nature portfolio

# Peer Review File

# COOPTION OF AN ANCESTRAL CLOACAL REGULATORY LANDSCAPEDURING DIGIT EVOLUTION

Corresponding Author: Professor Denis Duboule

Any redactions in this file are there to maintain patient confidentiality, the confidentiality of unpublished data, or to remove third-party material.

**This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.**

Version 0:

Reviewer comments:

Referee #2

(Remarks to the Author)
In this manuscript by Hintermann and Bolt et al, the authors establish that there is synteny and an ancestral TAD organization to the large 5' and 3' regions flanking the Hoxd cluster in zebrafish that was previously well-characterized in tetrapods, largely mice. They further relate this to the observation that HoxD cluster expression is similar in zebrafish fins and mouse limbs with hoxd9a and hoxd10a expressed in the anterior and hoxd11a-hoxd13a restricted to the posterior fin, reflecting a common proximal/anterior and distal/posterior axis.

They ask, is this pattern of expression in zebrafish fins controlled by 5' and 3' flanking regions in the same way as in mouse limbs, despite the fact the distal fin will not form digits? The authors first find that the 3DOM regulatory region is required for hoxd4a and hoxd10a but not hoxd13a expression in zebrafish fins, as in mouse limbs. However, the 5DOM is not required for hoxd13a expression in the fin as it is in mouse limbs. This leads the authors to ask what other function the 5DOM might play in Hox regulation in both species. They convincingly show that the 5DOM contains regulatory sequences that are required for hoxd13 expression in the zebrafish pseudocloacal region and in the mouse urogenital sinus, two homologous structures. The genetic analyses of this paper are strong. They then conclude, as stated in the title, that the regulatory landscape controlling Hoxd13 expression in the digits and genitals came about by "co-option" of a regulatory landscape that ancestrally functioned in the cloaca. I do not see support for this conclusion that is the major crux of the manuscript.

1)The authors see no effect of 5DOM deletion on hoxd13a expression in zebrafish fins, but they note that "previous transgenic results using components of the fish 5DOM sequences in transgenic mice showed sufficiency to drive expression (line159-160)". This references two papers where the Shubin laboratory demonstrated sufficiency of conserved fish enhancers identified in the 5' regulatory region of the HoxD cluster to drive reporter expression in mouse limbs. Notably, the sequences drove much higher and more distal (tetrapod-like) expression when originating from the gar genome than when originating from the zebrafish genome.

2)In the discussion, the authors acknowledge this "may reflect a secondary loss of several distal enhancers associated with teleost whole genome duplication… (which may be solved) with a comparable deletion and epigenetic characterization of 5DOM in more basal fish species such as gar… (lines 288-290)" Yes, this is entirely possible and could undercut the conclusion that the fin/limb regulatory control was co-opted from an ancestral pseudocloacal function.

3)Also in the discussion, they state that "most of the regulatory control (in zebrafish) likely resides within the gene cluster itself", but they do not provide the same sort of transgenic evidence for the sufficiency of reporter expression from within the zebrafish gene cluster, similar to the HoxD cluster BAC reporter that they tested in mice.

4)In the summary Figure 6, they seem to interpret the fact that 5DOM deletion in zebrafish causes a slight and transient loss of hoxd13a to suggest that the "5DOM lightly contributes to hoxda gene regulation in…paired fin buds". I don't know what 'lightly' means in a cis-regulatory or evolutionary context.

The authors have therefore not convincingly shown that limb/genital regulation of hoxd13a from within the 5DOM is a

derived state that wasn't also functioning to control fin expression in more basal fishes.

Second, an interpretation of cis-regulatory co-option requires evidence that the same enhancers functioning in the pseudoclocal region of zebrafish have acquired new function in external genitalia and limbs. Aside from the fact that the ancestral state is not adequately resolved to interpret co-option, detailed above, the authors have not shown that the same enhancers that function in the UGS and limbs of mice have pseudocloacal function in zebrafish.

1)Once the authors shifted to an effort to determine if 5DOM regulation of hoxd13a in zebrafish is conserved in the UGS of mice, they do not return to the function of specific enhancers in zebrafish. They show that chromatin accessibility in the mouse UGS maps to already identified enhancers, GT2, Island E, and CsB that were previously demonstrated to have function in mouse genital tubercle and/or distal limbs. They do not show that orthologs of these sequences are required for pseudocloacal development or sufficient for pseudocloacal expression in zebrafish. The 5DOM is a huge region to make such an assumption that these enhancers were co-opted from an ancestral function in the pseudocloacal region to a derived function in the genital tubercle and limb without any test of their function in zebrafish.

2)In the results section entitled 'An ancestral regulatory landscape for an ancient function', the authors simply test whether Hoxa13a, hoxa13b, and hoxd13a are together required for pseudocloacal development. This genotype combines hoxd13a and two paralogs of hoxa13 – three distinct loci. The result shows that Hox13 genes are necessary for properly terminating the digestive and urogenital systems, but it does not show that there is an ancestral regulatory landscape. One could similarly state that loss of the autopod after combined loss of HoxA13 and HoxD13 in mouse limbs suggests an ancestral regulatory function prior to the first vertebrate whole genome duplication.

3)They focus on the previously identified enhancers, but it seems that there is an opportunity for comparative genomics between zebrafish and mouse to more comprehensively identify enhancers functioning in the UGS and pseudocloacal region, assuming the 'posterior trunk' of the zebrafish that was used for ATAC-seq and ChIP-seq encompasses this region. Is the 5DOM regulatory control in both species indeed ancestral and using the same sequences?

In summary, the genetics of the paper are strong. The authors have compellingly shown that the 5' HoxD regulatory REGION is required for zebrafish pseudocloacal and mouse UGS development. Although they have not demonstrated necessity, they have shown that the tetrapod limb/genital enhancers are also sufficient to drive mouse UGS expression. The major weaknesses of the paper are in the evolutionary interpretations, and they are substantial weaknesses. I'm not sure if the authors intended to imply that the same ENHANCERS functioning in the zebrafish pseudocloacal region were co-opted to control expression in the limbs/external genitalia, but that is the definition of regulatory co-option as I understand it. An alternative explanation, which I wouldn't consider co-option, is that the 5DOM in zebrafish may have pseudocloacal enhancers, and tetrapods independently evolved UGS/GT/limb enhancers in the same region that is hundreds of kilobases in size. This could be due to ancestral accessibility of the 5DOM in posterior/distal regions of the body that would thus be well-suited to repeated regulatory evolution for development of these structures.

Minor points:
1)The authors switch from mouse HoxD to the zebrafish hoxda cluster without reference to the teleost whole genome duplication. This is important background for a general audience to understand the difference in nomenclature.
2)They show expression of hoxd4a and hoxd10a in Figure 1 but then interpret the experiment to suggest that 3DOM controls the expression of 'hoxd3a to hoxd10a'. This should be clarified to avoid overinterpretation, or expression of the other hoxda genes should also be shown.
3)"We tested these three putative enhancers in an enhancer-reporter assay…" should be clarified to state this was done in mice. The two previous sentences reference bony fish and mammals, and so the species of study is not clear.
4)Deletions in the mouse 5DOM are huge. There is a scalebar in the figure, but it would help to note these deletion sizes in the text.

Referee #3

(Remarks to the Author)
I really enjoyed reading the manuscript of Hintermann et al. They have used several different assays to document that the regulatory landscape 5DOM, involved in the evolution and shaping of digits by controlling the expression of posterior Hox genes, has an ancestral function in patterning the cloacal area. The study's relevance is heightened by the variety of approached they employed. Apart from very few typos in figure legends and M&M paragraphs (e.g. presence/absence of space between numbers and units of measurement), the text is well balanced among chapters and paragraphs, coherent, extremely clear also for non-evolution people. In brief, I found the text very explanatory, and in general the work done rigorous and solid, and the presentation (figures, tables, etc) of high quality. I, therefore, think that the present experimental manuscript add novel issues to our comprehension of how an important patterning mechanism evolved in vertebrates.

I feel in the fortunate condition of being able to recommend the acceptance in its exact form, not being able to suggest any further experiment or text revision.

Version 1:

Reviewer comments:

Referee #2

(Remarks to the Author)
First, my apologies that this review was delayed as I react and respond to the apparent dismantling of federal research support in the US.

I have carefully read the revised manuscript; it is much improved. The authors demonstrate a requirement for the 5' hoxd genes and the orthologous 5DOM regulatory region in zebrafish pseudocloacal and mouse urogenital development. They show that, although the specific cis-regulatory elements do not appear to be shared, the global cis-regulatory landscape appears to have been co-opted from an ancestral function in genital development to also control digit development in tetrapods. This helps to explain the deep evolutionary conservation of regulatory synteny in this cluster.

My earlier concerns about the communication of 'regulatory co-option' have been addressed by revisions to the text and with new data. Specifically, the end of the summary and the last few paragraphs of the introduction are much improved to more specifically and accurately communicate what is meant by 'regulatory landscape co-option'. The explanation of the whole genome duplication and subsequent loss of hoxdb is also helpful to those who may be less familiar with the details needed to understand the overall premise of this paper.

The additional data on fish enhancers and their subtle deletion phenotype is very helpful to clarify that this paper doesn't imply specific enhancer co-option but rather explains the function and preservation of the 5DOM region as a whole. I also like the introductory paragraph in the section titled, "An ancestral regulatory landscape for an ancient function", because it better establishes the emphasis on the question of why regulatory synteny might be preserved for the 5DOM if it isn't required for fin development in fishes.

I encourage the authors to maintain the clarity of these points if they are required to shorten the manuscript.

In the revision, I notice that replicate numbers are missing for most experiments and are surprisingly small for zebrafish experiments (e.g., wild type and the simpler mutant configurations in Figure 5). This should be addressed before acceptance.

Remaining minor points:

Line 214, 'allelic series' does not seem to be the appropriate language. Perhaps instead "these genetic configurations"

In figure 3C, I would also illustrate the product of the inversions to make it explicitly clear how 5DOM is moved far from the gene cluster in the second but not in the first configuration.

Referee #4

(Remarks to the Author)
In this work, the authors study the conservation of the 5' and 3' regulatory landscapes of Hoxd clusters in vertebrates. It is well known that these genes have a bipartite regulation, with the anterior genes mostly regulated by the 3' regulatory domain and the posterior genes by the 5' domain, which leads the expression of these genes to the digits in tetrapods. They find that this 5' domain contains enhancers that drive Hoxd13 to the urogenital sinus in mice and its homologous structure, the cloacal region, in zebrafish. They conclude that the current transcriptional regulation of Hoxd genes in distal limbs (digits in tetrapods) has been coopted from the regulation at the cloaca in the vertebrate common ancestor.
The article is well-written, clear and easy to follow. The authors have used a good variety of approaches to support their claims, and the conclusions are in general original and well-funded in the results. In my opinion this will be a very interesting study for the scientific community, and thus I happily endorse it for publication in Nature.
However, I did find several details that should be corrected. The most important of them is related to the enhancers responsible of hoxd13a expression in the zebrafish fin. In their deletion assays, the authors don't find a big part of this regulatory information, and the mutant fish with no 5' domain still show hoxd13a expression in the fin. Although the authors claim that the expression is milder in these mutant animals than in wildtypes, which seems to be true according to the pictures (Fig. S5), the expression is very variable and other staining methods would be necessary to be able of measuring and comparing it, such as HCR. In any case, it is obvious that most of the enhancers responsible of this regulation are out of the deleted region. According to the diagrams showed in Fig. 2a and S5a, the deleted region goes from the hox genes to the atp5mc3a/atf2 genes, not including them. However, in the HiC data showed in Fig. S2 (although it was not performed in fin tissue), we can observe that the TAD containing the 5' domain encompasses these two genes, and thus this region not deleted by the authors might contain part of this missing regulatory information. Furthermore, 4Cseq data reported in Acemel et al. 2016 (PMID: 26829752) in zebrafish support this hypothesis, showing strong contacts between hoxd13a promoter and the atp5mc3a/atf2 genes region. Although new deletion experiments would not change the overall results of the article, I think this issue should be considered by the authors and properly discussed in the manuscript.
Other minor comments are:
- L119: Fig. S2c should be Fig. S3c.
- L166: hoxda10a should be hoxd10a.

- L169: Apart from Fig. 2, Fig. S6b should also be called here. Indeed, Fig. 2 and Fig. 6 are a bit confusing, since very similar information is showed in both. For example, Fig. 2b and S6b are showing the same experiment at different zoom level. Young embryos showed in Fig. 2a would benefit of an inset with a zoom in as well. In general, I think these two figures could be better organized.
- L172: Is the expression of the genes hoxd10a and hoxd4a at 72h showed somewhere?
- L235, 236: I don't find that the ATAC peaks are "positioned at regions that are relatively depleted for the K3K27ac mark". Indeed, I see in Fig. 4a that they co-localize with an ATAC-enriched region. Besides, I don't see the point that the authors want to make, because this would mean that these are not active regions.
- L283-285: I do not agree with the authors about the number of peaks in the cloaca and pectoral fins cell types. I think the cloaca pseudobulk comes from a very small number of cells, and thus it shows a high background ratio, while the pectoral fins samples are more populated, and the signal/noise ratio is much better. I would trust in the peaks showed in fins, but I am not very sure about the cloaca, I think many of these 'peaks' are probably not real signal. In any case, even if I trust these data, I don't see "many more peaks" in the cloaca than in the fins.
- L322: Apart from calling the Fig. 5b, f, S14b, c, e and f should also be called. Alternatively, I would remove this figure calling and call them together in the next line, Fig. 5b, f and Fig. S14.
- L342, 343: Again the issue addressed in my first point.
- Materials and methods: Mouse ChIP-Seq section should be called just ChIP-Seq, since also zebrafish ChIP-Seq re-analysis is detailed here. Alternatively, this section could be split in two.
- L1197: bard should be bars.
- Figure 6 is not called in the text.
- Fig. S2 (and corresponding materials and methods): Is the CTCF ChIP-Seq signal shown in the bottom panel coming from the merge of 24 and 48hpf samples? This must be clarified.
- Fig. S7a: male and female symbols should be added to these diagrams.
- Fig. S7c, d: the organization of these panels is confusing, taking into account how they are called in the text. I would rather put both diagrams together in panel c, and add new letters d and e for the panels with pictures.
- Fig. S11a: looking at these tracks with a high zoom level and autoscale makes the reader think that very different signals are comparable, like head or cloaca and tailbud mesoderm, when they are actually very different, as can be observed in Fig. S11. I think this panel should be corrected, either increasing the region shown (decreasing the zoom level) so we could have other peaks as a reference, or deactivating the autoscale of the browser and fixing the scale in reasonable heights, like the ones used in Fig. S11.

Referee #5

(Remarks to the Author)
The manuscript by Hintermann, Bolt, Hawkins et al. investigates the regulatory evolution of the HoxD cluster in the context of fin/limb development. In mice, the HoxD cluster is flanked by two distinct regulatory domains, 3DOM and 5DOM, which control HoxD expression in the proximal (stylopod/zeugopod) and distal (autopod) fin/limb regions, respectively. Previous studies also have revealed several limb enhancers within these domains. Furthermore, the complete deletion of either the 3DOM or 5DOM domains abolishes HoxD gene expression in the corresponding limb spatial domains (proximal or distal, respectively).

In this study, the authors explore whether HoxD fin/limb regulation is conserved across vertebrates, by performing large deletions of the homologous domains in zebrafish. Deletion of the 3DOM results in effects similar to those observed in mice, characterized by the loss of the limb expression of hoxd10a and hoxd4a. This demonstrates that the enhancers controlling the expression of anterior hoxd genes (hoxd3a–hoxd10a) are located within the 3DOM, as in tetrapods, suggesting a deep evolutionary conservation for this regulatory domain.

Unexpectedly, the deletion of the 5DOM in zebrafish does not affect the expression of posterior hoxd genes, unlike in mice. This is evidenced by the expression pattern of hoxd13a, which remains largely unaffected, only showing minor defects at later stages. These results indicate functional differences in HoxD posterior limb regulation between vertebrate species, suggesting that in zebrafish the regulatory potential may be embedded within the HoxDa cluster itself.

Since the 3DOM and 5DOM also control HoxD expression in other developing tissues, the authors investigated expression patterns by whole-mount in situ hybridization (WISH) in embryos. Notably, these analyses revealed that Hoxd13 expression is completely abolished in the pseudocloacal region of zebrafish mutants carrying a 5DOM deletion. This prompted the authors to explore whether the 5DOM also regulates HoxD gene expression in the mouse urogenital sinus (UGS), a structure homologous to the zebrafish cloaca. Through the analysis of various mouse mutant alleles, they confirm that the 5DOM indeed drives UGS expression. Epigenetic experiments helped to pinpoint specific enhancers within this domain, which were validated through LacZ reporter assays.

To assess evolutionary conservation, the authors performed multispecies alignments and compare relevant ATAC-seq datasets, identifying putative enhancers in zebrafish and mice. However, these analyses reveal limited conservation between species. One conserved enhancer, CsB, previously shown to have regulatory activity, was deleted in zebrafish, revealing its limited functionality during fin or pseudocloacal development. Finally, the authors demonstrate that hox13 genes are essential for cloacal development by analyzing the phenotypes of combined deletions of hoxd13a, hoxa13a, and hoxa13b.
Overall, through a series of elegant genetic experiments, the authors establish that the 5DOM is necessary for controlling Hoxd13 expression in the cloaca/UGS, suggesting that this function is deeply conserved during vertebrate evolution.

Additionally, they demonstrate that posterior HoxD gene regulation in the fin/limbs differs between zebrafish and mice, with 5DOM sustaining expression only in the latter. Based on these findings, they propose that the ancestral function of the 5DOM was to regulate HoxD expression in the cloaca, and that it was later co-opted during evolution to support digit expression.

This is not a “typical” case of co-option, where a gene or particular enhancers may have evolved to support a new function. Instead, the authors argue that the regulatory landscape has evolved as a whole, which is an important concept for the field. However, the situation is complex, as the zebrafish 5DOM also contais distal regulatory potential in fins (as reported in Gehrke et al., PNAS 2015 and Schneider et al PNAS 2011), but that are mostly functionally dispensable. The authors argue that zebrafish may represents an partial step in the full co-option that is observed in tetrapods. While this is a compelling hypothesis, the evidence remains still preliminary.

The study only involves the functional characterization of two species, making it difficult to infer whether the zebrafish condition represents an ancestral state or a derived feature (this interpretation is also considered in the discussion). In this section, the authors also argue that their main conclusion is supported by the existence of Hoxd13 cloacal expression in lampreys, a species that lack fins (digits). While this implies that Hoxd13 cloacal expression likely predates fin/limb expression, it does not confirm that cloacal regulation was initially sustained by the existence of distal elements within a 5′ flanking regulatory domain. Additional evidence may be required to support this conclusion, which is central to the manuscript and in it is reflected in the title.

Given the technical challenges of performing large deletions in other relevant species (already performing those in zebrafish is an impressive achievement), alternative approaches could provide further support for the authors claims. For example, investigating whether Hoxd13 cloacal expression in lampreys is controlled by enhancers within a 5DOM-like structure could be informative. Identifying a similar 5′ flanking regulatory domain in cyclostomes, which lack fin/limbs, would support the hypothesis that cloacal distal regulation is ancestral. These experiments could be complex to interpret, due to the genome triplication of cyclostomes, but may result very informative. Experiments such as Hi-C and ATAC-seq in these species could help to address this question.

An interesting observation is that ATAC-seq comparisons between zebrafish fins (Figure S11) and mouse distal limbs (Bolt et al., Nat. Commun., 2022) reveal significantly more peaks in the latter. This may serve as an indirect indicator of regulatory potential. Performing similar analyses in other species, such as cartilaginous fishes or the spotted gar, may serve to strengthen the authors claims. If ATAC-seq profiles from these species also show a reduced number of fin-specific peaks in the 5DOM, similar to zebrafish, this would support the hypothesis that distal limb regulation has evolved later in tetrapods.

The results in the manuscript also suggest that distal HoxD expression in zebrafish is driven by regulatory elements within the hoxda cluster itself. This contrast with the mouse, where regulation is external to the cluster and located in the 5DOM (as shown in previous studies from the authors). This raises the possibility that posterior HoxD limb regulation was ancestrally located within the cluster and later relocated to the 5DOM in tetrapods. Alternatively, distal regulation may have been the ancestral state, with zebrafish having evolved a more proximal regulatory mechanism. For instance, do specific peaks within the zebrafish hoxda cluster support posterior gene expression? Are these regulatory regions functionally conserved outside tetrapods? Investigating other species may also help to clarify this issue. Are BACs from other relevant species (i.e. spotted gar, sharks, skates…) containing the HoxD cluster and that could be used to investigate the regulation within the cluster?

Minor comments:

1-Lines 235-236: “The ATAC peaks were positioned at regions that are relatively depleted for the H3K27ac mark (Fig. S9)”

I am not sure of what the authors mean with this statement. The ATAC-seq peaks, in particular Island E and CsB seem to be at positions that display abundance of H3K27 marks.

2-Figure 3d: Please indicate the Y axis. Are the expression data derived from the experiments presented in Figure s8?

3-Lines. 358-360: “Alternatively, it may reflect a secondary loss of several distal enhancers occurring in aftermath of the teleost whole genome duplication as such event may lead to the reorganization of paralogous regulatory landscapes”

Although enhancer loss is a recurrent phenomenon after WGD, the alteration of regulatory activity in zebrafish does not need to be necessarily associated with this evolutionary event. In fact, the alignment analyses performed by the authors actually point to a limited impact of WGD in the 5DOM. Thus, the secondary loss of regulatory elements could have also happened in the Actinopterygian lineage, independently of the WGD.

Version 2:

Reviewer comments:

Referee #5

(Remarks to the Author)
After carefully evaluating the revised version provided by the authors, which includes additional experiments (HCR) and some changes in data presentation and interpretation, I am happy to support the manuscript for its publication in Nature.

A particular strength of the study is the demonstration that changes in gene expression may not only result from variation of individual enhancers, but also from alterations in regulatory landscapes (TADs) that evolve as integrated units. Thus, the study encourages a more holistic view on the evolution of gene regulation, which is an important concept for the field. In addition, the generation of large deletions in zebrafish represents a significant technological advancement that opens new perspectives for large-scale comparative analyses of gene regulation, and may stimulate future research.

Although some related questions may still remain open, I agree with the authors in that the associated difficulties likely place them beyond the scope of the current study (though they may be worth exploring in the future).

# CO-OPTION OF AN ANCESTRAL CLOACAL REGULATORY LANDSCAPE DURING THE EVOLUTIONARY EMERGENCE OF DIGITS AND GENITALS

Aurélie Hintermann[1,@,*], Christopher Chase Bolt[2,¶,*], M. Brent Hawkins[3,*], Guillaume Valentin[2], Lucille Lopez-Delisle[2], Sandra Gitto[1], Paula Barrera Gómez[2], Bénédicte Mascrez[1], Thomas A. Mansour[4], Tetsuya Nakamura[4], Matthew P. Harris[3], Neil H. Shubin[5] and Denis Duboule[1,2,6,α]

**Responses to referee's comments**

---

Dear Editor,

Please find below our point-by-point answers to the two referees (referees #2 and #3) who had commented on our MS #2024-03-05222. One referee '*enjoyed reading*' the paper and felt in the '*fortunate condition of being able to recommend the acceptance in its exact form*'. The other referee was positive too, but had several comments and requests, and also found that the main issue was not completely substantiated.

These comments catalyzed us to add new experiments and modify the text in ways that we think improve the paper. Amongst new genomic analyses, we generated a *de novo* deletion of a regulatory sequence on top of a chromosome containing a new transgene labelling endogenous *hoxd13a* expression. This in part explains the rather long delay in submitting this revised version. Because of the additional work, we have added former author #3 as an additional shared first co-author. As main modifications (described below), we have now introduced three new supplementary figures, each of them addressing one of the main concerns of referee 2; the first is dealing with genomic alignments of conserved sequences and the 'zebrafish *versus* gar' argument. The second shows chromatin accessibility profiles on few, but specific, cells (made possible by mining novel datasets released in 2024) and the third reporting the deletion of one enhancer candidate in zebrafish and the resulting effects.

Referee comments *in extenso* are italicized and our responses are in plain text, for the sake of clarity.

---

**Referee #3 (Remarks to the Author):**

*I really enjoyed reading the manuscript of Hintermann et al. They have used several different assays to document that the regulatory landscape 5DOM, involved in the evolution and shaping of digits by controlling the expression of posterior Hox genes, has an ancestral function in patterning the cloacal area. The study's relevance is heightened by the variety of approached they employed. Apart from very few typos in figure legends and M&M paragraphs (e.g. presence/absence of space between numbers and units of measurement), the text is well balanced among chapters and paragraphs, coherent, extremely clear also for non-evolution people. In brief, I found the text very explanatory, and in general the work done rigorous and solid, and the presentation (figures, tables, etc) of high quality. I, therefore, think that the present experimental manuscript add novel issues to our comprehension of how an important patterning mechanism evolved in vertebrates.*

*I feel in the fortunate condition of being able to recommend the acceptance in its exact form, not being able to suggest any further experiment or text revision.*

We are grateful to this expert for this positive comment.

**Referee #2 (Remarks to the Author):**

*In this manuscript by Hintermann and Bolt et al, the authors establish that there is synteny and an ancestral TAD organization to the large 5' and 3' regions flanking the Hoxd cluster in zebrafish that was previously well-characterized in tetrapods, largely mice. They further relate this to the observation that HoxD cluster expression is similar in zebrafish fins and mouse limbs with hoxd9a and hoxd10a expressed in the anterior and hoxd11a-hoxd13a restricted to the posterior fin, reflecting a common proximal/anterior and distal/posterior axis.*

*They ask, is this pattern of expression in zebrafish fins controlled by 5' and 3' flanking regions in the same way as in mouse limbs, despite the fact the distal fin will not form digits? The authors first find that the 3DOM regulatory region is required for hoxd4a and hoxd10a but not hoxd13a expression in zebrafish fins, as in mouse limbs. However, the 5DOM is not required for hoxd13a expression in the fin as it is in mouse limbs.*

*This leads the authors to ask what other function the 5DOM might play in Hox regulation in both species. They convincingly show that the 5DOM contains regulatory sequences that are required for hoxd13 expression in the zebrafish pseudocloacal region and in the mouse urogenital sinus, two homologous structures. The genetic analyses of this paper are strong.*

Thank you for these positive comments and excitement over the findings.

*They then conclude, as stated in the title, that the regulatory landscape controlling Hoxd13 expression in the digits and genitals came about by "co-option" of a regulatory landscape that ancestrally functioned in the cloaca. I do not see support for this conclusion that is the major crux of the manuscript.*

*1) The authors see no effect of 5DOM deletion on hoxd13a expression in zebrafish fins,*

This is not our interpretation and we strived to make this clearer in the revision. We note that pectoral fin expression in the 5DOM deletions is attenuated, albeit variably so. Fig. 1c and Suppl. Fig. 5 where a collection of fins is shown precisely to illustrate the variability of this result.

*but they note that "previous transgenic results using components of the fish 5DOM sequences in transgenic mice showed sufficiency to drive expression (line159-160)". This references two papers where the Shubin laboratory demonstrated sufficiency of conserved fish enhancers identified in the 5' regulatory region of the HoxD cluster to drive reporter expression in mouse limbs. Notably, the sequences drove much higher and more distal (tetrapod-like) expression when originating from the gar genome than when originating from the zebrafish genome.*

As we note in the text, this comment is consistent with our results. The slightly attenuated expression shows that some regulatory activity resides in the zebrafish 5DOM but, since a significant amount of expression is retained, much of this activity, or compensating activity lie elsewhere. Our main point is that this situation is very different from mice where, unlike fish, a similar deletion abrogates expression entirely.

*2) In the discussion, the authors acknowledge this "may reflect a secondary loss of several distal enhancers associated with teleost whole genome duplication… (which may be solved) with a comparable deletion and epigenetic characterization of 5DOM in more basal fish species such as gar… (lines 288-290)"*

This is an excellent point but one where experiments would be challenging, or even impossible, to do at this point. Such molecular approaches have never been carried out in the gar (the large deletions we show are already novel in zebrafish), and establishing the necessary stable mutant lines is not feasible as this species takes about 3 years and considerable space to reach sexual maturity. Regarding the epigenetic approaches in the gar, this is equally challenging, since they would require complete multi-omic datasets to be able to identify (clustering analysis) the same cellular populations as we now show in zebrafish Suppl. Fig 11, see below). Indeed, the initial 'dissection approach' (that may be carried out in the gar, and still…) is not precise enough given too many contaminating cells as we now demonstrate in the same figure.

Despite the empirical challenges, the point raised by this referee is an important one, i.e., that of the potential problem created by the additional genome duplication in teleosts and the possibility that it reshuffled some enhancers between the paralogous copies, thus blurring our analyses. To address their question, we made an extensive genome comparison between teleosts and non-teleost (including the gar) fishes and we now show that, regarding the *hoxda* gene cluster, the patterns of conservation are virtually identical, with only rare and minor differences following phylogenetic distances. This was not completely unexpected, for zebrafish only has 7 copies of *Hox* clusters, the missing one being precisely the elusive *hoxdb* cluster. Under these conditions, this is no surprise that overall structure and conserved sequences within 5DOM, are almost identical between zebrafish and non-teleost species, relative to the amniote counterparts. To triple-check, we also made sure that none of those 5DOM sequences would be located elsewhere in the ZF genome. The final direct comparison between the *hoxd* loci supported this notion.

We thank this referee to have raised this concern, and we have now included a full new Supplementary Fig. 10 to illustrate these points. The text was amended accordingly.

*3) Also in the discussion, they state that "most of the regulatory control (in zebrafish) likely resides within the gene cluster itself", but they do not provide the same sort of transgenic evidence for the sufficiency of reporter expression from within the zebrafish gene cluster, similar to the HoxD cluster BAC reporter that they tested in mice.*

This referee is correct on this point.

Here, in essence is the logic behind our inference. There are two main ways to localize the control of a regulatory element; either by a transgenic approach, or by a targeted deletion approach. The former is the easiest one of course, yet it may be the source of numerous artifacts, in particular in 'novel' experimental systems such as zebrafish (copy number, insertion site impacting specificity etc), which have been well described in the literature. On the other hand, the deletion approach is by and large most significant, functionally speaking, even though one often has to deal with a subtractive readout, rather than with an additive one. This is why the latter approach has to be backed up with internal positive controls. This is exactly what we do in this work. We delete 3DOM with no distal effect on *hoxd13a*; we then delete 5DOM in both cases with positive controls. With these two deletions, the two entire TADs are removed (in two separate deletions), the only piece of DNA remaining in both deletions being the *hoxda* cluster. We naturally conclude that the regulation must be found there. Because there is a TAD border right after *hoxd13a* and the rest of the cluster (Gomez-Skarmeta lab and our previous work), which thus prevents (or at least greatly reduces) enhancer-promoter contacts, we logically speculate that an elusive fin regulation may be close to this latter gene.

*4) In the summary Figure 6, they seem to interpret the fact that 5DOM deletion in zebrafish causes a slight and transient loss of hoxd13a to suggest that the "5DOM lightly contributes to hoxda gene regulation in…paired fin buds". I don't know what 'lightly' means in a cis-regulatory or evolutionary context.*

We understand this comment and agree that the word 'lightly' is not very informative. We have changed this wording and better explain what we mean by this, i.e., that the effect of removing 5DOM variably attenuates expression. We believe the wording is clearer in the revised version.

*The authors have therefore not convincingly shown that limb/genital regulation of hoxd13a from within the 5DOM is a derived state that wasn't also functioning to control fin expression in more basal fishes.*

There are two responses to this important criticism. First: *hoxd13* is expressed in the cloaca of digit-less fishes all the way to the base of vertebrate phylogeny, including jawless fish. Second, we have performed new ATAC analyses and dataset mining (see below) that indicate that 5DOM in zebrafish indeed has regulatory activities, yet distinct from those seen in mice, thus leading to the suggestion for the co-option of a 'regulatory structure' (landscape) for distal regulation in mice (not of specific enhancers, as discussed below).

*Second, an interpretation of cis-regulatory co-option requires evidence that the same enhancers functioning in the pseudoclocal region of zebrafish have acquired new function in external genitalia and limbs.*

We appreciate this comment because it forces us to be clearer about our argument. Given the large evolutionary distance separating mouse and teleost fishes, we decided to consider the 5DOM as a whole, rather than its individual enhancers. At such distance, the conservation of non-coding sequences is minimal. Nevertheless, as previously documented (Hintermann et al. 2022), the overall regulatory outcome of a regulatory landscape can be conserved even if the function of individual enhancers is not. Therefore, we believe that basing our research on sequence conservation of individual enhancers would be less appropriate (thought it may be valid in more closely related species), and we decided to focus on the entire landscape.

Accordingly, we do not claim that fish sequences were co-opted to achieve a different function in amniotes. We claim that a *regulatory landscape* was co-opted, as a sort of regulatory playground where novel regulatory patterns can emerge faster and more efficiently than elsewhere (as discussed mechanistically in some references cited in the paper). This may be due to an efficient readout, ready to be used (*Hox* genes) and the presence of a 'regulatory architecture' (TADs) prone to efficient enhancer selection and improvement. In this process, there may be some enhancers co-opted to work in multiple, pleiotropic contexts, and we show examples of this in the paper, with the mouse limbs-genitals-UGS connection. The main message here is that the landscape has to be considered as a single, large regulatory unit, which evolved as such.

*Aside from the fact that the ancestral state is not adequately resolved to interpret co-option, detailed above, the authors have not shown that the same enhancers that function in the UGS and limbs of mice have pseudocloacal function in zebrafish.*

Although we maintain that the focus should be centered on whole-landscape regulation and co-option (as discussed above), to address these queries and to have a more global comparative view of sequence accessibility between zebrafish and mouse within 5DOM, we now include new data of aligned ATAC seq profiles either produced out of dissected material, or from mining new scATAC-seq datasets produced recently (new Suppl. Fig. 11, that includes now a panel that was removed from former Suppl. Fig. 4 to group all ATAC data together). Additionally, we created new deletion mutants of CsB that shows conserved modulation of expression in the cloaca as well as UGS, whereas no obvious effect in the fin.

We would like to note that functional attribution of individual enhancers to *Hox* cluster regulation is complex since, even between amniotes (birds and mammals) comparable enhancer functions can be carried out by distinct DNA sequences (e.g., Hintermann et al. 2022). Thus, the single enhancer 'trees' may not specifically define the "forest" of the landscape. We feel that our new experimental and *in silico* mapping data specifically addresses the nature of 5DOM enhancer conservation and lack of fin regulatory program within.

*1)Once the authors shifted to an effort to determine if 5DOM regulation of hoxd13a in zebrafish is conserved in the UGS of mice, they do not return to the function of specific enhancers in zebrafish. They show that chromatin accessibility in the mouse UGS maps to already identified enhancers, GT2, Island E, and CsB that were previously demonstrated to have function in mouse genital tubercle and/or distal limbs. They do not show that orthologs of these sequences are required for pseudocloacal development or sufficient for pseudocloacal expression in zebrafish. The 5DOM is a huge region to make such an assumption that these enhancers were co-opted from an ancestral function in the pseudocloacal region to a derived function in the genital tubercle and limb without any test of their function in zebrafish.*

Thank you for this comment. To directly address it, we have added data in which we deleted the zebrafish CsB counterpart from a fish chromosome where we had also introduced a reporter into *hoxd13a*. The deletion of CsB had little effect -if any-, on distal appendage expression. Interestingly, deletion of CsB in the zebrafish showed transient upregulation of *hoxd13a* activity – not loss –

suggesting that a simple necessity argument may not be sufficient for understanding output of the broader landscape. Notably, mouse also show similar expression regulation of *Hoxd13* when individual 5DOM enhancers are removed. Mouse CsB was never deleted separately in amniotes. Mouse CsB as a transgene seems to be active in the UGS, the distal limbs and the genitals, as well as in V0 interneurons, which was likely its ancestral function in both amniotes and fishes (Sordino et al., 1996).

We believe our added zebrafish experiment, which is now detailed in a new Suppl. Fig. 12, tells us that the concept of 'co-option' cannot be broken down to a single 'enhancer' sequence having common and distinct functional specificities between fish and amniotes. While this may be the case within amniotes, the idea of co-option between fishes and amniotes must be considered at a more global level, i.e., that of the general outcome of 5DOM, rather than from its components. We believe that the very different distributions of ATAC-seq peaks within fish and mouse 5DOM are an illustration of this novel way to consider developmental gene regulation (and evolution).

*2)In the results section entitled 'An ancestral regulatory landscape for an ancient function', the authors simply test whether Hoxa13a, hoxa13b, and hoxd13a are together required for pseudocloacal development. This genotype combines hoxd13a and two paralogs of hoxa13 – three distinct loci. The result shows that Hox13 genes are necessary for properly terminating the digestive and urogenital systems, but it does not show that there is an ancestral regulatory landscape.*

The clearest comparable effect of removing 5DOM either in fishes or in mice is upon the pseudo-cloacal region or the UGS, respectively. Therefore, we claim that this regulation (once again associated to 5DOM rather than to specific sequences) is an ancestral feature, as shown by the 5DOM deletions in mouse and zebrafish, and not by zebrafish *hox13* deletions.

Because the cloacal region is certainly more ancestral than digits or genitals, we then proposed that an ancestral regulation of 5DOM was associated with an ancient function of *Hoxd13* in this anatomical structure. To strengthen this argument, we had to show the functional necessity of *hoxd13a* in cloaca patterning, which was not yet established in zebrafish. This was the point here, not to re-establish that the 5DOM is ancestral regulatory landscape, which we had already shown.

We have now added supplementary explanations in the introduction to the result section 'An ancestral regulatory landscape for an ancient function*'.*

*One could similarly state that loss of the autopod after combined loss of HoxA13 and HoxD13 in mouse limbs suggests an ancestral regulatory function prior to the first vertebrate whole genome duplication.*

The combined loss of HOX13 functions in mouse induces a loss of autopods, but also of the external genitals and of the UGS, i.e., three main regulations associated with 5DOM (for the *HoxD* cluster, of course). Of these three structures, phylogenetic evidence shows that only a cloaca (which has hox expression, evolved prior to the 'first vertebrate genome duplication'.

*3)They focus on the previously identified enhancers, but it seems that there is an opportunity for comparative genomics between zebrafish and mouse to more comprehensively identify enhancers functioning in the UGS and pseudocloacal region, assuming the 'posterior trunk' of the zebrafish that was used for ATAC-seq and ChIP-seq encompasses this region. Is the 5DOM regulatory control in both species indeed ancestral and using the same sequences?*

We thank the reviewer for this point. In review, we now directly address this question with the new Suppl. Fig. 11, as described above.

*In summary, the genetics of the paper are strong. The authors have compellingly shown that the 5' HoxD regulatory REGION is required for zebrafish pseudocloacal and mouse UGS development. Although they have not demonstrated necessity, they have shown that the tetrapod limb/genital enhancers are also sufficient to drive mouse UGS expression.*

Thank you. We are pleased that the reviewer is convinced of our compelling data and strength of the supporting genetic data. Three different points in this short summary needing discussion.

◊ We do center our discussion on the broader region (defined as a 'landscape' in 2003).

◊ Secondly, we do show the *necessity* of this 'region' for the UGS as demonstrated by the almost complete loss of *Hoxd13* transcripts in the sub-deletions used in this paper, added to the fact that in the absence of HOX13 function, there is no UGS at all (papers by Warot et al., and Kondo et al., referred to in the paper).

◊ Finally, we show that some UGS enhancers are also functional in limbs or genitals. For instance, the best ever described distal limb enhancer within 5DOM (island II) is not at all required for expression in the UGS. If anything, the deletion of the sub-region including it slightly increases *Hoxd13* transcription.

*The major weaknesses of the paper are in the evolutionary interpretations, and they are substantial weaknesses. I'm not sure if the authors intended to imply that the same ENHANCERS functioning in the zebrafish pseudocloacal region were co-opted to control expression in the limbs/external genitalia, but that is the definition of regulatory co-option as I understand it.*

Thank you for this criticism. We did not want to imply that. We speak in favor of the alternative explanation mentioned by this expert reviewer just below. In fact, we think we did not do a good job in clearly explaining what we meant by the 'co-option of a regulatory landscape' (in the title) and hence this long discussion with referee 2 on this topic.

*An alternative explanation, which I wouldn't consider co-option, is that the 5DOM in zebrafish may have pseudocloacal enhancers, and tetrapods independently evolved UGS/GT/limb enhancers in the same region that is hundreds of kilobases in size.*

Yes, this is a possibility, but the tetrapod 5DOM had to start from somewhere and we suggest that the 5DOM regulation in an ancestral cloacal region (present e.g., in agnathans) provided the 'playground', the 'framework conditions' to subsequently evolve new regulations, by using or not using the same individual enhancers in the domain.

*This could be due to ancestral accessibility of the 5DOM in posterior/distal regions of the body that would thus be well-suited to repeated regulatory evolution for development of these structures.*

We agree: the 'posterior and distal' region is, indeed, the cloacal region.

***Minor points:***

*1) The authors switch from mouse HoxD to the zebrafish hoxda cluster without reference to the teleost whole genome duplication. This is important background for a general audience to understand the difference in nomenclature.*

Thank you. In response, we now strive better introduce this important point in the text.

*2) They show expression of hoxd4a and hoxd10a in Figure 1 but then interpret the experiment to suggest that 3DOM controls the expression of 'hoxd3a to hoxd10a'. This should be clarified to avoid overinterpretation, or expression of the other hoxda genes should also be shown.*

There are many complete accounts of this in the literature and we thought it would not be necessary to produce one more for this paper (or to use 20 years old datasets..). The best and most complete atlas of *Hox* gene expression in the ZF fins was published by Ahn and Ho (2008) and we now also refer to this paper for this specific purpose (formerly reference #16).

*3) "We tested these three putative enhancers in an enhancer-reporter assay…" should be clarified to state this*

*was done in mice. The two previous sentences reference bony fish and mammals, and so the species of study is not clear.*

Sorry for this. This is now clarified

*4) Deletions in the mouse 5DOM are huge. There is a scalebar in the figure, but it would help to note these deletion sizes in the text.*

These deletions are indeed the largest ever engineered in ZF. The exact sizes are now indicated in the plain text.

**CO-OPTION OF AN ANCESTRAL CLOACAL REGULATORY LANDSCAPE DURING THE EVOLUTIONARY EMERGENCE OF DIGITS AND GENITALS**

Aurélie Hintermann[1,@,*], Christopher C. Bolt[2,¶,*], M. Brent Hawkins[3,*], Guillaume Valentin[2], Lucille Lopez-Delisle[2], Madeline M. Ryan[3], Sandra Gitto[1], Paula Barrera Gómez[2], Bénédicte Mascrez[1], Thomas A. Mansour[4], Tetsuya Nakamura[4], Matthew P. Harris[3], Neil H. Shubin[5] and Denis Duboule[1,2,6,α]

**Responses to referee's comments**

---

Dear Editor,

Please find below our point-by-point answers to the referees who had commented on our revised version of MS #2024-03-05222. We were very pleased to see the positive response of former referee #2 to our revisions, as well as the equally positive comments of the new referees #4 and #5.

We were also glad to see that the additional work requested involves either the improvement of existing data, or of their presentation, rather than any new critically important experiments. Some of these were indeed suggested, yet the referees are well aware of the difficulty -sometimes the impossibility- to realize them and thus clearly made them optional at best. We discuss some of these issues below. We nevertheless completed the paper with some new data involving yet another collaborator (HCR analyses) and hence Madeline Ryan now appears as an additional author. We modified and recomposed several figures and added both a Supplementary Figure (S6) and a Supplementary Table (S1).

We also tried our best to streamline the paper and remove some elements, which were not essential to support the main messages.

Referee comments *in extenso* are italicized and our responses are in plain text, for the sake of clarity.

Best wishes,

Denis (on behalf of all authors)

---

In our responses below, we use the new (revised) numbering of Figures and Supplementary Figures and Tables, without having modified them in the comments of the referees, thus inducing some possible mismatching here and there in Figures and Tables numbering between the comments and our responses.

---

**Referee #2:**

*I have carefully read the revised manuscript; it is much improved. The authors demonstrate a requirement for the 5' hoxd genes and the orthologous 5DOM regulatory region in zebrafish pseudocloacal and mouse urogenital development. They show that, although the specific cis-regulatory elements do not appear to be shared, the global cis-regulatory landscape appears to have been co-opted from an ancestral function in genital development to also control digit development in tetrapods. This helps to explain the deep evolutionary conservation of regulatory synteny in this cluster.*

*My earlier concerns about the communication of 'regulatory co-option' have been addressed by revisions to the text and with new data. Specifically, the end of the summary and the last few paragraphs of the introduction are much improved to more*

*specifically and accurately communicate what is meant by 'regulatory landscape co-option'. The explanation of the whole genome duplication and subsequent loss of hoxdb is also helpful to those who may be less familiar with the details needed to understand the overall premise of this paper.*

*The additional data on fish enhancers and their subtle deletion phenotype is very helpful to clarify that this paper doesn't imply specific enhancer co-option but rather explains the function and preservation of the 5DOM region as a whole. I also like the introductory paragraph in the section titled, "An ancestral regulatory landscape for an ancient function", because it better establishes the emphasis on the question of why regulatory synteny might be preserved for the 5DOM if it isn't required for fin development in fishes.*

*I encourage the authors to maintain the clarity of these points if they are required to shorten the manuscript.*

We thank this referee for his/her positive comments on the revised version and these clarifications will surely be kept in any subsequent versions of the paper.

*In the revision, I notice that replicate numbers are missing for most experiments and are surprisingly small for zebrafish experiments (e.g., wild type and the simpler mutant configurations in Figure 5). This should be addressed before acceptance.*

This is a valid point. We have now added a table in the supplementary material (new Table S1) showing the various numbers and we have increased the sample size for the simpler mutant configurations and wild type animals. Considering that neither single homozygous mutants, nor triple heterozygous mutants showed any phenotypes, the numbers are highly significant, with p-value < 0.001 (Fisher Exact Test) in the two types of comparisons.

Accordingly, we have re-numbered Supplementary Tables to maintain a correct text sequence in numbers (Table S1 to Table S4)

*Remaining minor points:*

*Line 214, 'allelic series' does not seem to be the appropriate language. Perhaps instead "these genetic configurations"*

This was changed as proposed

*In figure 3C, I would also illustrate the product of the inversions to make it explicitly clear how 5DOM is moved far from the gene cluster in the second but not in the first configuration.*

This is a good suggestion, for we understand that these inversion alleles are not easily visualized by the readers. We have now added simplified schemes, trying at the same time not to make the figure unnecessarily too complex.

## Referee #4:

*In this work, the authors study the conservation of the 5' and 3' regulatory landscapes of Hoxd clusters in vertebrates. It is well known that these genes have a bipartite regulation, with the anterior genes mostly regulated by the 3' regulatory domain and the posterior genes by the 5' domain, which leads the expression of these genes to the digits in tetrapods. They find that this 5' domain contains enhancers that drive Hoxd13 to the urogenital sinus in mice and its homologous structure, the cloacal region, in zebrafish. They conclude that the current transcriptional regulation of Hoxd genes in distal limbs (digits in tetrapods) has been coopted from the regulation at the cloaca in the vertebrate common ancestor.*

*The article is well-written, clear and easy to follow. The authors have used a good variety of approaches to support their claims, and the conclusions are in general original and well-funded in the results. In my opinion this will be a very interesting study for the scientific community, and thus I happily endorse it for publication in Nature.*

*However, I did find several details that should be corrected. The most important of them is related to the enhancers responsible of hoxd13a expression in the zebrafish fin. In their deletion assays, the authors don't find a big part of this regulatory information, and the mutant fish with no 5' domain still show hoxd13a expression in the fin. Although the authors claim that the expression is milder in these mutant animals than in wildtypes, which seems to be true according to the pictures (Fig. S5), the expression is very variable and other staining methods would be necessary to be able of measuring and comparing it, such as HCR. In any case, it is obvious that most of the enhancers responsible of this regulation are out of the deleted region.*

As this referee points out in his/her last sentence, it is obvious that *most of the enhancers (…) are out of the deleted region*. This is also supported by a critically important observation previously published by Ahn and Ho (2008), perhaps not enough visible in the current manuscript, i.e., the fact the *Lunapark* (*lnp*) gene, which is embedded

into 5DOM is NOT expressed in fin buds, while strongly expressed in mouse distal limb buds (Spitz et al., 2003). Another evidence for the absence of strong limb elements in the fish 5DOM landscape.

We nonetheless followed the expert's recommendation and carried out some HCR experiments on both control and 5DOM mutant fin buds. The results nicely fit the previous WISH dataset and allowed for some quantification to be carried out. This is now shown in a new Figure S6. Accordingly, the main text was completed as well as the methods section.

*According to the diagrams showed in Fig. 2a and S5a, the deleted region goes from the hox genes to the atp5mc3a/atf2 genes, not including them. However, in the HiC data showed in Fig. S2 (although it was not performed in fin tissue), we can observe that the TAD containing the 5' domain encompasses these two genes, and thus this region not deleted by the authors might contain part of this missing regulatory information. Furthermore, 4Cseq data reported in Acemel et al. 2016 (PMID: 26829752) in zebrafish support this hypothesis, showing strong contacts between hoxd13a promoter and the atp5mc3a/atf2 genes region. Although new deletion experiments would not change the overall results of the article, I think this issue should be considered by the authors and properly discussed in the manuscript.*

Yes, this is a good point indeed. Few elements can be considered to answer this issue: First this 'upstream region' does not contain any limb regulatory activity in mammals, since deletion of 5DOM totally removes all activity (Montavon et al., 2011). In this sense, even if the fish region would indeed contain (a) distal enhancer(s), the general argument on the evolution of the landscape would not be different, since the amniote distal enhancers would have emerged *de novo* at a different place (unless one would argue for a transposition of this region closer to the *Hox* cluster in amniotes, which is highly unlikely when considering the surrounding genomic structures and micro-syntenic relationships). Secondly, the strong contacts seen in HiC (Fig. S2) are clearly with the CTCF site located close to the *Atf2* promoter and this is most probably the main anchor point to be considered, rather than any TAD calling algorithms, which are useful but notoriously variable according to the fixed initial conditions. Finally, in Woltering et al. (2014), we showed that a pufferfish BAC including this region was NOT able to elicit a distal limb expression when used as a transgene.

***Other minor comments are:***

*L119: Fig. S2c should be Fig. S3c.*

Thanks. This was changed.

*L166: hoxda10a should be hoxd10a.*

Thanks. This was changed.

*L169: Apart from Fig. 2, Fig. S6b should also be called here. Indeed, Fig. 2 and Fig. 6 are a bit confusing, since very similar information is showed in both. For example, Fig. 2b and S6b are showing the same experiment at different zoom level. Young embryos showed in Fig. 2a would benefit of an inset with a zoom in as well. In general, I think these two figures could be better organized.*

Yes, this is a good option. We thus transferred former panel S6b into figure 2. The text (main and legend to figure) was changed accordingly.

*L172: Is the expression of the genes hoxd10a and hoxd4a at 72h showed somewhere?*

No, the latest stage where we have WISH for these genes in Del5DOM embryos is 48hpf for *hoxd10a* and 60hpf for *hoxd4a*. Pictures have now been added as a panel **e** in Fig. S7.

*L235, 236: I don't find that the ATAC peaks are "positioned at regions that are relatively depleted for the K3K27ac mark". Indeed, I see in Fig. 4a that they co-localize with an ATAC-enriched region. Besides, I don't see the point that the authors want to make, because this would mean that these are not active regions.*

Sorry for this misunderstanding, due to the lack of precision in the text (also raised by referee #5). This sentence refers to Fig. S10 and not to figure 4a. Here we are talking about the two binding sites (GT2 and Island E) and not about the global profiles shown in Fig. 4.

Within any accessible site in the chromatin, the ATAC signal will be 'central', because the cut DNA will be between the nucleosomes i.e., where the tagmentation will occur (where the transposome can access). In contrast, because nucleosomes will be displaced 'laterally' (due to the binding of a factor), the signals for histone

modifications (e.g. H3K27 acetylation) will be on either sides on the ATAC signal. This 'depletion' of H3K27 acetylation from the exact position of the ATAC peak is a hallmark of 'positive' chromatin as found e.g. in enhancer sequences. This can be seen for these two elements (Fig. S10), yet not for CSB. Of course. these details cannot be seen from the global profiles usually displayed as comparisons, such as that shown in Fig. 4, where the two signals appear superposed.

We have now clarified this issue by modifying the sentence. It now reads: *Within these two elements, the main ATAC peak is positioned at a region that is relatively depleted for the H3K27ac mark (Fig. S10), which is a hallmark of active enhancer elements*[36].

*L283-285: I do not agree with the authors about the number of peaks in the cloaca and pectoral fins cell types. I think the cloaca pseudobulk comes from a very small number of cells, and thus it shows a high background ratio, while the pectoral fins samples are more populated, and the signal/noise ratio is much better. I would trust in the peaks showed in fins, but I am not very sure about the cloaca, I think many of these 'peaks' are probably not real signal. In any case, even if I trust these data, I don't see "many more peaks" in the cloaca than in the fins.*

Yes, we agree with this comment. This is the difficulty to work with a structure of less than 40 cells detected in the datasets analyzed. We tone down this part of the description by removing the '*many more peaks*'. However, we maintain that every signal for which we have repeated coverage amongst these 38 cells is significant target to look at meta-analysis and support functional experiments of this work. Both the text and the figure S13 have been modified accordingly.

*L322: Apart from calling the Fig. 5b, f, S14b, c, e and f should also be called. Alternatively, I would remove this figure calling and call them together in the next line, Fig. 5b, f and Fig. S14.*

Yes, we agree and moved the reference to these figures to the line below.

*L342, 343: Again, the issue addressed in my first point.*

Please see our answer above on the additional use of HCR

*Materials and methods: Mouse ChIP-Seq section should be called just ChIP-Seq, since also zebrafish ChIP-Seq re-analysis is detailed here. Alternatively, this section could be split in two.*

The title was changed to 'ChIP-seq'

*L1197: bard should be bars.*

Thanks!

*Figure 6 is not called in the text. Section.*

Thanks for noticing this. This has now been introduced in the discussion

*Fig. S2 (and corresponding materials and methods): Is the CTCF ChIP-Seq signal shown in the bottom panel coming from the merge of 24 and 48hpf samples? This must be clarified.*

It is now indicated in the legend of figure S2 that the CTCF signal is from 48hpf larvae (GSM534493)

*Fig. S7a: male and female symbols should be added to these diagrams.*

Sorry for this oversight. The symbols have been added.

*Fig. S7c, d: the organization of these panels is confusing, taking into account how they are called in the text. I would rather put both diagrams together in panel c, and add new letters d and e for the panels with pictures.*

Yes, this is a better option. The figure was reorganized accordingly

*Fig. S11a: looking at these tracks with a high zoom level and autoscale makes the reader think that very different signals are comparable, like head or cloaca and tailbud mesoderm, when they are actually very different, as can be observed in Fig. S11. I think this panel should be corrected, either increasing the region shown (decreasing the zoom level) so we could have other*

*peaks as a reference, or deactivating the autoscale of the browser and fixing the scale in reasonable heights, like the ones used in Fig. S11.*

Yes, we agree (even though we think that this referee is concerned by Fig. S12, rather than S11, as mentioned at first. We have deactivated the autoscale option for the bulk ATAC-seq panels and removed the single cell ATAC-seq panels from a new Fig. S13a, for we agree that they were not substantial enough to draw firm conclusions.

## Referee #5:

*The manuscript by Hintermann, Bolt, Hawkins et al. investigates the regulatory evolution of the HoxD cluster in the context of fin/limb development. In mice, the HoxD cluster is flanked by two distinct regulatory domains, 3DOM and 5DOM, which control HoxD expression in the proximal (stylopod/zeugopod) and distal (autopod) fin/limb regions, respectively. Previous studies also have revealed several limb enhancers within these domains. Furthermore, the complete deletion of either the 3DOM or 5DOM domains abolishes HoxD gene expression in the corresponding limb spatial domains (proximal or distal, respectively).*

*In this study, the authors explore whether HoxD fin/limb regulation is conserved across vertebrates, by performing large deletions of the homologous domains in zebrafish. Deletion of the 3DOM results in effects similar to those observed in mice, characterized by the loss of the limb expression of hoxd10a and hoxd4a. This demonstrates that the enhancers controlling the expression of anterior hoxd genes (hoxd3a–hoxd10a) are located within the 3DOM, as in tetrapods, suggesting a deep evolutionary conservation for this regulatory domain.*

*Unexpectedly, the deletion of the 5DOM in zebrafish does not affect the expression of posterior hoxd genes, unlike in mice. This is evidenced by the expression pattern of hoxd13a, which remains largely unaffected, only showing minor defects at later stages. These results indicate functional differences in HoxD posterior limb regulation between vertebrate species, suggesting that in zebrafish the regulatory potential may be embedded within the HoxDa cluster itself.*

*Since the 3DOM and 5DOM also control HoxD expression in other developing tissues, the authors investigated expression patterns by whole-mount in situ hybridization (WISH) in embryos. Notably, these analyses revealed that Hoxd13 expression is completely abolished in the pseudocloacal region of zebrafish mutants carrying a 5DOM deletion. This prompted the authors to explore whether the 5DOM also regulates HoxD gene expression in the mouse urogenital sinus (UGS), a structure homologous to the zebrafish cloaca. Through the analysis of various mouse mutant alleles, they confirm that the 5DOM indeed drives UGS expression. Epigenetic experiments helped to pinpoint specific enhancers within this domain, which were validated through LacZ reporter assays.*

*To assess evolutionary conservation, the authors performed multispecies alignments and compare relevant ATAC-seq datasets, identifying putative enhancers in zebrafish and mice. However, these analyses reveal limited conservation between species. One conserved enhancer, CsB, previously shown to have regulatory activity, was deleted in zebrafish, revealing its limited functionality during fin or pseudocloacal development. Finally, the authors demonstrate that hox13 genes are essential for cloacal development by analyzing the phenotypes of combined deletions of hoxd13a, hoxa13a, and hoxa13b.*

*Overall, through a series of elegant genetic experiments, the authors establish that the 5DOM is necessary for controlling Hoxd13 expression in the cloaca/UGS, suggesting that this function is deeply conserved during vertebrate evolution. Additionally, they demonstrate that posterior HoxD gene regulation in the fin/limbs differs between zebrafish and mice, with 5DOM sustaining expression only in the latter. Based on these findings, they propose that the ancestral function of the 5DOM was to regulate HoxD expression in the cloaca, and that it was later co-opted during evolution to support digit expression.*

*This is not a "typical" case of co-option, where a gene or particular enhancers may have evolved to support a new function. Instead, the authors argue that the regulatory landscape has evolved as a whole, which is an important concept for the field.*

We thank this expert for these positive words and for pointing to the 'atypical' nature of the co-option, which is what we would like to propose in this work.

*However, the situation is complex, as the zebrafish 5DOM also contains distal regulatory potential in fins (as reported in Gehrke et al., PNAS 2015 and Schneider et al PNAS 2011), but that are mostly functionally dispensable. The authors argue that zebrafish may represents a partial step in the full co-option that is observed in tetrapods. While this is a compelling hypothesis, the evidence remains still preliminary.*

We of course agree with this statement and we would even say that we can hardly see how to better demonstrate this hypothesis in a close future. Please note that this hypothesis is only one of the alternative discussed and should of course be considered as such. Also note that the CsB region was never shown to be fully distal in mouse transgenes. The region embedded into the larger GCR segment had a distal specificity, but the region alone did NOT work in the most distal part (the future digits) (Gonzalez et al., 2007). The transgenes reported in Montavon et al. (2011) to be superstrong in the distal part (in particular 'Island II') had no orthologs in zebrafish.

*The study only involves the functional characterization of two species, making it difficult to infer whether the zebrafish condition represents an ancestral state or a derived feature (this interpretation is also considered in the discussion). In this section, the authors also argue that their main conclusion is supported by the existence of Hoxd13 cloacal expression in lampreys, a species that lack fins (digits). While this implies that Hoxd13 cloacal expression likely predates fin/limb expression, it does not confirm that cloacal regulation was initially sustained by the existence of distal elements within a 5' flanking regulatory domain.*

We also agree with this, but this is not exactly what we are arguing in the discussion. What we propose is that 'expression' in the cloaca predated expression in any other structures. Then, because of this ancient expression, the associated regulatory landscape (TAD) was 'prone' to more rapidly evolve other regulations that anywhere else in the genome, due to 1) the high 'functional potential' of recruiting *Hox* genes in another emerging structure (due to their demonstrated versatility) and 2) the presence of a 'regulatory ground-state' (factors, co-factors, chromatin structures etc…) that made co-option of the landscape facilitated. This mechanism of 'facilitation' (a sort of 'pioneer enhancer sequence', rather than a pioneer factor) was proposed by Gonzalez et al. (2007, 'enhancer priming') and discussed in a few essays (see Darbellay and Duboule, 2016).

*Additional evidence may be required to support this conclusion, which is central to the manuscript and in it is reflected in the title.*

Yes, certainly, but please note that this is not a 'conclusion'. Instead, it is the main 'proposal' of the paper, as is often the case when dealing with the evolution of the regulatory genome.

*Given the technical challenges of performing large deletions in other relevant species (already performing those in zebrafish is an impressive achievement), alternative approaches could provide further support for the authors claims. For example, investigating whether Hoxd13 cloacal expression in lampreys is controlled by enhancers within a 5DOM-like structure could be informative*

Yes, we agree that this experiment would be fantastic and likely decisive for our argument. However, it is simply not feasible with current technology, similar to large deletions, as acknowledged by this expert. To get a first approximation, a multiomic approach would be necessary (to try and identify ATAC peaks in those cells expressing an elusive *hox13* gene and 'identified' as cloacal cells.. (see below). The results obtained with the zebrafish (38 cells characterized as cloacal cells) and the valid modifications asked by the referees (on peak specificity) show well the impossibility to run something similar with a non-model species, where even genome alignment would not be guaranteed. On the other hand, a bulk ATAC-seq approach would hardly be informative due to the very high dilution of the specific signal, if any.

Furthermore, it is quite clear that genome amplifications occurred after the separation between agnathans and gnathostomes, which makes a correspondence with the mouse and fish clusters elusive. This would represent the work of a laboratory for many years, with several technological developments.

*Identifying a similar 5' flanking regulatory domain in cyclostomes, which lack fin/limbs, would support the hypothesis that cloacal distal regulation is ancestral. These experiments could be complex to interpret, due to the genome triplication of cyclostomes, but may result very informative. Experiments such as Hi-C and ATAC-seq in these species could help to address this question*.

There again, we fully agree with the idea. But how to extract or deconvolute a HiC profile found in 38 cells when running an experiment with a few million cells? There again, same problem as in zebrafish, even though it is a model system. We do not show any HiC profile for cloacal cells in zebrafish and nobody would be able to do so nowadays, unless 100 thousand fish would be cell suspended to sort out cloacal cells in a way or another. Now let's assume that despite all these difficulties, some contacts point would be seen in HiC, even matching ATAC-seq peak in few cloacal cells, how to demonstrate their specificities in lamprey?

*An interesting observation is that ATAC-seq comparisons between zebrafish fins (Figure S11) and mouse distal limbs (Bolt et al., Nat. Commun., 2022) reveal significantly more peaks in the latter. This may serve as an indirect indicator of regulatory potential. Performing similar analyses in other species, such as cartilaginous fishes or the spotted gar, may serve to strengthen the authors claims. If ATAC-seq profiles from these species also show a reduced number of fin-specific peaks in the 5DOM, similar to zebrafish, this would support the hypothesis that distal limb regulation has evolved later in tetrapods.*

In response to the comment of referee #4 (lines 283-285; see above), we have now tone down the importance of this particular comparison (which was not a major point in any case). Again, the main value of this work is the functional approach, which is impossible to achieve in the species mentioned by this expert.

ATAC profiles can be useful to back up a functional approach but cannot be used by themselves to conclude on regulatory similarities between different species. Eventually, and in particular if ATAC profiles would be somewhat similar, we would need to delete 5DOM in these species, which is strictly impossible for now. We believe that to 'corroborate' the conclusions by adding non-controllable evidences from non-model species is not desirable from an epistemological viewpoint.

*The results in the manuscript also suggest that distal HoxD expression in zebrafish is driven by regulatory elements within the hoxda cluster itself. This contrast with the mouse, where regulation is external to the cluster and located in the 5DOM (as shown in previous studies from the authors). This raises the possibility that posterior HoxD limb regulation was ancestrally located within the cluster and later relocated to the 5DOM in tetrapods. Alternatively, distal regulation may have been the ancestral state, with zebrafish having evolved a more proximal regulatory mechanism. For instance, do specific peaks within the zebrafish hoxda cluster support posterior gene expression?*

This is a good point. However, it is very difficult to unambiguously address this issue, since *Hox* clusters (in vertebrates) are very dense in genes, promoters, regulatory sequences and are furthermore covered with large transcribed regions producing read-through transcripts, sometimes on both DNA strands. To isolate these enhancers would be very tricky (we tried doing this in the mouse for several years) and again, the main point of the paper is that zebrafish enhancers are (for the most) NOT in C-DOM, rather than determining where they are. This referee is right in his/her evolutionary statement, and a possibility exists that the regulation in ancestral fishes was positioned into 5DOM and was somewhat 'translocated' into the cluster itself. However, we would have anticipated this to be somehow associated with a reshuffling of this landscape, which is clearly not the case when considering the syntenic relationship (presence of *Evx2* and *Lnp*, distance, general topology…). We thus consider this possibility as unlikely.

*Are these regulatory regions functionally conserved outside tetrapods? Investigating other species may also help to clarify this issue. Are BACs from other relevant species (i.e. spotted gar, sharks, skates…) containing the HoxD cluster and that could be used to investigate the regulation within the cluster?*

This is again a fair possibility. We did use some BACs in the past to investigate this issue by using the pufferfish genome, because the entire 5DOM and the cluster would be present on one BAC (Woltering et al.). This approach however is made extremely complex by several factors. First, BAC transgenesis is not possible in non-model organisms (in fact we are not aware of successful achievements even in zebrafish). This means that the BACs would need to be introduced into transgenic mice, which raises many issues. For example, copy number must be controlled as well as the integrity of the integrated material, calling for a HR event rather than an 'old fashioned' injection experiment. Few reports have addressed this technological advance and its feasibility is still discussable. In addition, mice have loads of endogenous HOX proteins in their distal limb buds and any regulatory sequence containing a HOX binding site (which are highly degenerated…) may respond to these endogenous proteins, thus producing an artefactual distal signal.

The main added value of this paper is the genetic approach, in both mice and fish, the latter being a premiere in term of addressing the large-scale regulation of a locus in any fishes thus far. We would feel somewhat uneasy to go back to somewhat more 'usual' evo-devo approaches, even though we agree that they might be informative, to some extent.

***Minor comments:***

*Lines 235-236: "The ATAC peaks were positioned at regions that are relatively depleted for the H3K27ac mark (Fig. S9)". I am not sure of what the authors mean with this statement. The ATAC-seq peaks, in particular Island E and CsB seem to be at positions that display abundance of H3K27 marks.*

Apologies for being unclear on this part. The same comment was also raised by referee 4 and hence please see our response above. We believe we have now clarified this issue.

*Figure 3d: Please indicate the Y axis. Are the expression data derived from the experiments presented in Figure s8?*

The Y axes are RT-qPCR, which is now indicated both in the new Fig. 3 and in its legend. This experiment was different from that shown in Fig. S9, which is an RNA-seq dataset.

*Lines. 358-360: "Alternatively, it may reflect a secondary loss of several distal enhancers occurring in aftermath of the teleost whole genome duplication as such event may lead to the reorganization of paralogous regulatory landscapes"*

*Although enhancer loss is a recurrent phenomenon after WGD, the alteration of regulatory activity in zebrafish does not need to be necessarily associated with this evolutionary event. In fact, the alignment analyses performed by the authors actually point to a limited impact of WGD in the 5DOM. Thus, the secondary loss of regulatory elements could have also happened in the Actinopterygian lineage, independently of the WGD.*

Yes, this is a good point and it is indeed possible that loss of enhancers (if any) happened before the teleost genome amplification. Even though we do not favor this hypothesis (see the discussion), we now mention this possibility in a revised sentence that reads: `*Alternatively, it may reflect a secondary loss of distal enhancers either in the Actinopterygian lineage, or in aftermath of the teleost whole genome duplication as such event may lead to the reorganization of paralogous regulatory landscapes. Our phylogenetic analysis of sequence conservation within 5DOM however shows that the latter option is unlikely for this specific locus, which does not maintain any paralogous copy in zebrafish*`.