCF site starting the series
with the same orientation), yet this effect was local and did not
prevent Hoxd13lacZ from contacting the telomeric side of the
HoxD cluster where potential metanephros enhancers were assumed to be.

Deletion of the CTCF Site. Because of the apparent importance of
this particular CTCF binding site in preventing Hoxd13 from
responding to metanephros enhancers, we deleted it from a wild-
type genetic background. This deletion leads to an up-regulation
of Hoxd13 in developing metanephros, supporting the proposal
that the series of CTCF sites with a telomeric orientation is
necessary to block potentially deleterious contacts between
Hoxd13 and “anterior” enhancers. In fact, this specific CTCF site
maps exactly at the position where an elusive “polar silencer”
had been previously positioned, based on a genetic-only ap-
proach (19). The gain of expression was nevertheless transitory,
likely due to the presence of other CTCF sites that make this
boundary very resilient (20), and Hoxd13 transcripts were rapidly
down-regulated, which might account for the apparent absence of
renal phenotype in these animals.

However, a slightly more pronounced and stable presence of
Hoxd13 mRNAs in the developing metanephros, as induced by the
desequestering of the Hoxd13 locus following a large cen-
tromeric inversion, leads to severe alterations in metanephros
morphology already in heterozygous specimens. Although such a
correlation between the ectopic presence of Hoxd13 transcripts
and metanephros malformation (or even agenesis) was previously ob-
served, the causative involvement of these transcripts in triggering
the abnormal phenotype was never shown. Here, we demonstrate
that a mutation in the DNA-binding region of the HOXD13 protein
in cis with the (Nis-Iraga) inversion is sufficient to entirely rescue
the phenotype. This result indicates that HOXD13 is indeed
responsible for the observed alterations and that this dominant-
negative effect is likely mediated by DNA binding.

The Logic of Global Hox Gene Regulation. These latter results
demonstrate the critical importance of insulating this gene from
telomeric regulatory influences and thus help us understand the
general logic of the global regulation occurring at Hox loci.
Hox13 genes are positioned at the last position within vertebrate
Hox clusters and as such, they are the last to be activated at most
caudal positions. Due to the dominant-negative effect of their
proteins over other Hox functions (18), their activation partici-
pates in the termination of various axial systems such as the
limbs, the intestinal tract, and the major body axis, as seen by
various gain- and loss-of-function approaches (44–46). There-
while, it is essential that these genes remain associated with
their respective Hox clusters such that they are activated in co-
ordinated spatial and temporal manners, they must be prevented
to be transcribed too early and too anteriorly, which would have
dramatic effects for the developing body. We believe that this is
the reason behind the evolution of such strong and resilient
boundaries within Hox clusters, as illustrated by the unusually
high concentration of CTCF sites. A similar situation was re-
cently described for the agnathan Hox clusters, which seems to
have been through independent genome amplifications (47). Of
note, these clusters also implement temporal colinearity (48).

Along with the emergence of global enhancers located on ei-
ther side on the gene cluster that progressively form these
complex regulatory landscapes, these CTCF sites were used to
allocate various subgroups of Hox genes to particular telomeric
(anterior, early) enhancers (49), yet always by respecting the
insulation of Hox13 genes. This was shown for proximal limb
enhancers (17), for the intestinal cecum (21), and for the de-
veloping metanephros in this work. In this view, the bimodal
gene regulatory system [proposed by Andrey et al. (17) and
based on the presence of 2 TADs flanking the HoxD cluster]
illustrates the general logic of Hox gene regulation: on the one
hand, a positive “morphogenetic potential” exists for the primary
body axis (the ancestral structure where Hox genes’ functions
were initially deployed) to organize the body plan by regulating
the rostral side of the Hox gene clusters; while on the other hand,
a negative inhibitory function exists for those regulations located
in the other side and acting upon the termination of this genetic
system by activating Hox13 genes. Such a bimodal regulatory
strategy subsequently constrained and guided the evolution of
regulatory innovations on either side of the cluster, along with
the emergence of vertebrate morphological novelties.

Materials and Methods
A detailed description of the mouse stocks and of chromatin immunopre-
cipitation sequencing (ChIP-seq), ChIPmentation, 4C-seq, qRT-PCR, RNA-seq,
WISH, X-gal staining, histology, water consumption, and the genomic data
can be found in SI Appendix, Material and Methods.

Animal Experimentation. All experiments were performed in agreement with
the Swiss law on animal protection, under license no. GE 81/14 (to D.D.). The
use of animals in these experiments was approved by the “Commission Consultative de l’Expérimentation Animale du Canton de Genève.”

Accession Nos. RNA-seq, 4C-seq, ChIP-seq, and ChIPmentation datasets are available from the National Center for Biotechnology Information Gene Expression Omnibus database under accession no. GSE127870.

Acknowledgments. We thank Dr. A. Necsulea for help with bioinformatic analyses; S. Gitto and T.-H. Nguyen Huynh for technical help; the

Genomics Platform of Geneva University; and the Gene Expression Core Facility, the Transgenic Core Facility at the Center of Phenogenomics, and the Histology Core Facility at École Polytechnique Fédérale de Lausanne (EPFL). We thank Dr. L. Beccari for sharing material, and past and present members of the D.D. laboratories for discussions, sharing of reagents, and protocols. This work was supported by funds from the EPFL, the University of Geneva, the Swiss National Science Foundation (No. 310030B_138662), and the European Research Council grants SystemHox (No. 232790) and RegulHox (No. 588029) and the Claraz Foundation (to D.D.).