***Cis*-regulatory landscapes in development and evolution**

Ignacio Maeso^{1*}, Rafael D. Acemel¹, José Luis Gómez-Skarmeta^{1*}

¹Centro Andaluz de Biología del Desarrollo (CABD), Consejo Superior de Investigaciones Científicas/Universidad Pablo de Olavide, Carretera de Utrera Km1, Seville, Spain.

* Corresponding authors: nacho.maeso@gmail.com, jlgomska@upo.es

Abstract

The recent advances in our understanding of the 3D organization of the chromatin together with an almost unlimited ability to detect *cis*-regulatory elements genome-wide using different biochemical signatures has provided us with an unprecedented power to study gene regulation. It is now possible to profile the complete regulatory apparatus controlling the spatio-temporal expression of any given gene, the so-called gene Regulatory Landscapes (RLs). Here we review several studies over the last two years demonstrating the functional consequences of altering RL structure in development, disease and evolution. These works clearly shows that a deep understanding of transcriptional regulation is no longer conceivable without considering the 3D modular organization of animal genomes.

Introduction

Transcriptional regulation plays a major role in almost every biological process from embryogenesis to tissue homeostasis and disease. In agreement with this, *cis*-regulatory elements (CREs) occupy a much larger fraction of the genome than the protein coding genes. Until recently, due to the lack of direct empirical data, it was difficult to study how genetic information was distributed within the large amount of DNA that characterizes animal genomes. Not much could be done beyond arbitrarily subdividing the genome into ad-hoc categories: genic and intergenic, intronic and exonic, upstream and downstream. However, these categories are a poor reflection of the physical and functional compartmentalization of the chromatin within the cell nucleus and are very far from fully capturing the complexity of *cis*-regulatory interactions. Many works in the last decades have shown that regulatory DNA can act at long distances, often more

than 1 Mb, by contacting the promoters of target genes through chromatin loops (reviewed in [1]). Furthermore, these enhancer-promoter physical interactions occur in a highly regulated manner that precludes spurious activation of non-target neighboring genes. We have now a first glimpse of how animal genomes are folded thanks to the use of several chromosome conformation capture techniques that identify DNA-DNA contacts throughout the genome (reviewed in [2]). Over the last years, the analyses of all possible contacts between a single genomic region and the rest of the genome at high resolution (4C-seq), or between all chromatin regions at low resolution (Hi-C), have shown that chromatin is compartmentalized into structures known as topologically associating domains (TADs): megabase-scale genomic regions in which sequences preferentially contact one another [3-6]. TADs are largely invariant between different cell types and developmental stages [5,7,8] and, effectively, can be considered as the basic structural units of the genome: separating in the 3D space genomic regions that are immediately adjacent from a linear perspective but are isolated in terms of function and expression while bringing together interacting sequences that are hundreds of kilobases apart but belong to the same TAD (reviewed in [9-11]). TADs can be further subdivided from the perspective of each of their constituent genes. By profiling the contacts of a given gene promoter, 4C-seq technology allows the identification of the genomic territory within a TAD containing all the CREs that regulate that promoter, what is defined as the regulatory landscape (RL) of that particular gene [12]. Indeed, functional studies in mouse embryos using a transposon-associated sensor [13] have shown that these contact profiles correspond to large regulatory domains typically found around developmental genes [14-16]. This improved knowledge of the genomic organization in animals, much more realistic and biologically informed, poses new questions to our current understanding of transcriptional regulation. How are structural changes in these 3D genomic modules, TADs and RLs, impacting gene regulation? How are RLs modified during evolution? Several studies over the last two years are starting to pave the way for answering these questions.

RL restructuring in disease and development

Recent studies have shown how mutations affecting the structure and integrity of RLs can lead to spurious enhancer-promoter contacts (Figure 1). And as a consequence of these ectopic regulatory interactions, critical genes associated to human diseases may be aberrantly expressed. In fact, in the genomes of tumor cells in Tcell acute lymphoblastic

leukemia, many prominent proto-oncogenes are activated due to deletions eliminating boundary elements in their vicinity [1700]. In genes such as *TAL1* and *LMO2*, the absence of their surrounding boundaries cause large-scale rearrangements of their chromatin interactions, allowing these genes to contact CREs that would normally lie outside their RLs [1700]. Furthermore, disruption of boundary elements seems to be a general signature of many other cancer types, since somatic mutations found in whole genome sequences of cancer samples are strongly enriched at CTCF sites [1700,18]. Importantly, these mutated CTCF sites correspond to boundary regions since no enrichment is found in non-boundary CTCF sites [1700].

Disruption of TADs can also result in pathogenic morphological alterations caused by rewiring of the long-range regulatory architecture of key developmental genes. In the chromosomal region around *Epha4* genes, large-scale genomic rearrangements, including duplications, inversions and macrodeletions have been shown to cause several limb malformations in different patients and in mouse models engineered to mimic these mutations [1900]. In mouse, *Epha4* is expressed in the distal part of developing limb buds and this expression is driven by a cluster of long-range enhancer elements located within the *Epha4* TAD gene desert, at ~1.7 Mb from its promoter. Three other prominent developmental regulators are found in the same region, *Wnt6*, *Ihh* and *Pax3*, but they are located in different TADs, and their RLs are isolated from that of *Epha4*. Furthermore, these genes are either not expressed in mouse limb buds or their expression patterns are very different from that of *Epha4*. But in the different human and mouse chromosomal mutations mentioned above, the integrity of the TADs and RLs of these developmental genes and the relative locations of their boundaries are affected [1900]. This results in changes in the 3D contacts of either *Wnt6*, *Ihh* or *Pax3*, that now are able to interact with the *Epha4* limb bud enhancers, causing their misexpression and altering normal limb development [1900].

A similar situation is found in duplications of the *KCNJ2-SOX9* genomic region, which are associated to the Cook syndrome [20]. A study using a mouse model reproducing this duplication shows that this rearrangement generates a new TAD that expose the *Kcnj2* gene to the *Sox9* RL, causing the overexpression of *Kcnj2* in developing limbs in *Sox9* expressing territories and, as a consequence, limb malformations that phenocopy those observed in Cooks syndrome patients [2100].

RL modifications in evolution

By demonstrating that reshuffling TAD architecture can produce expression changes that cause limb morphological changes, even though aberrant [19·, 21·], these recent studies provide a suggestive link with the appearance of phenotypic variation during evolution. Related to this, it has been suggested that the evolutionary recruitment of hedgehog signaling during early limb development in jawed vertebrates could have occurred by a genomic rearrangement in the RL of *Shh* [22]. This would have allowed the interaction of the *Shh* promoter with its long-range limb bud enhancer, located in an intron of the neighboring gene *Lmbr1*, an evolutionary scenario reminiscent to the gain of ectopic limb bud expression by the previously mentioned hedgehog paralogue, *Ihh*, upon mutations altering its RL topology [19·].

But are RL rearrangements actually playing an important role in evolution? A recent comparison of Hi-C data across four mammalian species has shown that TADs are remarkably stable, with most evolutionary genomic rearrangements located at positions that do not compromise the integrity of individual TADs and their borders, at least since the last common ancestor of these placental mammals [23·]. This would be in agreement with recent evolutionary models that propose that genomic rearrangements would preferentially occur at certain genomic positions [24]. Indeed, there is at least one case, in the genomic region encompassing the Six genes clusters, in which the evolutionary conservation of TAD boundaries is even deeper, dating back to the origin of deuterostomes during the Cambrian [25·]. However, this is just an isolated example and inferring global trends about the paucity and the rate at which TADs are modified during evolution would be premature till more comparative chromatin interaction data across different animal phyla becomes available. In fact, works focusing on the 3D architecture of some of the best-studied genomic regions, the Hox clusters, have revealed the important contribution of RL changes to animal evolution.

Hox RLs and the evolution of vertebrate appendages

In tetrapods, HoxA and HoxD genes play a critical role during the development of proximal and distal limb structures [26-28]. These two Hox clusters are flanked 5' and 3' by extensive neighboring genomic regions containing multiple distal CREs that precisely regulate their expression in the developing limb territories [12,29]. Structurally, the 5' and 3' neighboring regions are organized into two adjacent TADs, with the Hox genes lying precisely between these two interaction compartments. This bipartite 3D chromatin configuration allows the interaction of two independent sets of

CREs with their target genes (reviewed in [30]). Thus, at the early phase of limb development, CREs within the 3' TAD of the HoxD cluster are active and control the expression of 3' Hox genes, which are involved in the patterning of proximal structures (i.e. the arm). In the second phase, the CREs of the 5' TAD become active to promote the expression of 5' Hox genes in distal limb bud regions, which will give rise to autopod structures such as the hands and the digits [12,29,31]. As development progresses, genes in the central part of the HoxD cluster (*Hoxd8-Hoxd11*) transit from the 3' to the 5' TAD changing their RL and their expression from proximal to distal limb territories [31]. This temporal regulatory switch is likely important to generate a limb territory with low levels of Hox expression in which the wrist and ankle are formed [31]. How this global temporal regulation switch is controlled? HOX13 proteins seem to play a critical role in this process. A recent study show that these proteins activate a global repressive mechanism in the 3' TAD while at the same time bind CREs in the 5' TAD likely promoting their activity and generating a positive feed-back loop [32•]. Interestingly, part of the RL of the 5' TAD required for limb development is also used for the development of the external genitalia, and the two structures make use of some shared and some specific enhancers, all located within the same topological 3D structure [33•]. These results suggest that this type of 3D regulatory architecture is a fertile environment for generating functional diversity of developmental loci by hijacking the regulatory machinery of certain enhancers from a certain cell type to another one [33•]. In fact, attesting to the plasticity of this 3D topology, recent studies have shown that different jawed vertebrate lineages, with very different appendage morphologies including limbless species such as snakes, have maintained this bipartite Hox architecture, and that their differences in Hox gene expression would be related to changes in the number and disposition of enhancers within the conserved Hox TADs [34•,35-37•,38].

When did this topology appear during evolution? Recent insights come from the study of the 3D organization of the cephalochordate amphioxus [39•], which contain a single Hox cluster. In contrast to vertebrates, the only amphioxus Hox cluster is organized into a single TAD that includes long-range contacts mostly from the 3' genomic region flanking the cluster. Therefore, these results suggest that the vertebrate Hox bipartite regulatory system is an evolutionary novelty that emerged by combining ancient long-range regulatory information from the anterior Hox TAD with newly incorporated

regulatory inputs at the posterior side of the cluster, which probably arose by some kind of genomic rearrangements that allowed novel interactions with a new 5' neighboring genomic region at the base of the vertebrate lineage. Interestingly, most of the regulatory elements located in this vertebrate-specific posterior TAD are required for the formation of vertebrate novelties such as the digits and external genitalia [12,33].

Functional dissection of RLs

The examples commented in the three previous sections highlight the modular nature of TADs, which can be affected by different mutational events merging, splitting or combining them (Figure 1). But how would be the effect of these RL rearranging mutations in comparison with changes affecting individual CREs? Is the evolution of each CRE affected by the rest of CREs located within the same RL?

The extent of functional crosstalk between CREs belonging to the same RL has been recently evaluated by two reports [40,41] trying to solve the super-enhancer conundrum [42]. Super-enhancers are putative distinct higher-order regulatory entities that were recently defined as enhancer-like elements characterized by their binding to the Mediator complex and other master regulators and by their large size, typically consisting of a cluster of several CREs spanning up to 12.5 kb [43-45]. However, it is a matter of debate if they really represent conceptually and functionally distinct entities from traditionally defined enhancers and it has been argued that they could be simply regarded as clusters of strong enhancers [40,42]. The expression of α -globin genes in mouse erythroid cells is controlled by a cluster of five CREs that fulfill all the defining properties of super-enhancers [40]. However, when each of its constituent CREs is deleted within their native genomic context, either alone or in different combinations, the activity, CRE biochemical signatures and long-range interactions of the remaining elements within the α -globin super-enhancer are not affected [40]. Furthermore, the enhancer activity of each of these elements seems to be redundant and in some cases, their contribution to gene expression is negligible. Thus, at least in this particular example, the clustering of several CREs within the same RL do not seem to generate any kind of singular composite or emerging regulatory functions. Instead, it may simply reflect that the presence of strong enhancers could perhaps favor the evolution of other nearby enhancers by generating a more permissive chromatin environment in its surrounding regions [40,46]. The functional dissection of the super-enhancer activating the expression of the *Wap* gene in mammary glands during pregnancy shows a

completely different picture [41·]. In this case, full induction of *Wap* transcription could only be achieved when all of the three constituent CREs of the super-enhancer were intact and the mutagenesis of one of these elements completely abolished the activity of the other two, incapacitating the entire super-enhancer. Although the α -globin and *Wap* examples may seem contradictory, they are not necessarily incompatible. Different sets of enhancers in different RLs can probably exhibit disparate behaviors, from high levels of redundancy and almost no cooperativity to hierarchically organized CREs with a high degree of interdependency. If the super-enhancer concept and its defining molecular properties are any good at diagnosing these different regulatory modalities is still questionable, and it is likely that new examples of RLs containing functionally interconnected CREs but without displaying super-enhancer features will be described at some point in the near future.

Conclusion and future perspectives

We have now a much more precise and realistic picture of how regulatory information is organized in the genome, from the specific molecular and chromatin signatures that characterize CREs to the way these elements are physically and functionally compartmentalized in the 3D space inside the cell nucleus. However, there is still work to be done before formulating general principles about how RLs and their constituent CREs are spatially and functionally assembled and coordinated, and how they are modified during evolution. In this regard, the generation of chromatin interaction data from a much broader taxonomic sampling will be necessary in combination with careful examination and detailed functional dissection of many more examples of RLs. The use of novel genomic editing tools, such as the CRISPR/CAS9 system, will undoubtedly contribute to this endeavor, as highlighted by the increasing number of works that have benefited from these approaches to shed light into the molecular mechanisms underlying TAD boundaries and the establishment of chromatin loops [19,47-49].

Acknowledgements

We apologize to those researchers whose work could not be discussed due to space limitations. This work was funded by grants from Ministerio de Economía y Competitividad (BFU2013-41322-P and BFU2014-55738-REDT) and the Andalusian Government (BIO-396).

Figure Legends

Figure 1. Mutations generating genomic rearrangements such as (a) inversions, (b) deletions and (c) formation of new boundary regions, alter the interactions and relative positions of TAD borders, changing the structure of RLs and gene expression. This can lead to oncogene activation in tumor cells, morphological alterations during development and regulatory novelties in animal evolution.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. de Laat W, Duboule D: **Topology of mammalian developmental enhancers and their regulatory landscapes.** *Nature* 2013, **502**:499-506.
2. Denker A, de Laat W: **The second decade of 3C technologies: detailed insights into nuclear organization.** *Genes Dev* 2016, **30**:1357-1382.
3. Hou C, Li L, Qin ZS, Corces VG: **Gene density, transcription, and insulators contribute to the partition of the Drosophila genome into physical domains.** *Mol Cell* 2012, **48**:471-484.
4. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J, et al.: **Spatial partitioning of the regulatory landscape of the X-inactivation centre.** *Nature* 2012, **485**:381-385.
5. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B: **Topological domains in mammalian genomes identified by analysis of chromatin interactions.** *Nature* 2012, **485**:376-380.
6. Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Parrinello H, Tanay A, Cavalli G: **Three-dimensional folding and functional organization principles of the Drosophila genome.** *Cell* 2012, **148**:458-472.
7. Ghavi-Helm Y, Klein FA, Pakozdi T, Ciglar L, Noordermeer D, Huber W, Furlong EE: **Enhancer loops appear stable during development and are associated with paused polymerase.** *Nature* 2014, **512**:96-100.
8. Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, et al.: **A 3D Map of the Human**

Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* 2014, **159**:1665-1680.

9. Dixon JR, Gorkin DU, Ren B: **Chromatin Domains: The Unit of Chromosome Organization.** *Mol Cell* 2016, **62**:668-680.
10. Nora EP, Dekker J, Heard E: **Segmental folding of chromosomes: a basis for structural and regulatory chromosomal neighborhoods?** *Bioessays* 2013, **35**:818-828.
11. Schwarzer W, Spitz F: **The architecture of gene expression: integrating dispersed cis-regulatory modules into coherent regulatory domains.** *Curr Opin Genet Dev* 2014, **27C**:74-82.
12. Montavon T, Soshnikova N, Mascrez B, Joye E, Thevenet L, Splinter E, de Laat W, Spitz F, Duboule D: **A regulatory archipelago controls Hox genes transcription in digits.** *Cell* 2011, **147**:1132-1145.
13. Ruf S, Symmons O, Uslu VV, Dolle D, Hot C, Ettwiller L, Spitz F: **Large-scale analysis of the regulatory architecture of the mouse genome with a transposon-associated sensor.** *Nat Genet* 2011, **43**:379-386.
14. Uslu VV, Petretich M, Ruf S, Langenfeld K, Fonseca NA, Marioni JC, Spitz F: **Long-range enhancers regulating Myc expression are required for normal facial morphogenesis.** *Nat Genet* 2014, **46**:753-758.
15. Symmons O, Uslu VV, Tsujimura T, Ruf S, Nassari S, Schwarzer W, Ettwiller L, Spitz F: **Functional and topological characteristics of mammalian regulatory domains.** *Genome Res* 2014, **24**:390-400.
16. Tsujimura T, Klein FA, Langenfeld K, Glaser J, Huber W, Spitz F: **A discrete transition zone organizes the topological and regulatory autonomy of the adjacent tfap2c and bmp7 genes.** *PLoS Genet* 2015, **11**:e1004897.

This publication is one of the first examples showing that altering the relative position of topological borders and regulatory elements with respect to their target genes can produce changes in gene expression. The authors study two adjacent RLs corresponding to two developmental genes, *Tfap2c* and *Bmp7*, separated by a weak boundary region. Inversions including this region and a set of *Bmp7* heart enhancers cause dramatic upregulation of *Tfap2c* in the developing heart and a drastic drop of *Bmp7* expression in this tissue.

17. Hnisz D, Weintraub AS, Day DS, Valton AL, Bak RO, Li CH, Goldmann J, Lajoie BR, Fan ZP, Sigova AA, et al.: **Activation of proto-oncogenes by disruption of chromosome neighborhoods.** *Science* 2016, **351**:1454-1458.

Removal of TAD boundaries causes aberrant overexpression of proto-oncogenes through their ectopic interactions with CREs. 26% of the recurrent small mutations in T-cell acute lymphoblastic leukemia overlap with boundary elements.

18. Katainen R, Dave K, Pitkanen E, Palin K, Kivioja T, Valimäki N, Gylfe AE, Ristolainen H, Hanninen UA, Cajuso T, et al.: **CTCF/cohesin-binding sites are frequently mutated in cancer.** *Nat Genet* 2015, **47**:818-821.

19. Lupianez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, Horn D, Kayserili H, Opitz JM, Laxova R, et al.: **Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions.** *Cell* 2015, **161**:1012-1025.

An elegant study showing the functional and morphological consequences of mutations rearranging the relative topology of TAD boundaries. The authors reproduce in mouse models large-scale deletions, inversions and duplications found in human patients that cause ectopic regulatory interactions and aberrant gene expression in the RLs of three important developmental genes, *Wnt6*, *Ihh* and *Pax3*.

20. Kurth I, Klopocki E, Stricker S, van Oosterwijk J, Vanek S, Altmann J, Santos HG, van Harssel JJ, de Ravel T, Wilkie AO, et al.: **Duplications of noncoding elements 5' of SOX9 are associated with brachydactyly-anonychia.** *Nat Genet* 2009, **41**:862-863.

21. Franke M, Ibrahim DM, Andrey G, Schwarzer W, Heinrich V, Schopflin R, Kraft K, Kempfer R, Jerkovic I, Chan WL, et al.: **Formation of new chromatin domains determines pathogenicity of genomic duplications.** *Nature* 2016.

The authors show that mutations causing large-scale duplications associated to human syndromes can result in the formation of new ectopic TADs (neo-TADs). These neo-TADs can encompass regions that would correspond to two different TADs in non-mutants. Mouse models of a duplication including the *Kcnj2* gene show the formation of a neo-TAD that allow the aberrant activation of *Kcnj2* in *Sox9* expression domains, causing limb malformations resembling the Cooks syndrome.

22. Irimia M, Royo JL, Burguera D, Maeso I, Gomez-Skarmeta JL, Garcia-Fernandez J: **Comparative genomics of the Hedgehog loci in chordates and the origins of Shh regulatory novelties.** *Scientific reports* 2012, **2**:433.

23. Vietri Rudan M, Barrington C, Henderson S, Ernst C, Odom DT, Tanay A, Hadjur S: **Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture.** *Cell Rep* 2015, **10**:1297-1309.

By comparing Hi-C data from four mammal species (macaque, mouse, rabbit and dog) the authors show that TADs are highly conserved and are kept as intact units by genomic rearrangements occurred during mammalian evolution.

24. Farre M, Robinson TJ, Ruiz-Herrera A: **An Integrative Breakage Model of genome architecture, reshuffling and evolution: The Integrative Breakage Model of genome evolution, a novel multidisciplinary hypothesis for the study of genome plasticity.** *Bioessays* 2015, **37**:479-488.

25. Gomez-Marin C, Tena JJ, Acemel RD, Lopez-Mayorga M, Naranjo S, de la Calle-Mustienes E, Maeso I, Beccari L, Aneas I, Viélas E, et al.: **Evolutionary comparison reveals that diverging CTCF sites are signatures of ancestral topological associating domains borders.** *Proc Natl Acad Sci U S A* 2015, **112**:7542-7547.

This study describes the most ancient TAD boundary identified to date, separating the RLs of Six genes since the last common ancestor of deuterostome phyla. Furthermore the authors identify the presence of convergently oriented CTCF-binding sites as a conserved signature of TAD borders.

26. Davis AP, Witte DP, Hsieh-Li HM, Potter SS, Capecchi MR: **Absence of radius and ulna in mice lacking *hoxa-11* and *hoxd-11*.** *Nature* 1995, **375**:791-795.
27. Fromental-Ramain C, Warot X, Messadecq N, LeMeur M, Dolle P, Chambon P: **Hoxa-13 and Hoxd-13 play a crucial role in the patterning of the limb autopod.** *Development* 1996, **122**:2997-3011.
28. Kmita M, Tarchini B, Zakany J, Logan M, Tabin CJ, Duboule D: **Early developmental arrest of mammalian limbs lacking HoxA/HoxD gene function.** *Nature* 2005, **435**:1113-1116.
29. Berlivet S, Paquette D, Dumouchel A, Langlais D, Dostie J, Kmita M: **Clustering of tissue-specific sub-TADs accompanies the regulation of HoxA genes in developing limbs.** *PLoS Genet* 2013, **9**:e1004018.
30. Lonfat N, Duboule D: **Structure, function and evolution of topologically associating domains (TADs) at HOX loci.** *FEBS Lett* 2015.
31. Andrey G, Montavon T, Mascres B, Gonzalez F, Noordermeer D, Leleu M, Trono D, Spitz F, Duboule D: **A switch between topological domains underlies HoxD genes collinearity in mouse limbs.** *Science* 2013, **340**:1234167.
32. Beccari L, Yakushiji-Kaminatsui N, Woltering JM, Necselea A, Lonfat N, Rodriguez-Carballo E, Mascres B, Yamamoto S, Kuroiwa A, Duboule D: **A role for HOX13 proteins in the regulatory switch between TADs at the HoxD locus.** *Genes Dev* 2016, **30**:1172-1186.

This study shows that HOX13 proteins themselves contribute to the switch of regulatory TADs that takes place in HoxD genes during limb development. HOX13 proteins globally repress the 3' TAD and at the same time, they activate the 5' TAD by interacting directly with 5' distal enhancers. In the absence of HOX13 proteins, the TAD switching does not occur and the proximal limb structure grows without wrist articulation, reminiscent of an ancestral fish-like situation.

33. Lonfat N, Montavon T, Darbellay F, Gitto S, Duboule D: **Convergent evolution of complex regulatory landscapes and pleiotropy at Hox loci.** *Science* 2014, **346**:1004-1006.

The same 3D regulatory topology of the HoxD cluster and a common set of enhancers are used in two different developmental contexts: the limbs and the external genitals. Although the two anatomical structures also have some tissue-specific enhancer interactions, the basic organization of the chromatin architecture is shared. A very similar architecture is found in the RL of the HoxA cluster, suggesting this organization could predate the duplication of the clusters. Thus, regulatory topologies could both favor and constrain the evolution of developmental genes.

34. Gehrke AR, Schneider I, de la Calle-Mustienes E, Tena JJ, Gomez-Marin C, Chandran M, Nakamura T, Braasch I, Postlethwait JH, Gomez-Skarmeta JL, et al.: **Deep conservation of wrist and digit enhancers in fish.** *Proc Natl Acad Sci U S A* 2015, **112**:803-808.

The authors combines 4C-seq, ATAC-seq, transgenic reporter assays and phylogenetic footprinting using gar sequences to identify several Hox enhancers, conserved across all bony vertebrates, but undetectable in direct tetrapod-teleost comparisons.

35. Braasch I, Gehrke AR, Smith JJ, Kawasaki K, Manousaki T, Pasquier J, Amores A, Desvignes T, Batzel P, Catchen J, et al.: **The spotted gar genome illuminates vertebrate evolution and facilitates human-teleost comparisons.** *Nat Genet* 2016, **48**:427-437.

36. Woltering JM, Noordermeer D, Leleu M, Duboule D: **Conservation and divergence of regulatory strategies at Hox Loci and the origin of tetrapod digits.** *PLoS Biol* 2014, **12**:e1001773.

37. Guerreiro I, Gitto S, Novoa A, Codourey J, Nguyen Huynh TH, Gonzalez F, Milinkovitch MC, Mallo M, Duboule D: **Reorganisation of Hoxd regulatory landscapes during the evolution of a snake-like body plan.** *Elife* 2016, **5**.

The authors show that the striking morphological features of the snake body plan are associated to changes in the expression and regulation of HoxD genes, which show a different usage and genomic distribution of mesodermal enhancers. However, these regulatory changes do not alter the characteristic bimodal HoxD chromatin structure, which is also present in snakes.

38. Nakamura T, Gehrke AR, Lemberg J, Szymaszek J, Shubin NH: **Digits and fin rays share common developmental histories.** *Nature* 2016, **537**:225-228.

39. Acemel RD, Tena JJ, Irastorza-Azcarate I, Marletaz F, Gomez-Marin C, de la Calle-Mustienes E, Bertrand S, Diaz SG, Aldea D, Aury JM, et al.: **A single three-dimensional chromatin compartment in amphioxus indicates a stepwise evolution of vertebrate Hox bimodal regulation.** *Nat Genet* 2016, **48**:336-341.

3D chromatin changes can generate regulatory and morphological novelties in evolution. The authors profile the 3D organization of the amphioxus Hox cluster to reconstruct the evolutionary history of vertebrate Hox RLs. They propose that the bipartite TAD structure of jawed vertebrates evolved from a single TAD in the chordates ancestor by the incorporation of a new TAD neighboring posterior Hox genes.

40. Hay D, Hughes JR, Babbs C, Davies JO, Graham BJ, Hanssen LL, Kassouf MT, Oudelaar AM, Sharpe JA, Suci MC, et al.: **Genetic dissection of the alpha-globin super-enhancer in vivo.** *Nat Genet* 2016, **48**:895-903.

By mutating and functionally characterizing each of the constituent CREs of a super-enhancer controlling the α -globin genes, this study shows that there is no evidence that this cluster of CREs acts as an integrated functional unit. The authors conclude that this super-enhancer is simply an assembly of conventional enhancers of varying strengths.

41. Shin HY, Willi M, Yoo KH, Zeng X, Wang C, Metser G, Hennighausen L: **Hierarchy within the mammary STAT5-driven Wap super-enhancer.** *Nat Genet* 2016, **48**:904-911.

The authors report the functional dissection of a super-enhancer region in the *Wap* locus. In contrast to [40], they found that the different CREs composing the super-enhancer are functionally and hierarchically interdependent.

42. Pott S, Lieb JD: **What are super-enhancers?** *Nat Genet* 2015, **47**:8-12.
43. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA: **Master transcription factors and mediator establish super-enhancers at key cell identity genes.** *Cell* 2013, **153**:307-319.
44. Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, Hoke HA, Young RA: **Super-enhancers in the control of cell identity and disease.** *Cell* 2013, **155**:934-947.
45. Loven J, Hoke HA, Lin CY, Lau A, Orlando DA, Vakoc CR, Bradner JE, Lee TI, Young RA: **Selective inhibition of tumor oncogenes by disruption of super-enhancers.** *Cell* 2013, **153**:320-334.
46. Maeso I, Tena JJ: **Favorable genomic environments for cis-regulatory evolution: A novel theoretical framework.** *Semin Cell Dev Biol* 2016, **57**:2-10.
47. Guo Y, Xu Q, Canzio D, Shou J, Li J, Gorkin DU, Jung I, Wu H, Zhai Y, Tang Y, et al.: **CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function.** *Cell* 2015, **162**:900-910.
48. de Wit E, Vos ES, Holwerda SJ, Valdes-Quezada C, Verstegen MJ, Teunissen H, Splinter E, Wijchers PJ, Krijger PH, de Laat W: **CTCF Binding Polarity Determines Chromatin Looping.** *Mol Cell* 2015, **60**:676-684.
49. Narendra V, Rocha PP, An D, Raviram R, Skok JA, Mazzoni EO, Reinberg D: **CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation.** *Science* 2015, **347**:1017-1021.

Figure 1

Oncogene activation, morphological alterations, regulatory novelties