Integrative and Comparative Biology

Integrative and Comparative Biology, volume 0, pp. 1–20
https://doi.org/10.1093/icb/icaf040

Society for Integrative and Comparative Biology

INVITED PAPER

Roles and Future Opportunities for Genomic Architecture in Understanding Repeated Evolution

Riley Kellermeyer [*,1], David Alvarez-Ponce[†], Javier Arsuaga[‡], Frédéric J. J. Chain[§], Megan Y. Dennis [¶], Mark J. Margres[‖], Richard P. Meisel [#], Wynn K. Meyer[**], Tae Seok Moon[††], Keriayn N. Smith[‡‡], Thomas D. Wolfe[§], Coral Y. Zhou [§§] and Suzanne E. McGaugh[*]

[*]Department of Ecology, Evolution, and Behavior, University of Minnesota, St. Paul, MN 55108, USA; [†]Biology Department, University of Nevada, Reno, Reno, NV 89557, USA; [‡]Department of Molecular and Cellular Biology and Department of Mathematics, University of California, Davis, Davis, CA 95616, USA; [§]Department of Biological Sciences, University of Massachusetts Lowell, Lowell MA 01854, USA; [¶]Department of Biochemistry and Molecular Medicine, Genome Center, MIND Institute, University of California, Davis, Davis, CA 95616, USA; [‖]Department of Integrative Biology, University of South Florida, Tampa, FL 33620, USA; [#]Department of Biology and Biochemistry, University of Houston, Houston, TX 77004, USA; [**]Department of Biological Sciences, Lehigh University, Bethlehem, PA 18015, USA; [††]Synthetic Biology Group, J. Craig Venter Institute, La Jolla, CA 92037, USA; [‡‡]School of Data Science and Society, Department of Genetics, School of Medicine, University of North Carolina at Chapel Hill , Chapel Hill, NC 27514, USA; [§§]Department of Molecular Biosciences, University of Kansas, Lawrence, KS 66045, USA

[1]E-mail: kell3262@umn.edu

Synopsis The trajectory of evolution is impacted by molecular constraints and biases that are difficult to validate experimentally. Repeated evolution of similar traits across the Tree of Life serves as a natural experiment to discern common factors that drive the evolution of these traits. The architecture of genomes in one-dimensional, two-dimensional, and three-dimensional space is emerging as a potential factor that may predict repeated phenotypic evolution. For example, chromatin packaging and the 3D organization of the genome within the nucleus can impose evolutionary constraints by predisposing genomic regions for particular types of mutations, while the evolution of genome sequence can also drive reorganization of chromatin. With the explosion of new library preparation and sequencing technologies that are accessible for non-model species, we envision a great opportunity to understand how genome architecture across phylogenetically disparate species may impact repeated phenotypic evolution. We provide examples of the known and potential avenues of phenotypic convergence at each level of genome architecture and how integration of these data can provide unique insights into the constraints, trajectory, and predictability of evolution.

Introduction

Repeated evolution is the independent evolution of the same trait in different evolutionary lineages (Stern 2013; Rosenblum et al. 2014; Sackton and Clark 2019). In some cases, the term "convergent evolution" is a clear descriptor, as is in the case of the anatomical convergence of bat, bird, and insect wings. In other cases, especially among closely related lineages, the term "convergence" is less clear for a variety of reasons. For example, independent evolution of the same trait in two species may have occurred through different genetic mechanisms, which would demonstrate phenotypic convergence via divergent molecular evolution.

Here, we discuss repeated evolution to encompass the appearance of similar phenotypic traits across distinct evolutionary lineages, encompassing both convergent and divergent molecular origins (Gompel and Prud'homme 2009). Importantly, repeated evolution is a powerful tool to identify how organisms adapt to environments and acquire new phenotypes (Losos 2011).

Understanding the genetic mechanisms underlying repeated evolution has revealed that similar func-

Advance Access publication May 15, 2025

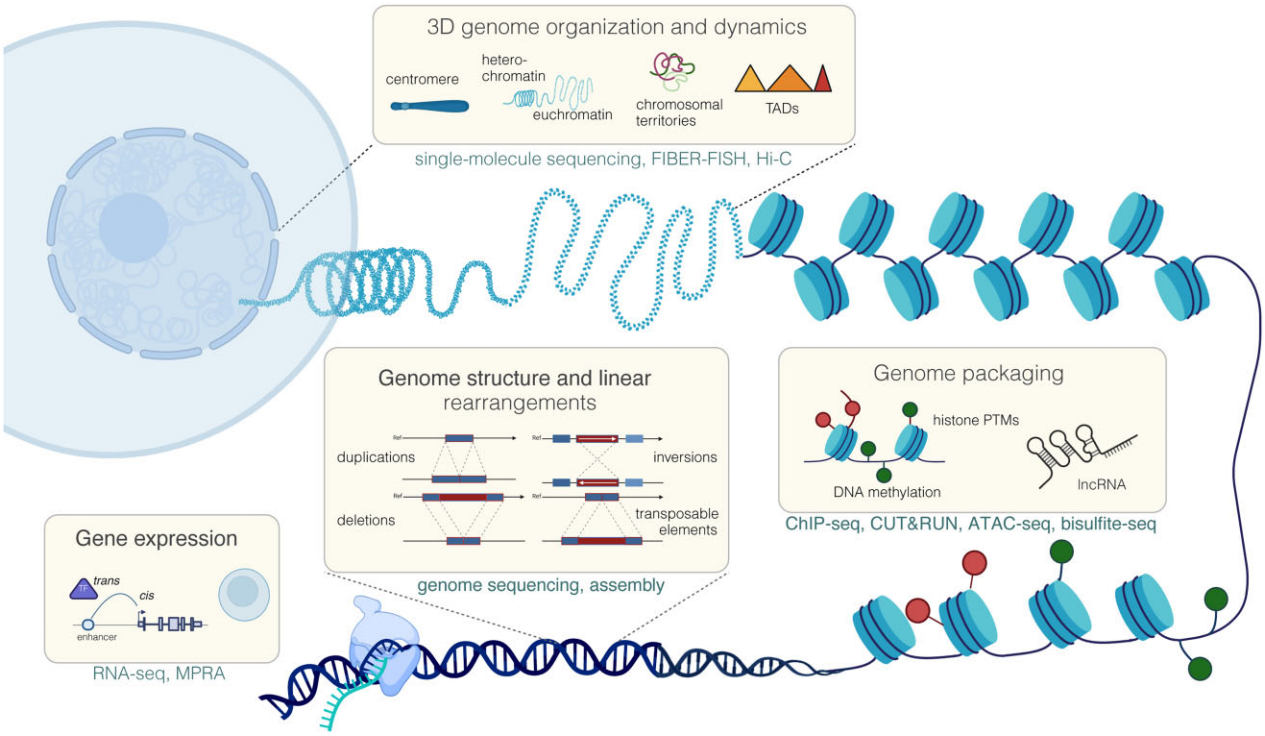

Fig. 1 Overview of repeated evolution and genome architecture. Overview of genome architecture across the one-dimensional linear genome, two-dimensional epigenome, and three-dimensional organization within the nucleus.

tional outcomes can be reached through multiple paths. For example, viviparity (live birth) has independently evolved in mammals, lizards, snakes, and sharks. While the phenotypic trait is convergent, the genetic basis of viviparity has little overlap in these species (Foster et al. 2022). These cases indicate that evolutionary constraints can be influenced by the historical contingencies and genetic background of each lineage. The specific mechanisms responsible for these constraints can be examined by identifying the genomic basis of convergent traits.

Our focus is on how genome architecture—a collective term that includes genome structure, packaging, and organization—relates to repeated evolution. We define genome structure as the linear (one-dimensional) DNA sequence, which changes through duplications, deletions, inversions, and chromosomal fusions/fissions. Genomic DNA is packaged into chromatin (two-dimensional), which includes wrapping of DNA around covalently modified histones to form nucleosomes. Chromatin is further organized into the three-dimensional nucleus through dynamic processes including loop extrusion and phase separation (Fig. 1). All aspects of genome architecture are interconnected; changes in the linear sequence can initiate a cascade of two- and three-dimensional changes, while changes in chromatin and nuclear organization can promote or inhibit sequence changes.

Studying the relationship between genome architecture and repeated phenotypic evolution can provide unique insights into how the organization of the genome may constrain evolution of DNA sequence. When we find repeated associations between specific genomic features and organismal traits—such as in epigenetic mechanisms of dosage compensation—it strengthens our understanding of the forces that constrain evolution. This evidence becomes particularly compelling when these associations arise through repeated evolution rather than shared inheritance. Through this lens, studying repeated evolution at the level of genome architecture helps illuminate both the flexibility and constraints of evolutionary processes.

The relationship between repeated evolution and genome architecture has largely been overlooked because measuring the underlying genomic features was not possible. There are several methodological obstacles to overcome as we improve this field of research. One primary obstacle is defining and measuring convergence itself. For instance, what constitutes "the same" genomic feature across different organisms? This question is particularly relevant for elements like long noncoding RNAs (lncRNAs), where traditional sequence identity measures may be less informative than structural features and GC content (Ross and Ulitsky 2022). Similarly, defining functional equivalence across species (whether features repress or activate the same genes

or bind the same proteins) remains a complex challenge that requires careful consideration of both experimental and computational approaches. Finally, a significant challenge lies in generating comparable data across evolutionarily divergent taxa, particularly when dealing with complex features like 3D genome organization. Newly emerging functional genomic techniques now offer the opportunity to characterize and compare genomic features to better understand their relationships with repeated evolution. Here, we discuss these limitations in understanding genome evolution and explore how new sequencing technologies may help compare genomic changes across the Tree of Life. We envision that the deployment of these techniques across species has the potential to revolutionize our understanding of the genetics of repeated evolution.

Section 1: gene expression

As the initial manifestation of a phenotype, gene expression is an essential link between repeated phenotypic evolution and the evolution of genome architecture. The regulation of RNA transcription occurs through two mechanisms: *cis* regulatory elements (CREs) and *trans* factors (Wittkopp et al. 2004; Signor and Nuzhdin 2018). CREs include DNA sequences, such as promoters and enhancers, which affect the transcription of genes on the same DNA molecule. *Trans* factors are proteins, RNAs, or other molecules that affect expression through DNA binding or signaling pathways. CREs are fixed in the haploid genotype, while *trans* factors can be encoded anywhere in the genome and can freely diffuse to affect gene expression across large distances (kilobases). Therefore, only *trans* factors can vary across cell/tissue types, development, and environments. However, both CREs and *trans* factors can be influenced by genome architecture by, for example, repositioning CREs relative to their target genes, modifying chromatin accessibility, or creating new regulatory domains through chromosome folding. Posttranscriptional regulation of gene expression can be further controlled by RNA modifications, such as methylation, splicing, editing, and polyadenylation.

There is growing evidence that the evolution of gene expression can be responsible for the repeated evolution of organismal phenotypes (Hart et al. 2018; Bittner et al. 2021). Genomic approaches quantifying transcriptomes cover a broad range of resolutions, from across entire organisms (e.g., bulk RNA-seq) down to single-cell and spatial scales. Work using hybrid cell lines/organisms and massively parallel reporter assays is further resolving the relative contributions of *cis* and *trans* effects on gene expression (Gallego Romero and Lea 2023; Dennis 2024), requiring cross-species comparisons across orthologous CREs, genes, and cell/tissue types.

Section 2: genome structure and linear rearrangements

Gene expression depends on the physical proximity of CREs and the genes they regulate within the three-dimensional space of the nucleus (Fig. 1). Genomic rearrangements—including duplications, deletions, and inversions—can alter this three-dimensional organization by rearranging DNA sequences in linear space, thereby affecting gene expression. The detection of these structural rearrangements has been dramatically improved with the advent of long-read sequencing technologies, which provide a more comprehensive and accurate view of chromosomal sequences compared to short-read methods (van Dijk et al. 2023). Of particular interest are the so-called fragile sites—chromosomal loci that are prone to being breakpoints for chromosomal rearrangements (Durkin and Glover 2007; van Dijk et al. 2023). When similar structural rearrangements are observed across different species or populations, that suggests the existence of shared genomic vulnerabilities or preferential breakpoints that can be leveraged during evolution to generate novel phenotypes.

Duplications and deletions

Perhaps the simplest example of sequence changes that affect gene expression is the duplication of a genomic region, which increases the dosage and possibly the expression of genes contained within that region (Birchler and Yang 2022; Zhang et al. 2022). Gene duplication can contribute to evolutionary divergence of gene expression between species, including via neofunctionalization and the evolution of new gene functions (Li et al. 2005). For example, genes within segmental duplications are enriched for differential expression between humans and chimpanzees (Blekhman et al. 2009). Notably, repeated evolution of antifreeze glycoproteins, which bind to ice and inhibit the formation of ice crystals, occurred through independent duplication events in Antarctic and Arctic fishes (Cheng and Chen 1999). This example demonstrates the potential for gene duplications to contribute to evolutionary convergence (Chen et al. 1997).

Chromosomal deletions have the potential for profound impacts on gene expression and repeated genomic regulation. Deletions in regulatory regions can have cascading effects on gene expression and can induce adaptive phenotypic changes. A famous example comes from the repeated loss of spines in sticklebacks that are the result of the deletion of the *Pitx1* enhancer

(Chan et al. 2010). This deletion occurred in regions of the chromosome that are more prone to breakage during replication (i.e., fragile sites). Convergent deletions can be identified with a forward genomics approach, such as the identification of deletions in gulonolactone (L-) oxidase (Gulo) that underlies loss of vitamin C synthesis in eight species across the Tree of Life (Hiller et al. 2012). Thus, understanding relevant forces impacting deletion likelihood has the potential to improve our understanding of the molecular basis of repeated phenotypic evolution.

Inversions

Chromosomal inversions can have important effects on organismal phenotypes and fitness (Wellenreuther and Bernatchez 2018; Berdan et al. 2023). One way in which inversions can affect phenotypes is by changing gene expression levels, causing genes to be differentially expressed between alternative chromosomal arrangements. Expression of genes near inversion breakpoints may be affected if CREs are moved away from gene bodies (Wesley and Eanes 1994; Lavington and Kern 2017). Alternatively, inversions can suppress recombination, resulting in divergent DNA sequence between arrangements (Navarro et al. 1997; Fuller et al. 2016, 2017). The resulting differences in DNA sequence between CREs can cause substantial gene expression differences across the entire inverted region (Fuller et al. 2016; Said et al. 2018). These gene expression differences can have compounding effects across the genome because regulatory networks include genes on inverted and non-inverted chromosomes (Naseeb et al. 2016). Given these substantial effects of inversions on gene expression, they have tremendous potential as mechanistic drivers of repeated evolution.

An illustrative example of chromosomal inversions associated with repeated evolution comes from the evolution of social behavior in ants. Both the Alpine silver ant, *F. selysi*, and the fire ant, *S. invicta*, have independently evolved polymorphisms for social behavior via polymorphic chromosomal inversions (Wang et al. 2013; Purcell et al. 2014). However, there is no homology in the inverted genomic regions between the species, indicating a lack of genetic convergence despite phenotypic (behavioral) convergence.

Transposable elements

Transposable elements (TEs) are mobile DNA sequences that shape genome evolution, particularly through their role in generating novel genes, altering gene expression, and promoting genome expansion (Galbraith and Hayward 2023). TEs preferentially insert into open, euchromatic regions near genes, where they can disrupt gene function, interrupt regulatory regions, and induce structural rearrangements.

Convergent disruption of pigmentation through TE insertions has been found to result in lighter or darker animals through a variety of mechanisms (Galbraith and Hayward 2023). The famous example of rapid evolution of melanism in the peppered moth, *Biston betularia*, during the industrial revolution was the result of a large TE insertion in the intron of *cortex*, a meiosis cell-cycle regulator (Van't Hof et al. 2016). This insertion resulted in increased expression of *cortex* in the wing imaginal disc and resulted in the darker morph (Bannasch et al. 2021). Darker pigmentation in vertebrates has repeatedly evolved via the insertion of TEs that reduce the expression of the Agouti Signaling Protein (ASIP) gene, a pigment production regulator (Ha et al. 2003). As ASIP inhibits the activity of the Melanocortin 1 Receptor (MC1R), TE insertions that disrupt *asip* expression result in increased MC1R activity, leading to a higher production of eumelanin and a darker individual (Trigo et al. 2021; Kamitaki et al. 2024). Convergent TE insertions in other genes result in temporal and spatial dark coat colors and patterns in dogs, as well as darker skin/fur in humans, mice, and cattle (Galbraith and Hayward 2023). Thus, TE insertions are a common mechanism for generating repeated color evolution.

Another compelling example of the repeated evolutionary impact of TEs comes from studies of canine breeds, where independent insertions of FGF4 retrogenes (intronless FGF4 genes that arose via duplication by an mRNA intermediate that was reverse transcribed by an enzyme encoded by an endogenous retrotransposon) on chromosomes 12 and 18 have led to repeated evolution of short-legged phenotypes across different dog breeds (Bannasch et al. 2022). These insertions each impact leg length, as some breeds like Cavalier King Charles Spaniels carry only the chromosome 12 insertion, while others like Cairn Terriers and West Highland White Terriers possess the insertion on chromosome 18, which also results in shorter legs (Dickinson and Bannasch 2020). Notably, dwarfism in humans is frequently attributed to one of the receptors for FGF4, suggesting convergent phenotypes through similar genetic mechanisms (Shiang et al. 1994).

Section 3: genome packaging

Gene expression can be controlled by DNA methylation, histone post-translational modifications (PTMs), and long non-coding RNAs (lncRNAs) interacting with chromatin. Chromatin modifications play crucial roles in regulating gene expression without altering the underlying DNA sequence. DNA methylation involves the

Table 1 Examples of convergent evolution in DNA cytosine methylation.

Type of methylation	Function	Independent evolutionary origins
Gene body	Unclear	Independently evolved in animals and flowering plants (Xiang et al. 2010; Zemach et al. 2010; Bewick and Schmitz 2017; Zilberman 2017)
Variable promoter	Gene silencing	Independently evolved in flowering plants, vertebrates, the demosponge *A. queenslandica*, the centipede *Strigamia maritima,* and the mealybug *Planococcus citri* (Newell-Price et al. 2000; de Mendoza et al. 2019; Lewis et al. 2020; Zhang et al. 2018)
Transposable elements	Transposable element silencing	Independently lost in oysters, hymenopterans, sea urchins, and tunicates (Keller et al. 2016; Strader et al. 2020; Zemach et al. 2010; Wang et al. 2014), or lost in animals and then regained in multiple animal lineages independently (de Mendoza et al. 2019)
Parental genomic imprinting	Diverse roles in growth and cellular proliferation, common regulatory pathways	Independent evolution in mammals and plants (Feil and Berger 2007)
Complete loss of cytosine methylation	Unclear	Independently lost in nematodes, myxozoans, and multiple lineages within insects (Urieli-Shoval et al. 1982; Wenzel et al. 2011; Bewick et al. 2017; Kyger et al. 2021; Engelhardt et al. 2022)

addition of methyl groups to specific DNA sites, particularly at CpG dinucleotides. Histone PTMs include methylation, acetylation, and other covalent modifications predominantly located at the residues in the N-terminal tails of histone proteins, which assemble into the nucleosomes that package DNA into chromatin. The degree of similarity in genomic content and nucleosome positioning among lineages is expected to impose evolutionary constraints that influence the likelihood of repeated evolution, similar to how the divergence between lineages can affect adaptive gene reuse (Bohutínská and Peichel 2024). Both DNA methylation and histone PTMs have emerged as informative systems for studying repeated evolution, where similar traits evolve independently in different lineages. Similarly, technologies that improve the detection and sequencing of lncRNAs allow us to investigate their potential as regulators of gene expression. Together, these regulatory systems can create changes in gene expression and have been repeatedly gained, lost, or modified across diverse species throughout evolution.

DNA methylation

In vertebrates, DNA cytosine methylation (the main form of cytosine methylation) is an essential mechanism involved in gene expression regulation, X chromosome inactivation, repression of repetitive elements, and genome imprinting (Yuasa 2002; Moore et al. 2013). Methyl groups are deposited on cytosines by a conserved family of DNA methyltransferases (Dnmts). The resulting methylcytosines are "read" by methyl-CpG binding domain proteins (MBDs), which recruit proteins that repress transcription.

Even though cytosine methylation is a highly conserved epigenomic mechanism, shared by most eukaryotes, patterns of cytosine methylation and the responsible enzymatic machinery can evolve surprisingly fast and significantly differ among taxa (Alvarez-Ponce et al. 2018; Singh et al. 2021; Sarkies 2022). For instance, while vertebrate and plant genomes are densely methylated, in invertebrates cytosine methylation is often sparse and mostly confined to gene bodies and TEs (de Mendoza et al. 2020); Zhang et al. 2018). In addition, while in vertebrates cytosine methylation mostly affects cytosines that are part of CpG dinucleotides, cytosine methylation in other contexts (CHG and CHH, where H represents any nucleotide) is relatively common in other taxa (e.g., Zhang et al. 2018). Moreover, while the function of cytosine methylation in vertebrates (especially mammals) and plants is well understood, its function in many other lineages is much less clear (de Mendoza et al. 2020; Matlosz et al. 2024), which hinders the interpretation of macroevolutionary cytosine methylation comparisons.

Methylation of different regions of the genome is present in certain evolutionary lineages but absent from others, and the phylogenetic distribution indicates convergent gains or losses (Table 1; Sarkies 2022). Examples include: (1) gene body methylation, whose function is unclear (Xiang et al. 2010; Zemach et al. 2010; Bewick and Schmitz 2017; Zilberman 2017); (2) variable promoter methylation (i.e., the promoters of some genes being more methylated than those of others, which often results in gene silencing; (Newell-Price et al. 2000; Zhang et al. 2018; de Mendoza et al. 2019; Lewis et al. 2020); (3) methylation of TEs, which results in their silencing (Zemach et al. 2010; Wang et al. 2014; Keller

et al. 2016; de Mendoza et al. 2019; Jansz 2019; Strader et al. 2020); (4) parental genomic imprinting (Feil and Berger 2007); and (5) complete loss of cytosine methylation (Urieli-Shoval et al. 1982; Wenzel et al. 2011; Bewick et al. 2017; Kyger et al. 2021; Engelhardt et al. 2022).

Some cases of convergent evolution at the level of phenotype have been linked to convergent changes in cytosine methylation. For example, Haghani et al. (2023) recently analyzed 15,456 samples from 348 mammalian species to generate a "phyloepigenetic" treewhich largely recapitulated the known mammalian phylogeny. They then used unsupervised clustering to identify groups of CpGs whose methylation status co-varied. Many of the 55 identified co-methylation modules correlated with life span. In other cases, phenotypes do not clearly associate with cytosine methylation. For instance, social and solitary insects do not exhibit significant differences in their methylomes (Bewick et al. 2017), and repeated adaptation of sticklebacks to freshwater environments does not seem to be explained by parallel evolved changes in cytosine methylation (Hu and Barrett 2023).

Notably, similar methylomes have evolved independently in different lineages, but the reasons and phenotypic consequences remain unknown. For instance, the demosponge *Amphimedon queenslandica* exhibits a highly dense methylome that resembles vertebrate methylomes in many aspects (80% of CpGs are methylated), while invertebrate methylomes, including those of other sponges, are often sparsely methylated. It is unclear why an organism with a small genome and only a few cell types would evolve a vertebrate-like methylome (de Mendoza et al. 2019). In contrast, different invertebrate lineages have independently lost their ability to methylate their DNA, including dipterans, most nematodes, and myxosporeans, among other lineages (Urieli-Shoval et al. 1982; Wenzel et al. 2011; Bewick et al. 2017; Kyger et al. 2021; Engelhardt et al. 2022). It remains unclear why these organisms lost cytosine methylation and how they evolved compensatory mechanisms for gene silencing (Sarkies et al. 2015; Chang and Liao 2017).

Cytosine methylation might promote convergent evolution not only through its effects on gene expression, but also through its mutagenic effects and its effects on higher-order 3D architecture. CpG dinucleotides are prone to C-to-T transition mutations due to deamination of methylated cytosines, which could be a source of convergent mutations in independent lineages (Ehrlich and Wang 1981; Hwang and Green 2004). In addition, cytosine methylation intrinsically alters chromatin structure, e.g., by reducing DNA flexibility and favoring heterochromatic states (Buitrago et al. 2021).

Histone PTMs

Histone PTMs consist of chemical changes, such as acetylation, methylation, phosphorylation, and ubiquitination, that are critical for regulating the 3D structure and function of the genome (Borg et al. 2021). These modifications affect chromatin accessibility, shifting the ability of nuclear molecules to physically contact genomic DNA, which can induce, enhance, or repress transcription (Reik 2007; Klemm et al. 2019; Han et al. 2023). Such changes in chromatin conformation and transcription can be heritable, as known in *Caenorhabditis elegans* (Özdemir and Steiner 2022). Changes in chromatin conformation also affect many cellular processes, including DNA replication (mitosis and meiosis) and genome stability (apoptosis, DNA damage, and repair) (Millán-Zambrano et al. 2022). Further, environmental stressors and conditions, such as temperature, can shape chromatin organization via histone modifications (Perrella et al. 2020; Kumar et al. 2021). The dynamic interplay between cellular environment, chromatin landscape, and gene expression patterns—where environmental stimuli can alter chromatin states and chromatin modifications themselves drive developmental trajectories—suggests a prominent role for histone PTMs in evolutionary phenomena like phenotypic and molecular convergence.

Similar histone PTM effects on related genes have the potential to convergently shape gene regulatory networks and other processes such as genomic imprinting in plants and animals (Feil and Berger 2007). For example, in *Capsella rubella*, independent mutations in the 5′ region of the *FLC* locus in two populations led to an increase in repressive histone PTMs and a decrease in activating histone PTMs regulating that locus (Yang et al. 2018). The result was a convergent reduction in flowering times via the reduced expression of the FLC transcription factor (Fig. 2). Evolutionary genetic studies that integrate the analysis of histone PTMs with downstream events of gene expression might help clarify the nature of molecular mechanisms of convergence.

TE insertions can create binding sites for transcription factors like CTCF, which modify chromatin accessibility (Diehl et al. 2020; Ichiyanagi et al. 2021; Fueyo et al. 2022; Choudhary et al. 2023). In addition, histone PTMs on TEs can spread to nearby euchromatic loci, affecting expression of neighboring genes and even mediating long-range chromosome interactions between euchromatic and pericentromeric regions (Lee and Karpen 2017; Lee et al. 2020; Di Stefano 2022). Convergent gene regulatory networks can emerge from independent insertions of TEs among diverged species (Ellison and Bachtrog 2013; Lucas et al. 2018; Ellison and Bachtrog 2019), raising the possibility that histone

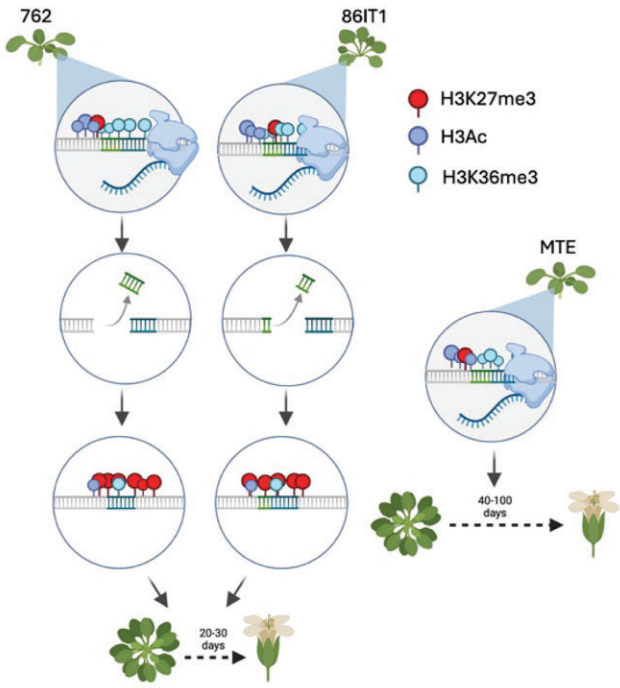

Fig. 2 Independent mutations in the 5′ UTR of the FLC locus in two *C. rubella* populations (sampled accessions 762 and 86IT1) were associated with convergent decreases in activating histone PTMs (H3Ac and H3K36me3) and an increase in a repressive PTM (H3K27me3) surrounding the locus. The FLC locus in 762 and 86IT1 had reduced expression, resulting in a shift in flowering time relative to a third population (accession MTE) lacking the deletion (Yang et al. 2018).

PTMs can shape this process. Prolonged changes in epigenetic states may spur long-term adaptation by remodeling nucleosomes and the underlying genetic mutations that influence nucleosome-positioning (Choi and Kim 2009).

Independently evolved sex chromosome regulators provide clear evidence for repeatedly evolved gene regulation that relies upon the same, convergently derived histone PTMs. Sex chromosomes have evolved independently in many animal lineages (Bachtrog et al. 2014). In each case, the existence of a heterogametic sex with only a single X (or Z) chromosome creates a stoichiometric imbalance that can be compensated via upregulation of the hemizygous sex chromosome. In *Drosophila*, transcriptional upregulation of X-linked genes in hemizygous males is accomplished by acetylation of histone H4 at lysine 16 (H4K16ac) by the histone acetyltransferase MOF (Lucchesi and Kuroda 2015). H4K16ac is also enriched on the mammalian X chromosome via a mechanism that may involve the homolog of MOF (Deng et al. 2013), suggesting repeated evolution of X up-regulation amongst independently evolved X chromosomes (Deng and Disteche 2019). Moreover, a newly evolved Z chromosome arm in the monarch butterfly also has enriched H4K16ac and is transcriptionally upregulated in hemizygous females (Gu et al. 2019). These patterns provide evidence for repeated evolution of H4K16ac as a critical component in dosage compensation of hemizygous sex chromosomes across animals. It is now within our reach to integrate multi-level processes of repeated evolution by studying the interactions between genomic sequence and histone modifiers that shape gene regulation.

Long non-coding RNAs

Long non-coding RNAs (lncRNAs) are molecules that influence diverse nuclear and cytoplasmic processes, including transcriptional regulation, the 3D organization of chromatin, translation control, and cell signaling (Noh et al. 2018; Rinn and Chang 2020; Andergassen and Rinn 2022; Mattick et al. 2023). Nuclear-localized lncRNAs modulate gene expression through interactions with chromatin to regulate genome packaging and gene silencing. These lncRNAs also interact with RNA binding proteins that generally recognize short sequence motifs (k-mers; Ross and Ulitsky 2022), suggesting that consistent protein-RNA interactions may drive functional convergence even when broader sequence conservation is difficult to detect due to the lack of extended conserved regions.

LncRNA genes evolve under different sequence constraints than protein-coding genes, allowing them to achieve similar functions through diverse sequence

paths (Necsulea et al. 2014; Palazzo and Koonin 2020). Even well-studied and functionally conserved lncRNAs like *XIST* and *JPX* show substantial sequence divergence between mouse and humans (Pontier and Gribnau 2011; Karner et al. 2020). This lack of extensive sequence conservation makes it challenging to infer lncRNA functions using conventional approaches that rely on identifying conserved domains (Kirk et al. 2018). However, newer analytical methods, such as k-mer content analysis and structure-based approaches (Kirk et al. 2018), have enabled the identification of functionally convergent lncRNAs across species despite their sequence diversity.

Several compelling examples highlight the potential of studying evolution through lncRNAs. In mammals, X chromosome inactivation is mediated by two different lncRNAs that evolved independently: *XIST* in eutherian mammals and *RSX* in metatherians like opossums (Grant et al. 2012; Furlan and Rougeulle 2016; Sprague et al. 2019; McIntyre et al. 2024). Despite their distinct evolutionary origins, both achieve similar functions in X-inactivation through chromosome coating and silencing. In addition, an independently evolved pair of lncRNAs (*roX1* and *roX2*) are required for dosage compensation of the X chromosome in *Drosophila* males (Franke and Baker 1999; Meller and Rattner 2002), although the nature of sex-specific X chromosome regulation differs greatly between mammals and *Drosophila* (Gu and Walters 2017). Similarly, the *EVX1AS-like* lncRNA in Madagascar geckos performs comparable developmental regulatory functions to human *EVX1AS* despite not being homologous (Olazagoitia-Garmendia et al. 2023).

Section 4: 3D genome organization and dynamics

The various layers of genome architecture discussed above coalesce to form a 3D genome that is precisely and dynamically organized and reorganized throughout the cell cycle and during major developmental transitions. Folding of chromosomes influences how different *cis* and *trans* components physically interact with one another, including the interactions between enhancers, promoters, and transcription factor binding sites (Kim and Shendure 2019). Genomic structure and spatial organization in the nucleus thus affect gene regulation and behave as a constraint on genome function. Here, we discuss two distinct features of meso-scale genome organization: chromosome topology (heterochromatin and centromeres) and nuclear organization (topologically associated domains). We emphasize the molecular machines that create these structures as well as their evolutionary history and potential.

Heterochromatin

Heterochromatin is often associated with the nuclear periphery and is also known to play a role in chromosome organization. These effects are detected at the level of chromosome territories, which are the broad scale organization of chromosomes during the early phase of the cell cycle. Chromosome territories are found across all domains of life (Cremer and Cremer 2010). In human lymphocytes, chromosome territories are organized by gene density, with gene rich chromosomes at the center of the nucleus and gene poor on the periphery, and by chromosome clusters with the cluster at the center of the nucleus and the cluster at the nucleolus (Cornforth et al. 2002; Arsuaga et al. 2004). These results, together with new Hi-C and microscopy observations of the genome suggest that interactions between heterochromatic regions, between heterochromatic regions and the lamina, and between heterochromatic regions and the nucleolus play a key role in the broad organization of genomes (Falk et al. 2019; Peng et al. 2023).

Satellite DNA repeats and TEs are often physically compartmentalized away from transcribed regions of the genome into heterochromatin domains within the nucleus. A key feature of heterochromatin in all organisms is di- or tri-methylation of histone H3 at lysine position 9 (H3K9me2/3). Associated with this mark are the families of methyltransferases that deposit this mark (Suv39 and SET) and heterochromatin protein 1 (HP1) that binds ands spread along chromatin containing H3K9me2/3 (Bell et al. 2023). Together, this set of molecular players establishes and maintains heterochromatin constitutively across the cell cycle and in all cell types. H3K9 methylation evolved shortly after the expansion of the long interspersed nuclear element 1 (LINE-1) retrotransposon and now makes up 20% of the human genome (Malik et al. 1999). While the ancestral role of H3K9 methylation may have been silencing of LINE-1, heterochromatin is now well conserved in both animals and plants, silencing large swaths of the genome beyond non-LTRs (Kabi and Filion 2021).

In general, heterochromatin is transcriptionally inert and can evolve rapidly, resulting in differences at the DNA sequence level even among closely related species (Hughes and Hawley 2009). The repetitive elements so often found in heterochromatin require careful regulation to replicate, repair, and recombine (Feng and Michaels 2015), leading to major expansions or deletions that can get transmitted into the germline during meiosis. For example, a 359-bp repetitive DNA element in *Drosophila* has diverged so rapidly that it is a source of reproductive isolation between two closely related species, *D. simulans* and *D. melanogaster* (Ferree

and Barbash 2009). Hybrid crosses between these two species induce lethality due to failure to maintain the 359-bp repeat during mitosis, resulting in lagging chromosomes and embryo death. Interestingly, the DNA "detangler" topoisomerase II was partially responsible for this effect by improperly localizing to the lagging 359-bp DNA element during anaphase in hybrid embryos (Ferree and Barbash 2009).

Satellite DNA regions can also co-evolve with the proteins that bind to them, leading to an "arms race" between DNA and protein elements that must cooperate to form silenced heterochromatin. For example, the satellite DNA binding factor OdsH from *D. mauritiana* binds to the heterochromatinized Y chromosome of *D. simulans*, whereas the *D. simulans* OdsH does not, leading to hybrid sterility between the two species (Bayes and Malik 2009). In addition, a series of molecular evolution studies of the HP1 gene family in over 40 species of *Drosophila* revealed that, while most HP1 genes are well conserved, the rapidly evolving HP1 genes are predominantly expressed in the germline (Levine et al. 2012). For example, the *rhino* gene is exclusively expressed in the female germline during oogenesis (Vermaak et al. 2005), raising the possibility that Rhino could compete in the centromere drive model of evolution, where only one of four meiotic products is destined to become a viable egg (Henikoff et al. 2001). Another HP1 isoform, HP1E, is expressed in the male germline in *D. melanogaster* and protects the paternal genome during mitosis in the early embryo (Levine et al. 2015). Thus, the independent evolution of HP1 duplicates across *Drosophila* suggests repeated evolution of paralogs in the germline across different species.

More generally, multiple HP1 paralogs are found in many other eukaryotes, each with its own unique function for regulating gene expression. Fission yeast have two paralogs of HP1, while humans have three: HP1α, which binds to constitutive heterochromatin, and HP1å and HP1β, both with roles in transcription activation (Fanti and Pimpinelli 2008; Bosch-Presegué et al. 2017). At the molecular level, the three paralogs differ significantly in two unstructured regions (Canzio et al. 2014), which have also been implicated in creating phase-separated heterochromatin droplets both *in vivo* and *in vitro* (Larson et al. 2017; Strom et al. 2017; Feric and Misteli 2021). Together, these studies suggest that HP1 genes have rapidly diversified to serve many different functions in the genome. Further comparative studies across lineages could reveal whether or not the diversification of HP1 proteins is related to repeated phenotypic evolution, possibly via differential regulation of chromatin to produce different gene expression outcomes.

Holocentric chromosomes

Centromeres are regions of chromosomes that regulate the partitioning of genetic material between daughter cells during cell division through physical linkage via the spindle (Kursel and Malik 2016). Centromeres are found in all eukaryotes and commonly occur in single genomic regions (i.e., where the kinetochore protein complex assembles), forming monocentric chromosomes. In some lineages, however, centromeric activity can be distributed along the entire length of the chromosome (Mola and Papeschi 2006; Melters et al. 2012; Escudero et al. 2016), resulting in holocentric chromosomes. Holocentric chromosomes were first described by (Schrader 1935), and although most eukaryotes have monocentric chromosomes, holocentric chromosomes may have independently evolved at least 19 times across ~800 species of plants (six origins) and animals (13 origins; Melters et al. 2012; Escudero et al. 2016; Mandrioli and Manicardi 2020). There are also multiple examples of reversions from holocentry to monocentry in both animals and plants, raising the possibility that holocentric chromosomes may in fact be the ancestral eukaryotic state (Escudero et al. 2016).

The repeated evolution of holocentric chromosomes across diverse eukaryotic lineages remains puzzling, particularly given the inherent meiotic challenges that require specialized solutions like inverted meiosis. While it was hypothesized that holocentricity could accelerate chromosomal evolution by facilitating fission and fusion events (Melters et al. 2012), studies investigating relationships between holocentricity and chromosome number have reached ambiguous conclusions (Escudero et al. 2016; Mandrioli and Manicardi 2020; Ruckman et al. 2020; Wright et al. 2024). Similarly, evidence for the impacts of holocentricity on diversification rates is mixed (Escudero et al. 2016), with potential effects on speciation in *Carex* (a flowering plant with a lot of variation in chromosome number; Tribble et al. 2025), and on reinforcement observed in lepidopterans (Lukhtanov et al. 2018) but not in other holocentric insects (Ruckman et al. 2020).

Understanding the evolutionary implications of centromere organization requires studying convergent traits across both holocentric and monocentric lineages, rather than comparing chromosome stability and diversification rates. Specific venomous lineages provide a unique model system in which to investigate the repeated evolution of holocentric chromosomes and phenotypic traits. Venoms are one of the most common and convergent functions among animals, with > 200,000 venomous species from >100 venom-origin events (Zancolli and Casewell 2020). Some spiders, scorpions, and centipedes are both venomous and have holo-

centric chromosomes (Melters et al. 2012; Escudero et al. 2016; Mandrioli and Manicardi 2020). Because most genes are intrinsically oriented near centromeres in holocentricity, venoms represent a unique opportunity for identifying correlations between centromere evolution and organismal phenotypes, possibly allowing us to discern rules and idiosyncrasies of centromeric constraint on subsequent trait evolution. For example, proximity to centromeres generally reduces evolutionary rates (Akhunov et al. 2003), but venom genes typically evolve very rapidly (Rokyta et al. 2013). Holocentric chromosomes lack a defined centromeric region, which means that no genes evolve slowly because of proximity to the centromere. Because scorpions exhibit both holocentric and monocentric chromosomes (Riess et al. 1978; Mattos et al. 2018), comparing venom genes across different centromeric states should elucidate how such organizations directly affect evolutionary rates. Ultimately, most of our knowledge on the molecular machinery of holocentric chromosomes is based on work in *C. elegans* (Dernburg 2001). To identify evolutionary biases associated with different chromosomal organizations, we must broaden our focus from model systems that represent a minimal part of the Tree of Life to non-model systems that enable extensive taxonomic sampling of centromere evolutionary dynamics.

Topologically associating domains and Lamina associated domains

Chromatin in eukaryotic genomes is organized into topologically associating domains (TADs), subdomains, loops, and insulation neighborhoods within the nucleus (Dixon et al. 2012; Dowen et al. 2014; Rao et al. 2014; Hafner et al. 2023). While TADs are not universal across species, with many plants lacking well-defined TAD structures, alternative patterns like A/B compartments and chromatin loops serve as similar organizing principles (Di Stefano and Nützmann 2021). TADs can contain both genomic regions that are close on a linear chromosome and segments of multiple chromosomes. TADs can be identified using chromatin conformation capture (3C) approaches (e.g., Hi-C sequencing (Lieberman-Aiden et al. 2009), which identifies DNA regions that are physically close in 3D space by cross-linking these regions and capturing the resulting DNA pairs. TADs are fundamental units of chromosome folding, conserved across cell types and within species (Dekker and Heard 2015; Dixon et al. 2016; Kentepozidou et al. 2020). TADs function to isolate heterochromatic regions from actively transcribed areas to prevent their silencing signals from spreading to active regions and regulate enhancer-promoter interactions (Phillips-Cremins and Corces 2013), such that their disruption through chromosomal rearrangements can alter gene expression and organismal phenotypes (Lupiáñez et al. 2015; Franke et al. 2016; Shanta et al. 2020; Galupa et al. 2022).

TADs work in concert with lamina associated domains (LADs), which are the interaction of heterochromatin and the nuclear lamina and are highly transcriptionally repressed. The most basic example of the role of heterochromatin in nuclear attachment is the Rabl configuration. The Rabl configuration is characterized by the attachment of centromeres and telomeres, both rich heterochromatic regions, to the nuclear envelope (Rabl 1885). Rabl appears to be specific to fungi and certain plants (Santos and Shaw 2004) but may be present in the early developmental stages of development in other organisms (Stevens et al. 2017). The sophisticated interaction between heterochromatin and the lamina results in LADs that range in size from 0.1 to 10 megabases (Guelen et al. 2008; Kind et al. 2015) and may play a dynamic role in gene regulation (Pascual-Reguant et al. 2018; Briand and Collas 2020).

TADs are established through loop extrusion by cohesin complexes, which requires both the conserved cohesin machinery and the placement of boundary elements like CTCF sites, whose positioning and binding motifs can vary across species (Hansen et al. 2018; Hehmeyer et al. 2023). Both TADs and the boundaries between them can be evolutionarily conserved (Dixon et al. 2012; Krefting et al. 2018; Fudenberg and Pollard 2019; Hoencamp et al. 2021), suggesting there are selective constraints against chromosomal rearrangements that disrupt TADs. These selective constraints could explain why evolutionarily conserved TAD boundaries are also found across cell types and contain alleles that are associated with phenotypic variation (McArthur and Capra 2021). Therefore, alteration of some TADs or their boundaries may have deleterious phenotypic effects, creating selective constraints that may limit the possible trajectories of chromosomal evolution.

Despite the selective constraint and conservation of TADs, there is also evolutionary turnover at the TAD boundaries that could be linked to the chromatin state of the TAD (Torosin et al. 2022; Okhovat et al. 2023). Evolutionary divergence of TADs and their boundaries provides additional mechanisms linking evolution of genome structure and repeated evolution (Sarni et al. 2020; Álvarez-González et al. 2022). For example, the breakpoints of chromosomal inversions in *Drosophila* occur more frequently at TAD boundaries than expected by chance (Wright and Schaeffer 2022), suggesting mutational biases for the breakpoints of structural rearrangements. Consistent with such biases, there is extensive evidence for breakpoint reuse of inversions that segregate as polymorphisms within populations or

across species (Pevzner and Tesler 2003; González et al. 2007; Puerma et al. 2016; Corbett-Detig et al. 2019; Orengo et al. 2019; Porubsky et al. 2021). Because of the relationships between TADs and gene expression, repeated evolution of TAD boundaries and chromosomal organization may create opportunities for convergent phenotypic evolution.

The dynamic nature of TADs during the cell cycle and developmental stages makes comparison between species challenging, especially when using existing datasets. Effective TAD comparison requires high-quality reference genomes, matching cell types, cells sorted for interphase, and even then is a challenging computational problem (Zufferey et al. 2018; Li et al. 2022; Sefer 2022). However, by combining lifeOver or other means to identify syntenic regions with methods such as C-InterSecture and Phylo-HMRF, cross-species comparisons are greatly facilitated (Nuriddinov and Fishman 2019; Li et al. 2022; Lukyanchikova et al. 2022). We emphasize that careful considerations of data origin are necessary and standardization of metadata will be essential for further comparison of TADs across evolutionary lineages.

Section 5: the future of genomic architecture in repeated evolution

Incorporating novel methods

Repeated phenotypic evolution across the Tree of Life has the potential to mechanistically link aspects of genome architecture to convergent phenotypes, capturing processes of micro- and macroevolution. Comprehensive study of repeated evolution needs to integrate phylogenetic modeling, genomic sequence data, gene expression analysis, measurements of genome architecture, and cellular and molecular phenotyping. Each of these fields has had technological advances that are primed for new applications, though some challenges remain.

There has been tremendous progress in incorporating phylogenetic models with whole-genome assemblies. Improvements in library preparation, sequencing technologies, and computational methods are democratizing genome assemblies across diverse species that have historically been difficult to obtain (e.g., due to inability to extract large quantities of high-molecular weight DNA or to highly repetitive genomes). Innovations in single-molecule long-read sequencing approaches are beginning to reveal gene expression, regulation, and chromatin organization of gene duplications and complex genomic regions that have been historically inaccessible from standard short-read methods (Stergachis et al. 2020; Zhong et al. 2023). Efforts to systematically produce publicly available quality reference genomes broadly representing the Tree of Life (Darwin Tree of Life Project Consortium 2022; Formenti et al. 2022; Lewin et al. 2022) are enabling more sophisticated analyses across phylogenetically distinct species.

Ubiquitous improvements in ATAC-seq and Hi-C have allowed for assessment of genome architecture in non-model organisms. ATAC-seq and Hi-C have become standard techniques for profiling chromatin in non-model organisms and *de novo* genome assembly, respectively, but remain cost-limiting at the spatial resolution needed for comparative studies. Additionally, some technologies are limited in their application to a broad range of species, such as PHi-C (polymer dynamics deciphered from Hi-C data), which simulates direct promoter-enhancer interactions, but is currently only available for two species (Laverré et al. 2022).

Even more challenging will be approaches to test whether genome architecture affects gene expression in a way that drives repeated evolution of organismal phenotypes. Comparative assessment of gene expression regulation requires careful experimental design. For example, antibody selection for cross-species ChIP-seq needs extensive validation to account for poor epitope conservation in even moderately diverged species (Kidder et al. 2011; Eder and Grebien 2022). Phylogenetic models of gene expression evolution have had substantial progression in the past decade (Dunn et al. 2013; Bertram et al. 2023; Dimayacyac et al. 2023), but require more work to test the comparison of models and to control for experimental artifacts.

There is significant potential to apply recent improvements in molecular techniques for non-model organisms to experimentally validate correlations between genomic architecture and phenotype. For example, CRISPR knockouts of CTCF, which defines TAD boundaries, have been obtained in both mice and tissue culture to show changes in gene expression (Rowley and Corces 2018). These knockouts can be tied to specific tissues and genes in tissue culture, such as CTCF knockout in the *HoxA* locus that regulates motor neurons (Narendra et al. 2015). Furthermore, protein structure prediction tools like AlphaFold can now model sequence-level changes across phylogenetically distinct species. For example, a recent study integrated convergent gene expression changes and protein variants to model genotype-phenotype associations on a macroevolutionary scale (Fukushima and Pollock 2023). Similarly, protein modeling has been applied to protein variants underlying repeated evolution of eye loss in subterranean animals (Kellermeyer et al. 2024). These major advancements in molecular techniques are now available for comparative studies, yet historically underapplied to the field of evolution. We emphasize the potential to use repeated phenotypic evolution to bridge

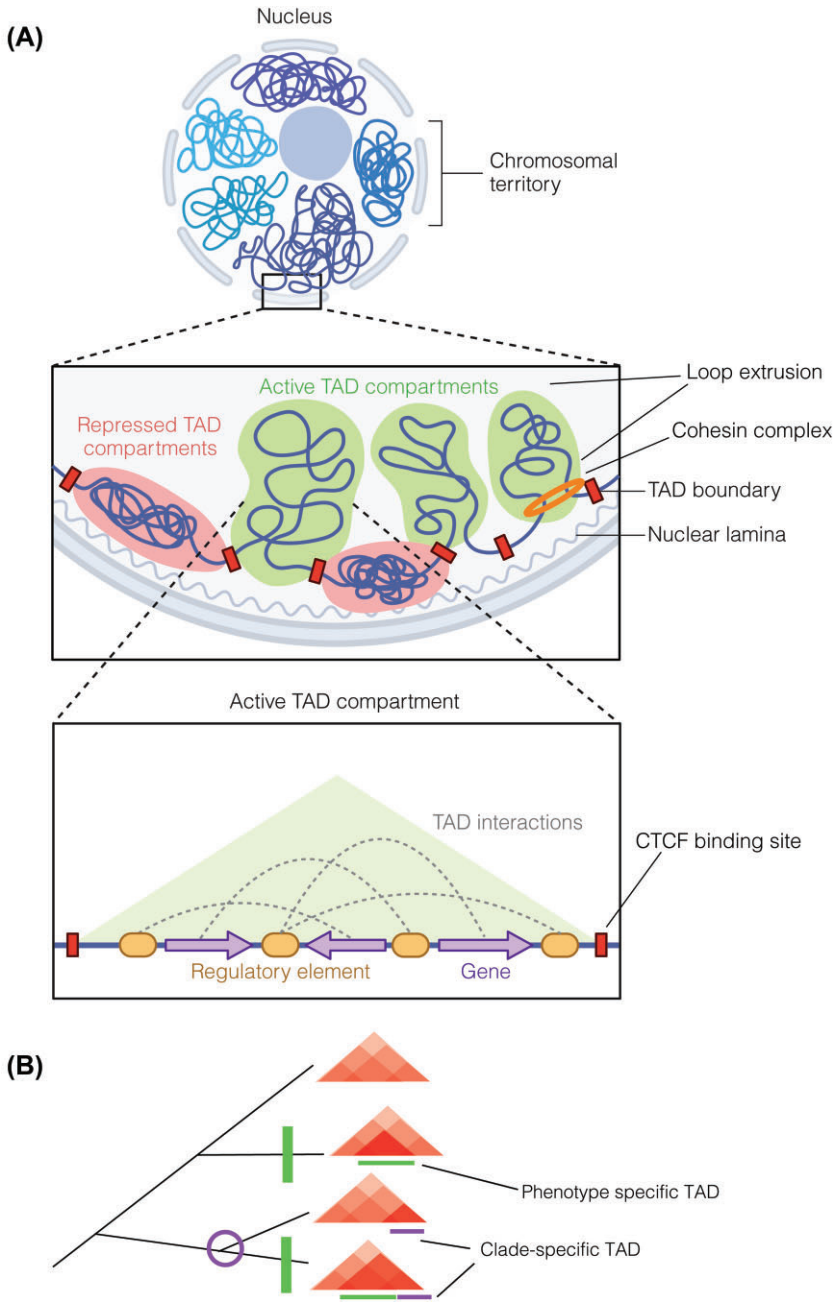

Fig. 3 3D organization of the genome. (A) Within the nucleus, TADs form fundamental organizational subunits. Cohesin complexes form loop extrusions that are transcriptionally active, and CTCF binding sites demarcate boundaries between TADs to isolate heterochromatic regions. (B) Representative schema to identify the TADs in the context of repeated evolution. Comparing Hi-C data between species can distinguish clade-specific TADs (circle) from phenotype-specific TADs (line) in the context of repeated evolution.

the gap between molecular biology and genome evolution.

Understanding the genome's evolution by building it

As an alternative to studying historical evolution via observation-based approaches, building genomes synthetically could reveal complementary insights into the links between genome architecture and repeated evolution (Moon 2023a). Computational models have been developed to simulate genomes that include architectural features, which may be a powerful tool in exploring architectural permutations (Brixi et al. 2025). Additionally, multiple research groups have constructed minimal genomes (Pósfai et al. 2006; Hutchison et al. 2016) or synthetic genomes (Gibson et al. 2010; Richardson et al. 2017). These engineering approaches

have provided some clues on the origin of life and evolution, although many more questions have yet to be answered.

Despite these technological advances in the synthetic genome field, it is still challenging to construct synthetic genomes, let alone to create a synthetic cell. Notably, engineering approaches have been copying nature's blueprint at a gigantic scale to build the "synthetic" cell. As an alternative, we propose using synthetic biology technologies such as genome engineering and DNA synthesis to expand the portfolio of synthetic genomes by constructing eukaryotic genomes that are more than minimalist replicas of yeast and Mycoplasma. We envision using artificial intelligence and other computational tools (Baek et al. 2021; Jumper et al. 2021; Kim et al. 2021; Michaud et al. 2022; Valeri et al. 2023), as well as all the insights gathered by performing large-scale experiments, to design and create an entirely synthetic genome. These experiments may raise ethical dilemmas, which will require new policies for biosafety and biosecurity (Moon 2023b). The design of these experiments could be informed by experiments linking genome architecture and repeated evolution, and they could also inform future data collection to those ends, creating positive feedback between engineering and biology toward a shared understanding of the relationships between genotypes and phenotypes.

Conclusion

Investigating the interplay between genome structure, packaging, and organization provides a transformative lens for understanding repeated evolution. By examining how structural rearrangements, chromatin packaging, and three-dimensional chromatin organization shape gene regulation and phenotypic traits, we gain insights into the evolutionary constraints and flexibility of genomes. Characterizing the mechanisms underlying repeated evolution can provide unique insights into the genetic and molecular basis of complex traits. Advances in sequencing technologies, such as long-read methods and Hi-C mapping, have opened new avenues to uncover these genomic underpinnings across phylogenetically diverse lineages. Moreover, the role of epigenetic systems like histone PTMs, DNA methylation, and lncRNAs highlights the dynamic relationship between environmental pressures, regulatory landscapes, and evolutionary outcomes. By integrating comparative genomic approaches with experimental and computational innovations, future research has the potential to unravel the complex, multiscale processes driving convergence. This synthesis not only deepens our understanding of evolutionary biology but also provides practical implications for synthetic biology and genomic engineering and illuminates the broader principles governing the evolution of life's diversity.

Acknowledgments

We thank Darin Rokyta for input on monomeric and holocentric centromeres. Figures 1 and 3 were adapted from BioRender.com—Ona, S. (2020). Genomic architecture; Huang, E. (2022). Chromosome Organization in Nucleus: TADs.

Funding

This project was funded by the National Science Foundation (NSF; MCB-2326865) for the Leveraging Innovations from Evolution (LIFE) initiative of 2023. We thank the following funding mechanisms for their support: NSF DEB-2316783 (S.M., R.K.); NSF MCB-2145885 (M.Y.D.); NSF DMS-2411979 (J.A.); NSF DBI-2213824 (R.P.M.); NSF MCB- 2144259 (F.J.J.C.); NSF DEB-2324456 (M.J.M.); NSF MCB-2243666 and NSF DBI-2243562 (K.N.S); NSF DBI-2233124 (W.K.M.).

References

Akhunov ED, Goodyear AW, Geng S, Qi L-L, Echalier B, Gill BS, Miftahudin, Gustafson JP, Lazo G, Chao S et al. 2003. The organization and rate of evolution of wheat genomes are correlated with recombination rates along chromosome arms. Genome Res 13:753–63.

Álvarez-González L, Arias-Sardá C, Montes-Espuña L, Marín-Gual L, Vara C, Lister NC, Cuartero Y, Garcia F, Deakin J, Renfree MB et al. 2022. Principles of 3D chromosome folding and evolutionary genome reshuffling in mammals. Cell Rep 41:111839.

Alvarez-Ponce D, Torres-Sánchez M, Feyertag F, Kulkarni A, Nappi T. 2018. Molecular evolution of DNMT1 in vertebrates: duplications in marsupials followed by positive selection. PLoS One 13:e0195162.

Andergassen D, Rinn JL. 2022. From genotype to phenotype: genetics of mammalian long non-coding RNAs in vivo. Nat Rev Genet 23:229–43.

Arsuaga J, Greulich-Bode KM, Vazquez M, Bruckner M, Hahnfeldt P, Brenner DJ, Sachs R, Hlatky L. 2004. Chromosome spatial clustering inferred from radiogenic aberrations. Int J Radiat Biol 80:507–15.

Bachtrog D, Mank JE, Peichel CL, Kirkpatrick M, Otto SP, Ashman T-L, Hahn MW, Kitano J, Mayrose I, Ming R et al. 2014. Sex determination: why so many ways of doing it? PLoS Biol 12:e1001899.

Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, Wang J, Cong Q, Kinch LN, Schaeffer RD et al. 2021. Accurate prediction of protein structures and interactions using a three-track neural network. Science 373:871–6.

Bannasch D, Batcher K, Leuthard F, Bannasch M, Hug P, Marcellin-Little DJ, Dickinson PJ, Drögemüller M, Drögemüller C, Leeb T. 2022. The effects of retrogenes on canine morphology. Genes 325: 13. https://doi.org/10.3390/genes13020325.

Bannasch DL, Kaelin CB, Letko A, Loechel R, Hug P, Jagannathan V, Henkel J, Roosje P, Hytönen MK, Lohi H et al. 2021. Dog colour patterns explained by modular promoters of ancient canid origin. Nat Ecol Evol 5:1415–23.

Bayes JJ, Malik HS. 2009. Altered heterochromatin binding by a hybrid sterility protein in Drosophila sibling species. Science 326:1538–41.

Bell O, Burton A, Dean C, Gasser SM, Torres-Padilla M-E. 2023. Heterochromatin definition and function. Nat Rev Mol Cell Biol 24:691–4.

Berdan EL, Barton NH, Butlin R, Charlesworth B, Faria R, Fragata I, Gilbert KJ, Jay P, Kapun M, Lotterhos KE et al. 2023. How chromosomal inversions reorient the evolutionary process. J Evol Biol 36:1761–82.

Bertram J, Fulton B, Tourigny JP, Peña-Garcia Y, Moyle LC, Hahn MW. 2023. CAGEE: computational analysis of gene expression evolution. Mol Biol Evol 40:msad106. https://doi.org/10.1093/molbev/msad106

Bewick AJ, Schmitz RJ. 2017. Gene body DNA methylation in plants. Curr Opin Plant Biol 36:103–10.

Bewick AJ, Vogel KJ, Moore AJ, Schmitz RJ. 2017. Evolution of DNA methylation across insects. Mol Biol Evol 34:654–65.

Birchler JA, Yang H. 2022. The multiple fates of gene duplications: deletion, hypofunctionalization, subfunctionalization, neofunctionalization, dosage balance constraints, and neutral variation. Plant Cell 34:2466–74.

Bittner NKJ, Mack KL, Nachman MW. 2021. Convergent patterns of gene expression and protein evolution associated with adaptation to desert environments in rodents. In bioRxiv(p. 2021.09.10.459863). https://doi.org/10.1101/2021.09.10.459863 bioRxiv.

Blekhman R, Oshlack A, Gilad Y. 2009. Segmental duplications contribute to gene expression differences between humans and chimpanzees. Genetics 182:627–30.

Bohutínská M, Peichel CL. 2024. Divergence time shapes gene reuse during repeated adaptation. Trends Ecol Evol 39:396–407.

Borg M, Jiang D, Berger F. 2021. Histone variants take center stage in shaping the epigenome. Curr Opin Plant Biol 61:101991.

Bosch-Presegué L, Raurell-Vila H, Thackray JK, González J, Casal C, Kane-Goldsmith N, Vizoso M, Brown JP, Gómez A, Ausió J et al. 2017. Mammalian HP1 isoforms have specific roles in heterochromatin structure and organization. Cell Rep 21:2048–57.

Briand N, Collas P. 2020. Lamina-associated domains: peripheral matters and internal affairs. Genome Biol 21:85.

Brixi G, Durrant MG, Ku J, Poli M, Brockman G, Chang D, Gonzalez GA, King SH, Li DB, Merchant AT et al. 2025. Genome modeling and design across all domains of life with Evo 2. In bioRxiv. https://doi.org/10.1101/2025.02.18.638918 bioRxiv.

Buitrago D, Labrador M, Arcon JP, Lema R, Flores O, Esteve-Codina A, Blanc J, Villegas N, Bellido D, Gut M et al. 2021. Impact of DNA methylation on 3D genome structure. Nat Commun 12:3243.

Canzio D, Larson A, Narlikar GJ. 2014. Mechanisms of functional promiscuity by HP1 proteins. Trends Cell Biol 24:377–86.

Chan YF, Marks ME, Jones FC, Villarreal G, Jr S, M. D, Brady SD, Southwick AM, Absher DM, Grimwood J et al. 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. Science 327:302–5.

Chang AY-F, Liao B-Y. 2017. Recruitment of histone modifications to assist mRNA dosage maintenance after degeneration of cytosine DNA methylation during animal evolution. Genome Res 27:1513–24.

Chen L, DeVries AL, Cheng CH. 1997. Convergent evolution of antifreeze glycoproteins in Antarctic notothenioid fish and Arctic cod. Proc Natl Acad Sci USA 94:3817–22.

Cheng CH, Chen L. 1999. Evolution of an antifreeze glycoprotein. Nature 401:443–4.

Choi JK, Kim Y-J. 2009. Implications of the nucleosome code in regulatory variation, adaptation and evolution. Epigenetics 4:291–5.

Choudhary MNK, Quaid K, Xing X, Schmidt H, Wang T. 2023. Widespread contribution of transposable elements to the rewiring of mammalian 3D genomes. Nat Commun 14:634.

Corbett-Detig RB, Said I, Calzetta M, Genetti M, McBroome J, Maurer NW, Petrarca V, Della Torre A, Besansky NJ. 2019. Fine-mapping complex inversion breakpoints and investigating somatic pairing in the *Anopheles gambiae* species complex using proximity-ligation sequencing. Genetics 213:1495–511.

Cornforth MN, Greulich-Bode KM, Loucas BD, Arsuaga J, Vázquez M, Sachs RK, Brückner M, Molls M, Hahnfeldt P, Hlatky L et al. 2002. Chromosomes are predominantly located randomly with respect to each other in interphase human cells. J Cell Biol 159:237–44.

Cremer T, Cremer M. 2010. Chromosome territories. Cold Spring Harb Perspect Biol 2:a003889.

Darwin Tree of Life Project Consortium. 2022. Sequence locally, think globally: the Darwin Tree of Life Project. Proc Natl Acad Sci USA 119:e2115642118. https://doi.org/10.1073/pnas.2115642118

de Mendoza A, Hatleberg WL, Pang K, Leininger S, Bogdanovic O, Pflueger J, Buckberry S, Technau U, Hejnol A, Adamska M et al. 2019. Convergent evolution of a vertebrate-like methylome in a marine sponge. Nat Ecol Evol 3:1464–73.

de Mendoza A, Lister R, Bogdanovic O. 2020. Evolution of DNA methylome diversity in eukaryotes. J Mol Biol 432:1687–705.

Dekker J, Heard E. 2015. Structural and functional diversity of topologically associating domains. FEBS Lett 589:2877–84.

Deng X, Berletch JB, Ma W, Nguyen DK, Hiatt JB, Noble WS, Shendure J, Disteche CM. 2013. Mammalian X upregulation is associated with enhanced transcription initiation, RNA half-life, and MOF-mediated H4K16 acetylation. Dev Cell 25:55–68.

Deng X, Disteche CM. 2019. Rapid transcriptional bursts upregulate the X chromosome. Nat Struct Mol Biol 26:851–3.

Dennis MY. 2024. Transforming our understanding of species-specific gene regulation. Cell Genomics 4:100540.

Dernburg AF. 2001. Here, there, and everywhere: kinetochore function on holocentric chromosomes. J Cell Biol 153:F33–8.

Di Stefano L. 2022. All quiet on the TE front? The role of chromatin in transposable element silencing. Cells 11:2501. https://doi.org/10.3390/cells11162501.

Di Stefano M, Nützmann H-W. 2021. Modeling the 3D genome of plants. Nucleus 12:65–81.

Dickinson PJ, Bannasch DL. 2020. Current understanding of the genetics of intervertebral disc degeneration. Front Vet Sci 7:431.

Diehl AG, Ouyang N, Boyle AP. 2020. Transposable elements contribute to cell and species-specific chromatin looping and gene regulation in mammalian genomes. Nat Commun 11:1796.
Dimayacyac JR, Wu S, Jiang D, Pennell M. 2023. Evaluating the performance of widely used phylogenetic models for gene expression evolution. Genome Biol Evolut 15:evad211. https://doi.org/10.1093/gbe/evad211
Dixon JR, Gorkin DU, Ren B. 2016. Chromatin domains: the unit of chromosome organization. Mol Cell 62:668–80.
Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485:376–80.
Dowen JM, Fan ZP, Hnisz D, Ren G, Abraham BJ, Zhang LN, Weintraub AS, Schujiers J, Lee TI, Zhao K et al. 2014. Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. Cell 159:374–87.
Dunn CW, Luo X, Wu Z. 2013. Phylogenetic analysis of gene expression. Integr Comp Biol 53:847–56.
Durkin SG, Glover TW. 2007. Chromosome fragile sites. Annu Rev Genet 41:169–92.
Eder T, Grebien F. 2022. Comprehensive assessment of differential ChIP-seq tools guides optimal algorithm selection. Genome Biol 23:119.
Ehrlich M, Wang RY. 1981. 5-Methylcytosine in eukaryotic DNA. Science 212:1350–7.
Ellison C, Bachtrog D. 2019. Contingency in the convergent evolution of a regulatory network: dosage compensation in Drosophila. PLoS Biol 17:e3000094.
Ellison CE, Bachtrog D. 2013. Dosage compensation via transposable element mediated rewiring of a regulatory network. Science 342:846–50.
Engelhardt J, Scheer O, Stadler PF, Prohaska SJ. 2022. Evolution of DNA methylation across ecdysozoa. J Mol Evol 90: 56–72.
Escudero M, Márquez-Corro JI, Hipp AL. 2016. The phylogenetic origins and evolutionary history of holocentric chromosomes. Systematic Botany 41:580–5.
Falk M, Feodorova Y, Naumova N, Imakaev M, Lajoie BR, Leonhardt H, Joffe B, Dekker J, Fudenberg G, Solovei I et al. 2019. Heterochromatin drives compartmentalization of inverted and conventional nuclei. Nature 570:395–9.
Fanti L, Pimpinelli S. 2008. HP1: a functionally multifaceted protein. Curr Opin Genet Dev 18:169–74.
Feil R, Berger F. 2007. Convergent evolution of genomic imprinting in plants and mammals. Trends Genet 23:192–9.
Feng W, Michaels SD. 2015. Accessing the inaccessible: the organization, transcription, replication, and repair of heterochromatin in plants. Annu Rev Genet 49:439–59.
Feric M, Misteli T. 2021. Phase separation in genome organization across evolution. Trends Cell Biol 31:671–85.
Ferree PM, Barbash DA. 2009. Species-specific heterochromatin prevents mitotic chromosome segregation to cause hybrid lethality in Drosophila. PLoS Biol 7:e1000234.
Formenti G, Theissinger K, Fernandes C, Bista I, Bombarely A, Bleidorn C, Ciofi C, Crottini A, Godoy JA, Höglund J et al. 2022. The era of reference genomes in conservation genomics. Trends Ecol Evol 37:197–202.
Foster CSP, Van Dyke JU, Thompson MB, Smith NMA, Simpfendorfer CA, Murphy CR, Whittington CM. 2022. Different genes are recruited during convergent evolution of pregnancy and the placenta. Mol Biol Evol 39:msac077. https://doi.org/10.1093/molbev/msac077
Franke A, Baker BS. 1999. The rox1 and rox2 RNAs are essential components of the compensasome, which mediates dosage compensation in Drosophila. Mol Cell 4:117–22.
Franke M, Ibrahim DM, Andrey G, Schwarzer W, Heinrich V, Schöpflin R, Kraft K, Kempfer R, Jerković I, Chan W-L et al. 2016. Formation of new chromatin domains determines pathogenicity of genomic duplications. Nature 538:265–9.
Fudenberg G, Pollard KS. 2019. Chromatin features constrain structural variation across evolutionary timescales. Proc Natl Acad Sci USA 116:2175–80.
Fueyo R, Judd J, Feschotte C, Wysocka J. 2022. Roles of transposable elements in the regulation of mammalian transcription. Nat Rev Mol Cell Biol 23:481–97.
Fukushima K, Pollock DD. 2023. Detecting macroevolutionary genotype-phenotype associations using error-corrected rates of protein convergence. Nat Ecol Evol 7:155–70.
Fuller ZL, Haynes GD, Richards S, Schaeffer SW. 2016. Genomics of natural populations: how differentially expressed genes shape the evolution of chromosomal inversions in *Drosophila pseudoobscura*. Genetics 204:287–301.
Fuller ZL, Haynes GD, Richards S, Schaeffer SW. 2017. Genomics of natural populations: evolutionary forces that establish and maintain gene arrangements in *Drosophila pseudoobscura*. Mol Ecol 26:6539–62.
Furlan G, Rougeulle C. 2016. Function and evolution of the long noncoding RNA circuitry orchestrating X-chromosome inactivation in mammals. WIREs RNA 7:702–22.
Galbraith JD, Hayward A. 2023. The influence of transposable elements on animal colouration. Trends Genet 39:624–38.
Gallego Romero I, Lea AJ. 2023. Leveraging massively parallel reporter assays for evolutionary questions. Genome Biol 24:26.
Galupa R, Picard C, Servant N, Nora EP, Zhan Y, van Bemmel JG, El Marjou F, Johanneau C, Borensztein M, Ancelin K et al. 2022. Inversion of a topological domain leads to restricted changes in its gene expression and affects interdomain communication. Development 149:dev200568. https://doi.org/10.1242/dev.200568
Gibson DG, Glass JI, Lartigue C, Noskov VN, Chuang R-Y, Algire MA, Benders GA, Montague MG, Ma L, Moodie MM et al. 2010. Creation of a bacterial cell controlled by a chemically synthesized genome. Science 329:52–6.
Gompel N, Prud'homme B. 2009. The causes of repeated genetic evolution. Dev Biol 332:36–47.
González J, Casals F, Ruiz A. 2007. Testing chromosomal phylogenies and inversion breakpoint reuse in Drosophila. Genetics 175:167–77.
Grant J, Mahadevaiah SK, Khil P, Sangrithi MN, Royo H, Duckworth J, McCarrey JR, VandeBerg JL, Renfree MB, Taylor W et al. 2012. Rsx is a metatherian RNA with xist-like properties in X-chromosome inactivation. Nature 487:254–8.
Gu L, Reilly PF, Lewis JJ, Reed RD, Andolfatto P, Walters JR. 2019. Dichotomy of dosage compensation along the neo Z chromosome of the monarch butterfly. Curr Biol 29:4071–4077.e3.
Gu L, Walters JR. 2017. Evolution of sex chromosome dosage compensation in animals: a beautiful theory, undermined by facts and bedeviled by details. Genome Biol Evolut 9: 2461–76.

Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W et al. 2008. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature 453:948–51.
Ha T, Naysmith L, Waterston K, Oh C, Weller R, Rees JL. 2003. Defining the quantitative contribution of the melanocortin 1 receptor (MC1R) to variation in pigmentary phenotype. Ann NY Acad Sci 994:339–47.
Hafner A, Park M, Berger SE, Murphy SE, Nora EP, Boettiger AN. 2023. Loop stacking organizes genome folding from TADs to chromosomes. Mol Cell 83:1377–1392.e6.
Haghani A, Li CZ, Robeck TR, Zhang J, Lu AT, Ablaeva J, Acosta-Rodríguez VA, Adams DM, Alagaili AN, Almunia J et al. 2023. DNA methylation networks underlying mammalian traits. Science 381:eabq5693.
Han M-H, Issagulova D, Park M. 2023. Interplay between epigenome and 3D chromatin structure. BMB Rep 56:633–44.
Hansen AS, Cattoglio C, Darzacq X, Tjian R. 2018. Recent evidence that TADs and chromatin loops are dynamic structures. Nucleus 9:20–32.
Hart JC, Ellis NA, Eisen MB, Miller CT. 2018. Convergent evolution of gene expression in two high-toothed stickleback populations. PLoS Genet 14:e1007443.
Hehmeyer J, Spitz F, Marlow H. 2023. Shifting landscapes: the role of 3D genomic organizations in gene regulatory strategies. Curr Opin Genet Dev 81:102064.
Henikoff S, Ahmad K, Malik HS. 2001. The centromere paradox: stable inheritance with rapidly evolving DNA. Science 293:1098–102.
Hiller M, Schaar BT, Indjeian VB, Kingsley DM, Hagey LR, Bejerano G. 2012. A "forward genomics" approach links genotype to phenotype using independent phenotypic losses among related species. Cell Rep 2:817–23.
Hoencamp C, Dudchenko O, Elbatsh AMO, Brahmachari S, Raaijmakers JA, van Schaik T, Sedeño Cacciatore Á, Contessoto VG, van Heesbeen RGHP, van den Broek B et al. 2021. 3D genomics across the tree of life reveals condensin II as a determinant of architecture type. Science 372:984–9.
Hu J, Barrett RDH. 2023. The role of plastic and evolved DNA methylation in parallel adaptation of threespine stickleback (*Gasterosteus aculeatus*). Mol Ecol 32:1581–91.
Hughes SE, Hawley RS. 2009. Heterochromatin: a rapidly evolving species barrier. PLoS Biol 7:e1000233.
Hutchison CA, 3rd, Chuang R-Y, Noskov VN, Assad-Garcia N, Deerinck TJ, Ellisman MH, Gill J, Kannan K, Karas BJ, Ma L et al. 2016. Design and synthesis of a minimal bacterial genome. Science 351:aad6253.
Hwang DG, Green P. 2004. Bayesian Markov chain Monte Carlo sequence analysis reveals varying neutral substitution patterns in mammalian evolution. Proc Natl Acad Sci USA 101:13994–4001.
Ichiyanagi T, Katoh H, Mori Y, Hirafuku K, Boyboy BA, Kawase M, Ichiyanagi K. 2021. B2 SINE copies serve as a transposable boundary of DNA methylation and histone modifications in the mouse. Mol Biol Evol 38:2380–95.
Jansz N. 2019. DNA methylation dynamics at transposable elements in mammals. Essays Biochem 63:677–89.
Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A et al. 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596:583–9.
Kabi M, Filion GJ. 2021. Heterochromatin: did H3K9 methylation evolve to tame transposons? Genome Biol 22:325.
Kamitaki N, Hujoel MLA, Mukamel RE, Gebara E, McCarroll SA, Loh P-R. 2024. A sequence of SVA retrotransposon insertions in ASIP shaped human pigmentation. Nat Genet 56:1583–91.
Karner H, Webb C-H, Carmona S, Liu Y, Lin B, Erhard M, Chan D, Baldi P, Spitale RC, Sun S. 2020. Functional conservation of LncRNA JPX despite sequence and structural divergence. J Mol Biol 432:283–300.
Keller TE, Han P, Yi SV. 2016. Evolutionary transition of promoter and gene body DNA methylation across invertebrate-vertebrate boundary. Mol Biol Evol 33:1019–28.
Kellermeyer R, Seidel C, Redwine WB, Moran RL, Bertho S, Ornelas-García CP, Alegre D, Weaver K, Unruh J, Troutwine B et al. 2024. Long-term hybridization in a karst window reveals the genetic basis of eye loss in cavefish. In bioRxiv. https://doi.org/10.1101/2024.10.25.620266 bioRxiv.
Kentepozidou E, Aitken SJ, Feig C, Stefflova K, Ibarra-Soria X, Odom DT, Roller M, Flicek P. 2020. Clustered CTCF binding is an evolutionary mechanism to maintain topologically associating domains. Genome Biol 21:5.
Kidder BL, Hu G, Zhao K. 2011. ChIP-Seq: technical considerations for obtaining high-quality data. Nat Immunol 12:918–22.
Kim GB, Gao Y, Palsson BO, Lee SY. 2021. DeepTFactor: a deep learning-based tool for the prediction of transcription factors. Proc Natl Acad Sci USA 118:e2021171118. https://doi.org/10.1073/pnas.2021171118.
Kim S, Shendure J. 2019. Mechanisms of interplay between transcription factors and the 3D genome. Mol Cell 76:306–19.
Kind J, Pagie L, de Vries SS, Nahidiazar L, Dey SS, Bienko M, Zhan Y, Lajoie B, de Graaf CA, Amendola M et al. 2015. Genome-wide maps of nuclear lamina interactions in single human cells. Cell 163:134–47.
Kirk JM, Kim SO, Inoue K, Smola MJ, Lee DM, Schertzer MD, Wooten JS, Baker AR, Sprague D, Collins DW et al. 2018. Functional classification of long non-coding RNAs by k-mer content. Nat Genet 50:1474–82.
Klemm SL, Shipony Z, Greenleaf WJ. 2019. Chromatin accessibility and the regulatory epigenome. Nat Rev Genet 20:207–20.
Krefting J, Andrade-Navarro MA, Ibn-Salem J. 2018. Evolutionary stability of topologically associating domains is associated with conserved gene regulation. BMC Biol 16:1–12.
Kumar S, Kaur S, Seem K, Kumar S, Mohapatra T. 2021. Understanding 3D genome organization and its effect on transcriptional gene regulation under environmental stress in plant: a chromatin perspective. Front Cell Dev Biol 9:774719.
Kursel LE, Malik HS. 2016. Centromeres. Curr Biol 26:R487–90.
Kyger R, Luzuriaga-Neira A, Layman T, Milkewitz Sandberg TO, Singh D, Huchon D, Peri S, Atkinson SD, Bartholomew JL, Yi SV et al. 2021. Myxosporea (Myxozoa, Cnidaria) lack DNA cytosine methylation. Mol Biol Evol 38:393–404.
Larson AG, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, Burlingame AL, Agard DA, Redding S, Narlikar GJ. 2017. Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. Nature 547:236–40.
Laverré A, Tannier E, Necsulea A. 2022. Long-range promoter-enhancer contacts are conserved during evolution and contribute to gene expression robustness. Genome Res. 32:280–96.
Lavington E, Kern AD. 2017. The effect of common inversion polymorphisms in(2L)t and in(3R)Mo on patterns of tran-

scriptional variation in Drosophila melanogaster. G3 7:3659–68.
Lee YCG, Karpen GH. 2017. Pervasive epigenetic effects of euchromatic transposable elements impact their evolution. eLife 6:e25762.https://doi.org/10.7554/eLife.25762
Lee YCG, Ogiyama Y, Martins NMC, Beliveau BJ, Acevedo D, Wu C-T, Cavalli G, Karpen GH. 2020. Pericentromeric heterochromatin is hierarchically organized and spatially contacts H3K9me2 islands in euchromatin. PLoS Genet 16: e1008673.
Levine MT, McCoy C, Vermaak D, Lee YCG, Hiatt MA, Matsen FA, Malik HS. 2012. Phylogenomic analysis reveals dynamic evolutionary history of the Drosophila heterochromatin protein 1 (HP1) gene family. PLoS Genet 8:e1002729.
Levine MT, Vander Wende HM, Malik HS. 2015. Mitotic fidelity requires transgenerational action of a testis-restricted HP1. eLife 4:e07378.
Lewin HA, Richards S, Lieberman Aiden E, Allende ML, Archibald JM, Bálint M, Barker KB, Baumgartner B, Belov K, Bertorelle G et al. 2022. The Earth BioGenome Project 2020: starting the clock. Proc Natl Acad Sci USA 119:e2115635118. https://doi.org/10.1073/pnas.2115635118
Lewis SH, Ross L, Bain SA, Pahita E, Smith SA, Cordaux R, Miska EA, Lenhard B, Jiggins FM, Sarkies P. 2020. Widespread conservation and lineage-specific diversification of genome-wide DNA methylation patterns across arthropods. PLoS Genet 16:e1008864.
Li D, He M, Tang Q, Tian S, Zhang J, Li Y, Wang D, Jin L, Ning C, Zhu W et al. 2022. Comparative 3D genome architecture in vertebrates. BMC Biol 20:99.
Li W-H, Yang J, Gu X. 2005. Expression divergence between duplicate genes. Trends Genet 21:602–7.
Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO et al. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326:289–93.
Losos JB. 2011. Convergence, adaptation, and constraint. Evolution 65:1827–40.
Lucas BA, Lavi E, Shiue L, Cho H, Katzman S, Miyoshi K, Siomi MC, Carmel L, Ares M, Jr, Maquat LE. 2018. Evidence for convergent evolution of SINE-directed Staufen-mediated mRNA decay. Proc Natl Acad Sci USA 115:968–73.
Lucchesi JC, Kuroda MI. 2015. Dosage compensation in Drosophila. Cold Spring Harb Perspect Biol 7:a019398. https://doi.org/10.1101/cshperspect.a019398
Lukhtanov VA, Dinca V, Friberg M, Šíchová J, Olofsson M, Vila R, Marec F, Wiklund C. 2018. Versatility of multivalent orientation, inverted meiosis, and rescued fitness in holocentric chromosomal hybrids. Proc Natl Acad Sci U S A. 115(41):E9610–E9619. https://doi.org/10.1073/pnas.1802610115.
Lukyanchikova V, Nuriddinov M, Belokopytova P, Taskina A, Liang J, Reijnders MJMF, Ruzzante L, Feron R, Waterhouse RM, Wu Y et al. 2022. Anopheles mosquitoes reveal new principles of 3D genome organization in insects. Nat Commun 13:1960.
Lupiáñez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, Horn D, Kayserili H, Opitz JM, Laxova R et al. 2015. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell 161:1012–25.
Malik HS, Burke WD, Eickbush TH. 1999. The age and evolution of non-LTR retrotransposable elements. Mol Biol Evol 16:793–805.
Mandrioli M, Manicardi GC. 2020. Holocentric chromosomes. PLoS Genet 16:e1008918.
Matlosz S, Franzdóttir SR, Pálsson A, Jónsson ZO. 2024. DNA methylation reprogramming in teleosts. Evol Dev 26: e12486.
Mattick JS, Amaral PP, Carninci P, Carpenter S, Chang HY, Chen L-L, Chen R, Dean C, Dinger ME, Fitzgerald KA et al. 2023. Long non-coding RNAs: definitions, functions, challenges and recommendations. Nat Rev Mol Cell Biol 24:430–47.
Mattos VF, Carvalho LS, Carvalho MA, Schneider MC. 2018. Insights into the origin of the high variability of multivalent-meiotic associations in holocentric chromosomes of Tityus (Archaeotityus) scorpions. PLoS One 13:e0192070.
McArthur E, Capra JA. 2021. Topologically associating domain boundaries that are stable across diverse cell types are evolutionarily constrained and enriched for heritability. Am Hum Genet 108:269.
McIntyre KL, Waters SA, Zhong L, Hart-Smith G, Raftery M, Chew ZA, Patel HR, Graves JAM, Waters PD. 2024. Identification of the RSX interactome in a marsupial shows functional coherence with the xist interactome during X inactivation. Genome Biol 25:134.
Meller VH, Rattner BP. 2002. The roX genes encode redundant male-specific lethal transcripts required for targeting of the MSL complex. EMBO J 21:1084–91.
Melters DP, Paliulis LV, Korf IF, Chan SWL. 2012. Holocentric chromosomes: convergent evolution, meiotic adaptations, and genomic analysis. Chromosome Res 20:579–93.
Michaud JM, Madani A, Fraser JS. 2022. A language model beats alphafold2 on orphans. Nat Biotechnol 40:1576–7.
Millán-Zambrano G, Burton A, Bannister AJ, Schneider R. 2022. Histone post-translational modifications—cause and consequence of genome function. Nat Rev Genet 23:563–80.
Mola LM, Papeschi AG. 2006. Holokinetic chromosomes at a glance. J Basic Appl Gen 17:17–33.
Moon TS. 2023. SynHEAL: synthesis of Health equity, advancement, and leadership. ACS Synth. Biol. 12:1583–5.
Moon TS. 2023a. Seven governing principles in biology. Front Synth Biol 1:1296513. https://doi.org/10.3389/fsybi.2023.1296513
Moore LD, Le T, Fan G. 2013. DNA methylation and its basic function. Neuropsychopharmacol 38:23–38.
Narendra V, Rocha PP, An D, Raviram R, Skok JA, Mazzoni EO, Reinberg D. 2015. CTCF establishes discrete functional chromatin domains at the hox clusters during differentiation. Science 347:1017–21.
Naseeb S, Carter Z, Minnis D, Donaldson I, Zeef L, Delneri D. 2016. Widespread impact of chromosomal inversions on gene expression uncovers robustness via phenotypic buffering. Mol Biol Evol 33:1679–96.
Navarro A, Betrán E, Barbadilla A, Ruiz A. 1997. Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. Genetics 146:695–709.
Necsulea A, Soumillon M, Warnefors M, Liechti A, Daish T, Zeller U, Baker JC, Grützner F, Kaessmann H. 2014. The

evolution of lncRNA repertoires and expression patterns in tetrapods. Nature 505:635–40.
Newell-Price J, Clark AJ, King P. 2000. DNA methylation and silencing of gene expression. Trends Endocrinol Metab 11:142–8.
Noh JH, Kim KM, McClusky WG, Abdelmohsen K, Gorospe M. 2018. Cytoplasmic functions of long noncoding RNAs. WIREs RNA 9:e1471.
Nuriddinov M, Fishman V. 2019. C-InterSecture-a computational tool for interspecies comparison of genome architecture. Bioinformatics 35:4912–21.
Okhovat M, VanCampen J, Nevonen KA, Harshman L, Li W, Layman CE, Ward S, Herrera J, Wells J, Sheng RR et al. 2023. TAD evolutionary and functional characterization reveals diversity in mammalian TAD boundary properties and function. Nat Commun 14:1–13.
Olazagoitia-Garmendia A, Senovilla-Ganzo R, García-Moreno F, Castellanos-Rubio A. 2023. Functional evolutionary convergence of long noncoding RNAs involved in embryonic development. Commun Biol 6:908.
Orengo DJ, Puerma E, Aguadé M. 2019. The molecular characterization of fixed inversions breakpoints unveils the ancestral character of the Drosophila guanche chromosomal arrangements. Sci Rep 9:1706.
Özdemir I, Steiner FA. 2022. Transmission of chromatin states across generations in *C. elegans*. Semin Cell Dev Biol 127:133–41.
Palazzo AF, Koonin EV. 2020. Functional long non-coding RNAs evolve from junk transcripts. Cell 183:1151–61.
Pascual-Reguant L, Blanco E, Galan S, Le Dily F, Cuartero Y, Serra-Bardenys G, Di Carlo V, Iturbide A, Cebrià-Costa JP, Nonell L et al. 2018. Lamin B1 mapping reveals the existence of dynamic and functional euchromatin lamin B1 domains. Nat Commun 9:3420.
Peng T, Hou Y, Meng H, Cao Y, Wang X, Jia L, Chen Q, Zheng Y, Sun Y, Chen H et al. 2023. Mapping nucleolus-associated chromatin interactions using nucleolus hi-C reveals pattern of heterochromatin interactions. Nat Commun 14:350.
Perrella G, Zioutopoulou A, Headland LR, Kaiserli E. 2020. The impact of light and temperature on chromatin organization and plant adaptation. J Exp Bot 71:5247–55.
Pevzner P, Tesler G. 2003. Human and mouse genomic sequences reveal extensive breakpoint reuse in mammalian evolution. Proc Natl Acad Sci USA 100:7672–7.
Phillips-Cremins JE, Corces VG. 2013. Chromatin insulators: linking genome organization to cellular function. Mol Cell 50:461–74.
Pontier DB, Gribnau J. 2011. Xist regulation and function explored. Hum Genet 130:223–36.
Porubsky D, Höps W, Ashraf H, Hsieh P, Rodriguez-Martin B, Yilmaz F, Ebler J, Hallast P, Maggiolini FAM, Harvey WT et al. 2021. Haplotype-resolved inversion landscape reveals hotspots of mutational recurrence associated with genomic disorders. In bioRxiv(p.2021.12.20.472354). https://doi.org/10.1101/2021.12.20.472354 bioRxiv.
Pósfai G, Plunkett G, 3rd, Fehér T, Frisch D, Keil GM, Umenhoffer K, Kolisnychenko V, Stahl B, Sharma SS, de Arruda M et al. 2006. Emergent properties of reduced-genome *Escherichia coli*. Science 312:1044–6.
Puerma E, Orengo DJ, Aguadé M. 2016. Multiple and diverse structural changes affect the breakpoint regions of polymorphic inversions across the Drosophila genus. Sci Rep 6:36248.
Purcell J, Brelsford A, Wurm Y, Perrin N, Chapuisat M. 2014. Convergent genetic architecture underlies social organization in ants. Curr Biol 24:2728–32.
Rabl C. 1885. Über zellteilung. In: Gegenbaur C, editors. Morphologisches Jahrbuch 10:214–330.
Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES et al. 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 159:1665–80.
Reik W. 2007. Stability and flexibility of epigenetic gene regulation in mammalian development. Nature 447:425–32.
Richardson SM, Mitchell LA, Stracquadanio G, Yang K, Dymond JS, DiCarlo JE, Lee D, Huang CLV, Chandrasegaran S, Cai Y et al. 2017. Design of a synthetic yeast genome. Science 355:1040–4.
Riess RW, Barker KR, Biesele JJ. 1978. Nuclear and chromosomal changes during sperm formation in the scorpion, *Centruroides Vittatus* (say). Caryologia 31:147–60.
Rinn JL, Chang HY. 2020. Long noncoding RNAs: molecular modalities to organismal functions. Annu Rev Biochem. 89:283–308.
Rokyta DR, Wray KP, Margres MJ. 2013. The genesis of an exceptionally lethal venom in the timber rattlesnake (*Crotalus horridus*) revealed through comparative venom-gland transcriptomics. Bmc Genom 14:394.
Rosenblum EB, Parent CE, Brandt EE. 2014. The molecular basis of phenotypic convergence. Annu Rev Ecol Evol Syst 45:203–26.
Ross CJ, Ulitsky I. 2022. Discovering functional motifs in long noncoding RNAs. WIREs RNA 13:e1708.
Rowley MJ, Corces VG. 2018. Organizational principles of 3D genome architecture. Nat Rev Genet 19:789–800.
Ruckman SN, Jonika MM, Casola C, Blackmon H. 2020. Chromosome number evolves at equal rates in holocentric and monocentric clades. PLoS Genet 16:e1009076.
Sackton TB, Clark N. 2019. Convergent evolution in the genomics era: new insights and directions. Phil Trans R Soc B 374:20190102.
Said I, Byrne A, Serrano V, Cardeno C, Vollmers C, Corbett-Detig R. 2018. Linked genetic variation and not genome structure causes widespread differential expression associated with chromosomal inversions. Proc Natl Acad Sci USA 115:5492–7.
Santos AP, Shaw P. 2004. Interphase chromosomes and the Rabl configuration: does genome size matter? J Microsc 214:201–6.
Sarkies P, Selkirk ME, Jones JT, Blok V, Boothby T, Goldstein B, Hanelt B, Ardila-Garcia A, Fast NM, Schiffer PM et al. 2015. Ancient and novel small RNA pathways compensate for the loss of piRNAs in multiple independent nematode lineages. PLoS Biol 13:e1002061.
Sarkies P. 2022. Encyclopaedia of eukaryotic DNA methylation: from patterns to mechanisms and functions. Biochem Soc Trans 50:1179–90.
Sarni D, Sasaki T, Irony Tur-Sinai M, Miron K, Rivera-Mulia JC, Magnuson B, Ljungman M, Gilbert DM, Kerem B. 2020. 3D genome organization contributes to genome instability at fragile sites. Nat Commun 11:3613.
Schrader F. 1935. Notes an the mitotic behavior of long chromosomes. Cytologia 6:422–30.

Sefer E. 2022. A comparison of topologically associating domain callers over mammals at high resolution. BMC Bioinf 23: 127.
Shanta O, Noor A, Sebat J. Human Genome Structural Variation Consortium (HGSVC), 2020. The effects of common structural variants on 3D chromatin structure. Bmc Genom 21:95.
Shiang R, Thompson LM, Zhu YZ, Church DM, Fielder TJ, Bocian M, Winokur ST, Wasmuth JJ. 1994. Mutations in the transmembrane domain of FGFR3 cause the most common genetic form of dwarfism, achondroplasia. Cell 78: 335–42.
Signor SA, Nuzhdin SV. 2018. The evolution of gene expression in cis and trans. Trends Genet 34:532–44.
Singh D, Sun D, King AG, Alquezar-Planas DE, Johnson RN, Alvarez-Ponce D, Yi SV. 2021. Koala methylomes reveal divergent and conserved DNA methylation signatures of X chromosome regulation. Proc Biol Sci 288:20202244.
Sprague D, Waters SA, Kirk JM, Wang JR, Samollow PB, Waters PD, Calabrese JM. 2019. Nonlinear sequence similarity between the and long noncoding RNAs suggests shared functions of tandem repeat domains. RNA 25:1004–19.
Stergachis AB, Debo BM, Haugen E, Churchman LS, Stamatoyannopoulos JA. 2020. Single-molecule regulatory architectures captured by chromatin fiber sequencing. Science 368: 1449–54.
Stern DL. 2013. The genetic causes of convergent evolution. Nat Rev Genet 14:751–64.
Stevens TJ, Lando D, Basu S, Atkinson LP, Cao Y, Lee SF, Leeb M, Wohlfahrt KJ, Boucher W, O'Shaughnessy-Kirwan A et al. 2017. 3D structures of individual mammalian genomes studied by single-cell hi-C. Nature 544:59–64.
Strader ME, Kozal LC, Leach TS, Wong JM, Chamorro JD, Housh MJ, Hofmann GE. 2020. Examining the role of DNA methylation in transcriptomic plasticity of early stage sea urchins: developmental and maternal effects in a kelp forest herbivore. Front Mar Sci 7:205. https://doi.org/10.3389/fmars.2020.00205
Strom AR, Emelyanov AV, Mir M, Fyodorov DV, Darzacq X, Karpen GH. 2017. Phase separation drives heterochromatin domain formation. Nature 547:241–5.
Torosin NS, Golla TR, Lawlor MA, Cao W, Ellison CE. 2022. Mode and tempo of 3D genome evolution in drosophila. Mol Biol Evol 39:msac216. https://doi.org/10.1093/molbev/msac216
Tribble CM, Márquez-Corro JI, May MR, Hipp AL, Escudero M, Zenil-Ferguson R. 2025. Macroevolutionary inference of complex modes of chromosomal speciation in a cosmopolitan plant lineage. New Phytol 245:2350–61.
Trigo BB, Utsunomiya ATH, Fortunato AAAD, Milanesi M, Torrecilha RBP, Lamb H, Nguyen L, Ross EM, Hayes B, Padula RCM et al. 2021. Variants at the ASIP locus contribute to coat color darkening in Nellore cattle. Genet Sel Evol 53:40.
Urieli-Shoval S, Gruenbaum Y, Sedat J, Razin A. 1982. The absence of detectable methylated bases in Drosophila melanogaster DNA. FEBS Lett 146:148–52.
Valeri JA, Soenksen LR, Collins KM, Ramesh P, Cai G, Powers R, Angenent-Mari NM, Camacho DM, Wong F, Lu TK et al. 2023. BioAutoMATED: an end-to-end automated machine learning tool for explanation and design of biological sequences. Cell Syst 14:525–542.e9.
van Dijk EL, Naquin D, Gorrichon K, Jaszczyszyn Y, Ouazahrou R, Thermes C, Hernandez C. 2023. Genomics in the long-read sequencing era. Trends Genet 39:649–71.
Van't Hof AE, Campagne P, Rigden DJ, Yung CJ, Lingley J, Quail MA, Hall N, Darby AC, Saccheri IJ. 2016. The industrial melanism mutation in British peppered moths is a transposable element. Nature 534:102–5.
Vermaak D, Henikoff S, Malik HS. 2005. Positive selection drives the evolution of rhino, a member of the heterochromatin protein 1 family in Drosophila. PLoS Genet 1:96–108.
Wang J, Wurm Y, Nipitwattanaphon M, Riba-Grognuz O, Huang Y-C, Shoemaker D, Keller L. 2013. A Y-like social chromosome causes alternative colony organization in fire ants. Nature 493:664–8.
Wang X, Li Q, Lian J, Li L, Jin L, Cai H, Xu F, Qi H, Zhang L, Wu F et al. 2014. Genome-wide and single-base resolution DNA methylomes of the Pacific oyster Crassostrea gigas provide insight into the evolution of invertebrate CpG methylation. Bmc Genom 15:1119.
Wellenreuther M, Bernatchez L. 2018. Eco-evolutionary genomics of chromosomal inversions. Trends Ecol Evol 33:427–40.
Wenzel D, Palladino F, Jedrusik-Bode M. 2011. Epigenetics in *C. elegans*: facts and challenges. Genesis 49:647–61.
Wesley CS, Eanes WF. 1994. Isolation and analysis of the breakpoint sequences of chromosome inversion In(3L)Payne in Drosophila melanogaster. Proc Natl Acad Sci USA 91:3132–6.
Wittkopp PJ, Haerum BK, Clark AG. 2004. Evolutionary changes in cis and trans gene regulation. Nature 430:85–8.
Wright CJ, Stevens L, Mackintosh A, Lawniczak M, Blaxter M. 2024. Comparative genomics reveals the dynamics of chromosome evolution in Lepidoptera. Nat Ecol Evol 8:777–90.
Wright D, Schaeffer SW. 2022. The relevance of chromatin architecture to genome rearrangements in Drosophila. Phil Trans R Soc B 3771856:20210206.
Xiang H, Zhu J, Chen Q, Dai F, Li X, Li M, Zhang H, Zhang G, Li D, Dong Y et al. 2010. Single base-resolution methylome of the silkworm reveals a sparse epigenomic map. Nat Biotechnol 28:516–20.
Yang L, Wang H-N, Hou X-H, Zou Y-P, Han T-S, Niu X-M, Zhang J, Zhao Z, Todesco M, Balasubramanian S et al. 2018. Parallel evolution of common allelic variants confers flowering diversity in. Plant Cell 30:1322–36.
Yuasa Y. 2002. DNA methylation in cancer and ageing. Mech Ageing Dev 123:1649–54.
Zancolli G, Casewell NR. 2020. Venom systems as models for studying the origin and regulation of evolutionary novelties. Mol Biol Evol 37:2777–90.
Zemach A, McDaniel IE, Silva P, Zilberman D. 2010. Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science 328:916–9.
Zhang D, Leng L, Chen C, Huang J, Zhang Y, Yuan H, Ma C, Chen H, Zhang YE. 2022. Dosage sensitivity and exon shuffling shape the landscape of polymorphic duplicates in Drosophila and humans. Nat Ecol Evol 6:273–87.
Zhang H, Lang Z, Zhu J-K. 2018. Dynamics and function of DNA methylation in plants. Nat Rev Mol Cell Biol 19: 489–506.
Zhong J-Y, Niu L, Lin Z-B, Bai X, Chen Y, Luo F, Hou C, Xiao C-L. 2023. High-throughput pore-C reveals the single-allele

topology and cell type-specificity of 3D genome folding. Nat Commun 14:1250.

Zilberman D. 2017. An evolutionary case for functional gene body methylation in plants and animals. Genome Biol 18:87.

Zufferey M, Tavernari D, Oricchio E, Ciriello G. 2018. Comparison of computational methods for the identification of topologically associating domains. Genome Biol 19:217.