EvoDevo

HYPOTHESIS

Open Access

The application of irreversible genomic states to define and trace ancient cell type homologies

Oleg Simakov[1*] and Günter P. Wagner[2,3,4]

Abstract

Homology, or relationship among characters by common descent, has been notoriously difficult to assess for many morphological features, and cell types in particular. The ontogenetic origin of morphological traits means that the only physically inherited information is encoded in the genomes. However, the complexity of the underlying gene regulatory network and often miniscule changes that can impact gene expression, make it practically impossible to postulate a clear demarcation line for what molecular signature should "define" a homologous cell type between two deeply branching animals. In this Hypothesis article, we propose the use of the recently characterized irreversible genomic states, that occur after chromosomal and sub-chromosomal mixing of genes and regulatory elements, to dissect regulatory signatures of each cell type into irreversible and reversible configurations. While many of such states will be non-functional, some may permanently impact gene expression in a given cell type. Our proposal is that such evolutionarily irreversible, and thus synapomorphic, functional genomic states can constitute a criterion for the timing of the origin of deep evolutionary cell type homologies. Our proposal thus aims to close the gap between the clearly defined homology of the individual genomic characters and their genomic states to the homology at the phenotypic level through the identification of the underlying evolutionarily irreversible and regulatory linked states.

Keywords Chromosome evolution, Enhancer–promoter contacts, Fusion-with-mixing, Entropy, Irreversibility

Homology at the organismal level, including developmental processes [1, 2], organ and cell type homology [3–5] has been one of the most discussed and debated concepts in the field of evolution and development [6]. At the molecular level, lately fueled by the technological advancement in single cell transcriptomics, the "molecular signature" of many morphological or developmental traits could be identified [3, 7–11]. However, complex interplay between regulatory networks during development [12] and the transcription factor logic associated with cell types [3] makes it difficult to identify a clear shared set of genes that identify a given cell type for clades that have been separated from each other for hundreds of millions of years.

High genomic evolvability and complexity of genomic features that can impact gene regulatory networks and their phenotypic outcomes is, arguably, the main unresolved problem when attempting to identify the mechanistic basis of homologies at the organismal level

*Correspondence:
Oleg Simakov
oleg.simakov@univie.ac.at
[1] Department of Neurosciences and Developmental Biology, University of Vienna, Djerassiplatz 1, 1030 Vienna, Austria
[2] Department of Evolutionary Biology, University of Vienna, Djerassiplatz 1, 1030 Vienna, Austria
[3] Systems Biology Institute, Yale University, New Haven, CT 06520, USA
[4] Hagler Institute for Advanced Studies, Texas A&M, College Station, TX 77843, USA

and among distantly related clades. Many genomic changes, from small nucleotide substitutions, enhancer evolution to duplications of genes and regulatory sites [13–15] can lead to evolutionary novelty. However, these changes can also equally lead to loss of such novel characters. It is often unclear how often such events can or have occurred on the vast macro-evolutionary (above clade-level) time-scale, and how often regulatory wiring can be reversed to the ancestral states [16]. Furthermore, it has been found that the genetic basis for clearly homologous characters can be different, in particular in terms of its inducing factors during development [17], while the core identity mechanisms are much more conserved [18].

At the 64th Phyletic Symposium in Jena last year, the lively discussion highlighted again the particular problems when assessing the homology of phenotypes and developmental processes that comprise the hourglass model [7, 19–21]. Ideally, a clear signature of a cell type or a developmental process rooted in the underlying genomic information needs to be identified.

In this hypothesis we argue for a possible link between genome structure, cell type identity, and developmental process homology at the macro-evolutionary level. We propose the identification of irreversible or highly "entangled" genomic states [22] that may affect gene expression and that can be related to the origin of a particular cell type, cell type family, or the underlying developmental process. We suggest that irreversibility is the key property in any identification of distant organismal homologies, as, considering the macro-evolutionary time-scales, there is otherwise no theoretical boundary to (re)evolve various gene expression patterns. Only gene regulation that is linked to a specific irreversible genomic state can be treated as a stable synapomorphic character within a given clade. Evolutionary irreversibility of a given genomic configuration also implies that without the information from the outgroup species about the possible ancestral state, we irrevocably lose this ancestral state information.

Over the recent years, accumulation of chromosomal-scale genomes has enabled us to trace the evolution of whole chromosomes and study their composition as a function of orthologous gene families [23, 24]. One of the key identified properties for animal chromosomal evolution was the so-called "fusion-with-mixing" (Fig. 1): when two ancestrally conserved chromosomes undergo fusion, the genes on them

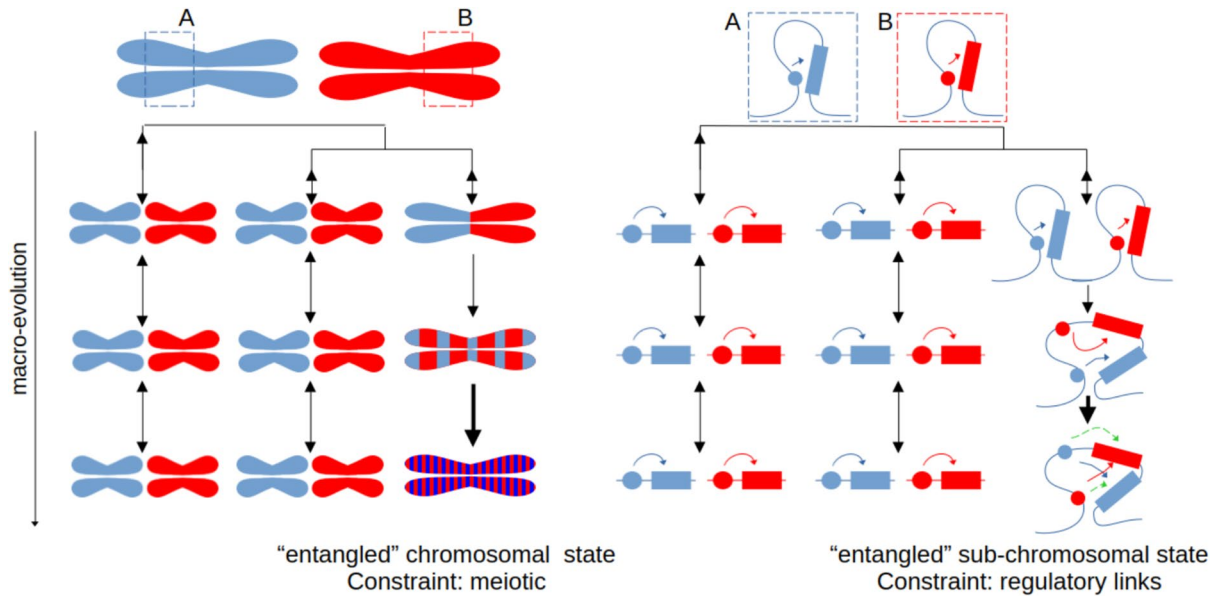

Fig. 1 Irreversible genomic states at chromosomal (left) and sub-chromosomal (right) scales, their origin, and occasional function. Left: chromosomal fusion-with-mixing occurs via, e.g., Robertsonian translocation, followed by intra-chromosomal inversions. The information about the ancestral two states (two separate chromosomes) is lost after the mixing and if no plesiomorphic (outgroup) information is available. Right: similar mixing can be observed for the more functionally relevant enhancer–promoter (E–P) contacts within a single chromosome (labeled as region "A" and region "B", with E–P links shown in blue and red, respectively). E–P links are mixed via intra-chromosomal inversions and translocations within an interactive environment (mediated, e.g., through DNA loop-extrusion) making it unlikely for random inversions to disentangle them into the original state without breaking functional E–P contacts. Over longer time-scales, this entanglement may lead to the evolution of novel (green arrows) persistent E–P links. Black vertical arrows indicate possible evolutionary transitions between homologous states (two-way arrow: reversible; one-way arrow: irreversible). Slightly thicker one-way arrow for the final mixed state suggests higher level of its entropic mixing

mix through intra-chromosomal translocations and the resulting state (mixed chromosome) cannot be reverted to the original two states comprising the two ancestral gene complements [23]. This chromosomal-scale mixing happens either when whole chromosomes or their arms fuse ("algebraic") [23, 24], or after large genomic rearrangement events that break chromosomal homologies ("non-algebraic") [25–29]. It thus constitutes a very strong synapomorphic character that, once established, cannot be reverted and is expected to be observed in all descendants of a particular lineage. This property has already been utilized to shed new light onto highly debated phylogenetic positions [30].

The observed maintenance of chromosomal-scale linkages can be largely explained by meiotic constraints [31–33] and so far little evidence exists of any regulatory function [34]. On the other hand, how genes explore their local, sub-chromosomal, interactive environments and the impact of this process for the evolution of novel gene regulation has already been suggested (e.g., the addition of hundreds of novel topological and co-regulated units after large-scale genome rearrangements [35]). More recent insights suggest that, due to the ongoing process of intra-chromosomal translocations, hundreds of stable regulatory interactions can emerge even within fully retained (unfused) chromosomes since their origin in the metazoan ancestor [22]. The constraint for the maintenance of such sub-chromosomal linkages can be diverse. Mixing of enhancer–promoter (E–P) interactions within a topologically interacting space (e.g., via loop-extrusion or topologically associating domains, TADs [36–39], as well as loop or meta-loop structures [40]) may create an entangled configuration that is very unlikely to be unmixed by random inversions, as these would otherwise break functional E–P contacts (Fig. 1). This constraint thus leads to mixing that is analogous to the chromosomal fusion-with-mixing, but on a much smaller, sub-chromosomal, level. Similar to the deeply conserved chromosomal-level synteny, such constraints may result in retention of unrelated genes and their regulatory regions within specific genomic neighborhoods (Fig. 1). This prediction is corroborated both by the frequently observed micro-synteny in animal genomes, including genomic regulatory blocks, the bystander model [41–43], as well as co-expressed or co-regulated regions [44–46]. These results were also recently complemented by the emergent data on genomic topological structure presence and conservation across animals [36, 38, 47], as well as by the findings that translocations usually happen at the TAD boundaries [48].

In this "mixing" view, the observed maintenance of local linkages does not imply an immediate functional advantage, with any potential synergistic function evolving after this initial entanglement (Fig. 1). Novel functional interactions may thus arise with substantial delay after the original entanglement. However, the irreversibility of this evolutionary process (no separation into ancestrally separate regulatory units) enables us to screen such states for specific changes in gene expression and, eventually, to gene expression that is associated with cell type development or function. This implies that gene sets that define each cell type or a developmental stage can be analyzed in terms of their synapomorphic states or regulatory entanglements. Phylogenetic dating of such regulatory entanglements and quantification of their irreversibility in particular can indicate at what evolutionary node the novelty arose and that a scenario of re-ancestralization (unmixing or disentanglement) can be ruled out. It also enables to define the probability of convergence of such mixed states, as has been done for chromosomal-level fusion-with-mixing events, which will depend on the number of the involved genomic elements that undergo mixing [24, 30].

The exact quantification and the methodology of the identification of such states begins to emerge [22], building on novel interdisciplinary applications, including topological theories [49] in macro-evolution. Such ideas may establish a fruitful testing ground for many deep evolutionary phenotype homology hypotheses. A signature of such entangled states would comprise of a set of enhancers and their target gene(s) located within one interactive region (e.g., a TAD or a loop) where the homologous regions in multiple outgroup species are located either in separate interactive environments or on different chromosomes [35] (Fig. 1). The changes in the gene expression associated with a particular cell type function and development should also be tested in this context and it should be expected that the regulatory entanglement facilitated the emergence of a more complex regulatory logic, e.g., by creating additive sub-functionalized enhancers or entangled super-enhancers within that region [50, 51]. Furthermore, regulatory entanglement may also lead to the accumulation of several such states within the same genomic locus, i.e., two already entangled states undergoing further fusion and mixing. In the developmental context, an example may constitute the vertebrate HoxD cluster and complex regulation of the C-TAD and T-TAD regions around it, compared to the plesiomorphic invertebrate state [41, 52–54]. While complex enhancer logic in many systems and loci has been reported, including enhancers that act over larger genomic distances [55], the determination of how much of it arose through regulatory entanglement is lacking. It is important to note that, similar to chromosomal rearrangement events in some clades that break the ancestral metazoan chromosomal homologies

[25], sub-chromosomally mixed clusters such as Hox can also break apart and scramble within or between chromosomes (e.g., Hox cluster scrambling in some metazoan lineages [42, 56, 57]). However, as at the chromosomal level, sub-chromosomal fission products are not homologous to the ancestral pre-fusion or pre-mixing state, i.e., they cannot theoretically revert back to any of the proposed ancestral Hox configurations [58].

An interesting emerging aspect of this thinking is the question whether genome expansions and contractions can facilitate the origin of the entanglements. In particularly small genomes [59, 60], regulation is often restricted to very proximal regions or even introns. The role of genome topology has also been discussed or disputed in some model organisms [40, 61, 62]. Such genome compaction may facilitate the evolution of segregated states, with proximal or intronic regulation [60]. This, in turn, may enable the observed very fast genomic reshuffling in these clades. Fast turn-over rates of regulatory and coding sequence evolution in such genomes also highlights that the process of separating genes into distinct regulatory units does not mean "disentanglement" and reintroduction of the ancestral state but rather the evolution of a new homologous unit. Contrary to compact genomes, in larger and expanding genomes, accumulation of transposable elements and increased E–P distances may lead to the formation of constrained entangled states that are susceptible for single inversion or translocation events to break existing regulatory links. This may lead to the “fossilization” of such entangled states and, counterintuitively, maintenance of a more ancestral genome architecture, as has been observed for some of the larger and expanded animal genomes (e.g., [24, 63]). Finally, duplications and losses at local and whole genome level can substantially impact the mixing dynamics, producing very different entangled regions in phyla that experienced one or another type of such evolutionary modification (e.g., [64]). Importantly, loss of entangled states or their decomposition into new regulatory units does not mean loss of phenotypes. Rather, it defines the limit on the ability to trace their homology and, if such configurations can be linked to phenotypes, the homology of phenotypes.

In general, this logic may be seen similar to what has been described for phylostratigraphy approaches using orphan (novel) genes and other genomic changes that are evolutionary very rare [65, 66]. However, novel genes by definition have no homology relationship to the outgroups where those genes are not found. Thus, such characters have limited implications for ancient cell type origination. In this context, it is also important to note that we do not propose that irreversible states resulting from mixing of genes and regulatory elements are the only driving force in phenotype or specifically cell type evolution. Clearly, many studies have shown a tremendous multitude of genomic changes that can result in changes in the gene expression. For core regulatory complexes (CoRCs), co-evolution and co-adaptation among transcription factor proteins has been proposed which may comprise a novel "mixed" state [3, 67]. However, without a clear understanding of transition and reversibility properties of such changes across macro-evolution their implications for the cell type constituting molecular signatures are limited. We can thus envisage the next macro-evolutionary genomics frontier that will encompass studying irreversibility properties in the evolution of a plethora of such states that define cell type and organ development. If our hypothesis is true, we would also expect that many recently uncovered regulatory changes that are associated with major innovations (e.g., [68–72]) are likely embedded into larger entangled environments, which, in turn, facilitate their long-term maintenance.

In summary, we propose that to be able to identify homology at the macro-evolutionary scale, irreversibility of the underlying genomic states, if present, may comprise a key criterion. We argue that viewing genomes as "fields", i.e., studying positional characters and their regulatory entanglement along chromosomes, may provide a testable concept to help address this long-lasting question. Knowledge and phylogenetic dating of entangled genomic states will be useful in directing sequencing efforts of species that may still retain some of the ancestral unmixed configurations, helping quantify how much of the ancestral information has been lost due to this irreversible mixing. Finally, further dissection of such states across animals will help us refine the concept of homology and identify what levels of development or the resulting morphological organization are most applicable to study in terms of their underlying genomic configurations.

Acknowledgements
Authors would like to thank all members of the Simakov lab, Ludwik Gąsiorowski, and Eve Seuntjens for valuable discussions and input on the manuscript. We thank three anonymous reviewers for their very valuable and constructive comments on the manuscript.

Author contributions
OS and GPW wrote the manuscript.

Funding
O.S. was supported by the European Research Council’s Horizon 2020: European Union Research and Innovation Programme, grant No. 945026. GPW acknowledges research funding from the Hagler Institute for Advanced Studies at Texas A&M University.

Availability of data and materials
No datasets were generated or analysed during the current study.

Declarations

Competing interests
The authors declare no competing interests.

Received: 27 January 2025 Accepted: 21 April 2025
Published online: 03 May 2025

References

1. von Baer KE. Über Entwicklungsgeschichte der Thiere. Königsberg: Bornträger; 1828.
2. Hall BK. Phylotypic stage or phantom: is there a highly conserved embryonic stage in vertebrates? Trends Ecol Evol. 1997;12:461–3.
3. Arendt D, Musser JM, Baker CVH, Bergman A, Cepko C, Erwin DH, et al. The origin and evolution of cell types. Nat Rev Genet. 2016;17:744–57.
4. Valentine JW. Cell types, numbers, and body plan complexity. In: Hall BK, editor. Keywords and concepts in evolutionary developmental biology. Cambridge: Harvard University Press; 2006. p. 35–43.
5. Wagner GP. Homology, genes, and evolutionary innovation. 1st ed. Princeton: Princeton University Press; 2014.
6. Gould SJ. Ontogeny and phylogeny. Cambridge: Harvard University Press; 1985.
7. Kalinka AT, Varga KM, Gerrard DT, Preibisch S, Corcoran DL, Jarrells J, et al. Gene expression divergence recapitulates the developmental hourglass model. Nature. 2010;468:811–4.
8. Martín-Durán JM, Hejnol A. A developmental perspective on the evolution of the nervous system. Dev Biol. 2021;475:181–92.
9. Sebé-Pedrós A, Chomsky E, Pang K, Lara-Astiaso D, Gaiti F, Mukamel Z, et al. Early metazoan cell type diversity and the evolution of multicellular gene regulation. Nat Ecol Evol. 2018;2:1176–88.
10. Sebé-Pedrós A, Saudemont B, Chomsky E, Plessier F, Mailhé MP, Renno J, et al. Cnidarian cell type diversity and regulation revealed by whole-organism single-cell RNA-Seq. Cell. 2018;173:1520-1534.e20.
11. Paganos P, Voronov D, Musser J, Arendt D, Arnone MI. Single cell RNA sequencing of the Strongylocentrotus purpuratus larva reveals the blueprint of major cell types and nervous system of a non-chordate deuterostome. bioRxiv. 2021. https://doi.org/10.1101/2021.03.16.435574v3.
12. Davidson EH, Erwin DH. Gene regulatory networks and the evolution of animal body plans. Science. 2006;311:796–7.
13. Sommer-Trembo C, Santos ME, Clark B, Werner M, Fages A, Matschiner M, et al. The genetics of niche-specific behavioral tendencies in an adaptive radiation of cichlid fishes. Science. 2024;384:470–5.
14. Touceda-Suárez M, Kita EM, Acemel RD, Firbas PN, Magri MS, Naranjo S, et al. Ancient genomic regulatory blocks are a source for regulatory gene deserts in vertebrates after whole-genome duplications. Mol Biol Evol. 2020;37:2857–64.
15. Villar D, Berthelot C, Aldridge S, Rayner TF, Lukk M, Pignatelli M, et al. Enhancer evolution across 20 mammalian species. Cell. 2015;160:554–66.
16. Babonis LS, Enjolras C, Reft AJ, Foster BM, Hugosson F, Ryan JF, et al. Single-cell atavism reveals an ancient mechanism of cell type diversification in a sea anemone. Nat Commun. 2023;14:885.
17. McColgan Á, DiFrisco J. Understanding developmental system drift. Development. 2024;151:dev203054.
18. Wagner GP. The developmental genetics of homology. Nat Rev Genet. 2007;8:473–9.
19. Haeckel E, Haeckel E. Generelle morphologie der organismen. Allgemeine grundzüge der organischen formen-wissenschaft, mechanisch begründet durch die von Charles Darwin reformirte descendenztheorie . Berlin: G. Reimer; 1866. https://catalog.loc.gov/vwebv/search?searchCode=LCCN&searchArg=43047785&searchType=1&permalink=y.
20. Levit GS, Hoßfeld U, Naumann B, Lukas P, Olsson L. The biogenetic law and the Gastraea theory: from Ernst Haeckel's discoveries to contemporary views. J Exp Zool B Mol Dev Evol. 2022;338:13–27.
21. Uesaka M, Kuratani S, Irie N. The developmental hourglass model and recapitulation: an attempt to integrate the two models. J Exp Zool B Mol Dev Evol. 2022;338:76–86.
22. Schultz DT, Blümel A, Destanović D, Sarigol F, Simakov O. Topological mixing and irreversibility in animal chromosome evolution. bioRxiv. 2024. https://doi.org/10.1101/2024.07.29.605683.
23. Simakov O, Marlétaz F, Yue JX, O'Connell B, Jenkins J, Brandt A, et al. Deeply conserved synteny resolves early events in vertebrate evolution. Nat Ecol Evol. 2020;4:820–30.
24. Simakov O, Bredeson J, Berkoff K, Marletaz F, Mitros T, Schultz DT, et al. Deeply conserved synteny and the evolution of metazoan chromosomes. Sci Adv. 2022;8:5884.
25. Albertin CB, Medina-Ruiz S, Mitros T, Schmidbaur H, Sanchez G, Wang ZY, et al. Genome and transcriptome mechanisms driving cephalopod evolution. Nat Commun. 2022;13:2427.
26. Leibowitz ML, Zhang C-Z, Pellman D. Chromothripsis: a new mechanism for rapid karyotype evolution. Annu Rev Genet. 2015;49:183–211.
27. Lewin TD, Liao IJ-Y, Luo Y-J. Annelid comparative genomics and the evolution of massive lineage-specific genome rearrangement in bilaterians. Mol Biol Evol. 2024;41:msae172.
28. Schultz DT, Heath-Heckman EAC, Winchell CJ, Kuo D-H, Yu Y-S, Oberauer F, et al. Acceleration of genome rearrangement in clitellate annelids. bioRxiv. 2024. https://doi.org/10.1101/2024.05.12.593736.
29. Vargas-Chávez C, Benítez-Álvarez L, Martínez-Redondo GI, Álvarez-González L, Salces-Ortiz J, Eleftheriadi K, et al. A punctuated burst of massive genomic rearrangements and the origin of non-marine annelids. bioRxiv. 2024. https://doi.org/10.1101/2024.05.16.594344.
30. Schultz DT, Haddock SHD, Bredeson JV, Green RE, Simakov O, Rokhsar DS. Ancient gene linkages support ctenophores as sister to other animals. Nature. 2023;618:110–7.
31. Lv J, Havlak P, Putnam NH. Constraints on genes shape long-term conservation of macro-synteny in metazoan genomes. BMC Bioinform. 2011;12:S11.
32. Muller HJ. Bearing of the "Drosophila" work on systematics. The New Systematics. 1940; 185–268. https://ci.nii.ac.jp/naid/10004957361/. Accessed 08 Feb 2021.
33. Wright S. On the probability of fixation of reciprocal translocations. Am Nat. 1941;75:513–22.
34. Clarence T, Robert NSM, Sarigol F, Fu X, Bates PA, Simakov O. Robust 3D modeling reveals spatiosyntenic properties of animal genomes. iScience. 2023;26:106136.
35. Schmidbaur H, Kawaguchi A, Clarence T, Fu X, Hoang OP, Zimmermann B, et al. Emergence of novel cephalopod gene regulation and expression through large-scale genome reorganization. Nat Commun. 2022;13:2172.
36. Acemel RD, Lupiáñez DG. Evolution of 3D chromatin organization at different scales. Curr Opin Genet Dev. 2023;78:102019.
37. Harmston N, Ing-Simmons E, Tan G, Perry M, Merkenschlager M, Lenhard B. Topologically associating domains are ancient features that coincide with Metazoan clusters of extreme noncoding conservation. Nat Commun. 2017;8:441.
38. Rogers TF, Simakov O. Emerging questions on the mechanisms and dynamics of 3D genome evolution in spiralians. Brief Funct Genomics. 2023;22:elad043.
39. Szalay M-F, Majchrzycka B, Jerković I, Cavalli G, Ibrahim DM. Evolution and function of chromatin domains across the tree of life. Nat Struct Mol Biol. 2024;31:1824–37.
40. Mohana G, Dorier J, Li X, Mouginot M, Smith RC, Malek H, et al. Chromosome-level organization of the regulatory genome in the *Drosophila* nervous system. Cell. 2023;186:3826-3844.e26.
41. Robert NSM, Sarigol F, Zimmermann B, Meyer A, Voolstra CR, Simakov O. Emergence of distinct syntenic density regimes is associated with early metazoan genomic transitions. BMC Genomics. 2022;23:143.
42. Simakov O, Marletaz F, Cho S-JJ, Edsinger-Gonzales E, Havlak P, Hellsten U, et al. Insights into bilaterian evolution from three spiralian genomes. Nature. 2013;493:526–31.
43. Zimmermann B, Robert NSM, Technau U, Simakov O. Ancient animal genome architecture reflects cell type identities. Nat Ecol Evol. 2019;3:1289–93.
44. Engström PG, Ho Sui SJ, Drivenes O, Becker TS, Lenhard B. Genomic regulatory blocks underlie extensive microsynteny conservation in insects. Genome Res. 2007;17:1898–908.

45. Irimia M, Tena JJ, Alexis MS, Fernandez-Miñan A, Maeso I, Bogdanović O, et al. Extensive conservation of ancient microsynteny across metazoans due to cis-regulatory constraints. Genome Res. 2012;22:2356–67.
46. Kikuta H, Laplante M, Navratilova P, Komisarczuk AZ, Engström PG, Fredman D, et al. Genomic regulatory blocks encompass multiple neighboring genes and maintain conserved synteny in vertebrates. Genome Res. 2007;17:545–55.
47. Acemel RD, Maeso I, Gómez-Skarmeta JL. Topologically associated domains: a successful scaffold for the evolution of gene regulation in animals. Wiley Interdiscip Rev Dev Biol. 2017;6:e265.
48. Álvarez-González L, Burden F, Doddamani D, Malinverni R, Leach E, Marín-García C, et al. 3D chromatin remodelling in the germ line modulates genome evolutionary plasticity. Nat Commun. 2022;13:2608.
49. Erwin DH. The topology of evolutionary novelty and innovation in macroevolution. Philos Trans R Soc Lond B Biol Sci. 2017;372:20160422.
50. Choi J, Lysakovskaia K, Stik G, Demel C, Söding J, Tian TV, et al. Evidence for additive and synergistic action of mammalian enhancers during cell fate determination. eLife. 2021;10:e65381.
51. Rickels R, Shilatifard A. Enhancer logic and mechanics in development and disease. Trends Cell Biol. 2018;28:608–30.
52. Acemel RD, Tena JJ, Irastorza-Azcarate I, Marlétaz F, Gómez-Marín C, de la Calle-Mustienes E, et al. A single three-dimensional chromatin compartment in amphioxus indicates a stepwise evolution of vertebrate Hox bimodal regulation. Nat Genet. 2016;48:336–41.
53. Montavon T, Soshnikova N, Mascrez B, Joye E, Thevenet L, Splinter E, et al. A regulatory archipelago controls hox genes transcription in digits. Cell. 2011;147:1132–45.
54. Rodríguez-Carballo E, Lopez-Delisle L, Willemin A, Beccari L, Gitto S, Mascrez B, et al. Chromatin topology and the timing of enhancer function at the HoxD locus. Proc Natl Acad Sci U S A. 2020;117:31231–41.
55. Batut PJ, Bing XY, Sisco Z, Raimundo J, Levo M, Levine MS. Genome organization controls transcriptional dynamics during development. Science. 2022;375:566–70.
56. Ikuta T, Yoshida N, Satoh N, Saiga H. Ciona intestinalis Hox gene cluster: its dispersed structure and residual colinear expression in development. Proc Natl Acad Sci. 2004;101:15118–23.
57. Seo H-C, Edvardsen RB, Maeland AD, Bjordal M, Jensen MF, Hansen A, et al. Hox cluster disintegration with persistent anteroposterior order of expression in Oikopleura dioica. Nature. 2004;431:67–71.
58. Butts T, Holland PWH, Ferrier DEK. The urbilaterian super-hox cluster. Trends Genet. 2008;24:259–62.
59. Martín-Durán JM, Vellutini BC, Marlétaz F, Cetrangolo V, Cvetesic N, Thiel D, et al. Conservative route to genome compaction in a miniature annelid. Nat Ecol Evol. 2021;5:231–42.
60. Plessy C, Mansfield MJ, Bliznina A, Masunaga A, West C, Tan Y, et al. Extreme genome scrambling in marine planktonic Oikopleura dioica cryptic species. Genome Res. 2024;34:426–40.
61. Ghavi-Helm Y, Jankowski A, Meiers S, Viales RR, Korbel JO, Furlong EEM. Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression. Nat Genet. 2019;51:1272–82.
62. Maeso I, Tena JJ. Favorable genomic environments for cis-regulatory evolution: a novel theoretical framework. Semin Cell Dev Biol. 2016;57:2–10.
63. Meyer A, Schloissnig S, Franchini P, Du K, Woltering J, Irisarri I, et al. Giant lungfish genome elucidates the conquest of land by vertebrates. Nature. 2021. https://doi.org/10.1038/s41586-021-03198-8.
64. Maeso I, Irimia M, Tena JJ, González-Pérez E, Tran D, Ravi V, et al. An ancient genomic regulatory block conserved across bilaterians and its dismantling in tetrapods by retrogene replacement. Genome Res. 2012;22:642–55.
65. Bleidorn C. Rare genomic changes. In: Bleidorn C, editor. Phylogenomics: an introduction. Cham: Springer International Publishing; 2017. p. 195–211.
66. Domazet-Lošo T, Brajković J, Tautz D. A phylostratigraphy approach to uncover the genomic history of major adaptations in metazoan lineages. Trends Genet. 2007;23:533–9.
67. Park Y, Nnamani MC, Maziarz J, Wagner GP. Cis-regulatory evolution of forkhead box O1 (FOXO1), a terminal selector gene for decidual stromal cell identity. Mol Biol Evol. 2016;33:3161–9.
68. Irimia M, Royo JL, Burguera D, Maeso I, Gómez-Skarmeta JL, Garcia-Fernandez J. Comparative genomics of the Hedgehog loci in chordates and the origins of Shh regulatory novelties. Sci Rep. 2012;2:433.
69. Letelier J, de la Calle-Mustienes E, Pieretti J, Naranjo S, Maeso I, Nakamura T, et al. A conserved Shh cis-regulatory module highlights a common developmental origin of unpaired and paired fins. Nat Genet. 2018;50:504–9.
70. Letelier J, Terriente J, Belzunce I, Voltes A, Undurraga CA, Polvillo R, et al. Evolutionary emergence of the rac3b/rfng/sgca regulatory cluster refined mechanisms for hindbrain boundaries formation. Proc Natl Acad Sci U S A. 2018;115:E3731–40.
71. Marlétaz F, de la Calle-Mustienes E, Acemel RD, Paliou C, Naranjo S, Martínez-García PM, et al. The little skate genome and the evolutionary emergence of wing-like fins. Nature. 2023;616:495–503.
72. Real FM, Haas SA, Franchini P, Xiong P, Simakov O, Kuhl H, et al. The mole genome reveals regulatory rearrangements associated with adaptive intersexuality. Science. 2020;370:208–14.

Publisher's Note