The relationship between genome structure and function

A. Marieke Oudelaar[1] ✉ *and Douglas R. Higgs*[2] ✉

Abstract | Precise patterns of gene expression in metazoans are controlled by three classes of regulatory elements: promoters, enhancers and boundary elements. During differentiation and development, these elements form specific interactions in dynamic higher-order chromatin structures. However, the relationship between genome structure and its function in gene regulation is not completely understood. Here we review recent progress in this field and discuss whether genome structure plays an instructive role in regulating gene expression or is a reflection of the activity of the regulatory elements of the genome.

Epigenetic
Pertains to changes in chromatin that register, signal or perpetuate altered activity states without changing the primary DNA sequence.

Chromatin
The complex of DNA and proteins that makes up chromosomes; chromatin consists of nucleosomes formed of ~150 bp of DNA wrapped around a histone octamer, which can be further packaged into higher-order structures.

An important aim in biology is to understand the process by which a single fertilized egg develops into a complex multicellular organism comprising hundreds of different cell types with distinct phenotypic and functional characteristics. In principle, we know that this is achieved by appropriately switching on and off different subsets of genes during development and differentiation to establish the complex networks of gene expression that characterize each cell type. This process is encoded in our inherited DNA, which is decoded by proteins that recognize and bind specific sequences in the genome to regulate gene expression. Subsequent epigenetic changes facilitate or inhibit gene activity and, importantly, provide a memory of previous patterns of gene expression associated with cell fate decisions; such modifications play an important role in maintaining cell identity[1].

Over the past 50 years it has emerged that, broadly speaking, there are three fundamental genomic DNA elements controlling gene expression: promoters, enhancers and boundary elements (FIG. 1). Promoters are typically located near the transcription start sites (TSSs) of structural genes. Enhancers and boundary elements may lie within genes, but are more frequently distributed at various distances (1–1,000 kb) flanking the genes in non-coding DNA. Silencers, as opposed to silenced regulatory elements, may represent an additional fundamental element, but at present these are relatively poorly characterized[2,3]. The genes and their regulatory elements occupy only about 5–15% of the human genome[4,5]; the role (if any) of other non-coding sequences is currently unknown. The entire ~6.5 Gb of DNA in a diploid human genome (the equivalent of 2 m of DNA) must be compacted and folded into each cell nucleus. This remarkable degree of packaging is achieved via the incorporation of DNA into chromatin, successive higher-order structures and ultimately chromosomes.

In the past decade, our understanding of both the structural organization of the genome and gene regulatory mechanisms has progressed tremendously[6–9]. The development of various innovative approaches to map genome architecture[10,11] (BOX 1), together with decreasing sequencing costs, has given insight into the nuclear structure of the genome across a wide range of cell populations. Integration of polymer physics and perturbations of proteins with key roles in regulating genome architecture have increased our understanding of the principles underlying the nuclear organization of the genome[12]. Furthermore, the application of high-throughput genome editing and the expanded documentation of naturally occurring mutations affecting the sequences and order of regulatory elements have clarified their roles in regulating gene expression[13].

It has become clear that chromatin exists in a dynamic state and that higher-order chromatin structures differ with local gene activity and epigenetic chromatin modifications. However, an important unanswered question concerns the relationship between the higher-order structure of the genome and the regulation of gene expression. Are chromatin structures encoded independently and do they play an instructive role to regulate or facilitate gene expression, or are these structures an emergent property of the dynamic nuclear processes played out on the regulatory elements of the genome? If the former is correct, we need to discover the rules that govern the independent formation of nuclear structure; if the latter is correct, we need to concentrate on the regulatory elements and processes from which chromatin architecture emerges to improve our understanding of how the genetic code is deciphered and functions in the 3D nucleus.

To explore this question, we first give an overview of the fundamental regulatory elements and the organizational principles of mammalian genomes.

[1]*Max Planck Institute for Biophysical Chemistry, Göttingen, Germany.*

[2]*Laboratory of Gene Regulation, MRC Weatherall Institute of Molecular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, UK.*

✉*e-mail: marieke.oudelaar@mpibpc.mpg.de; doug.higgs@imm.ox.ac.uk*

https://doi.org/10.1038/s41576-020-00303-x

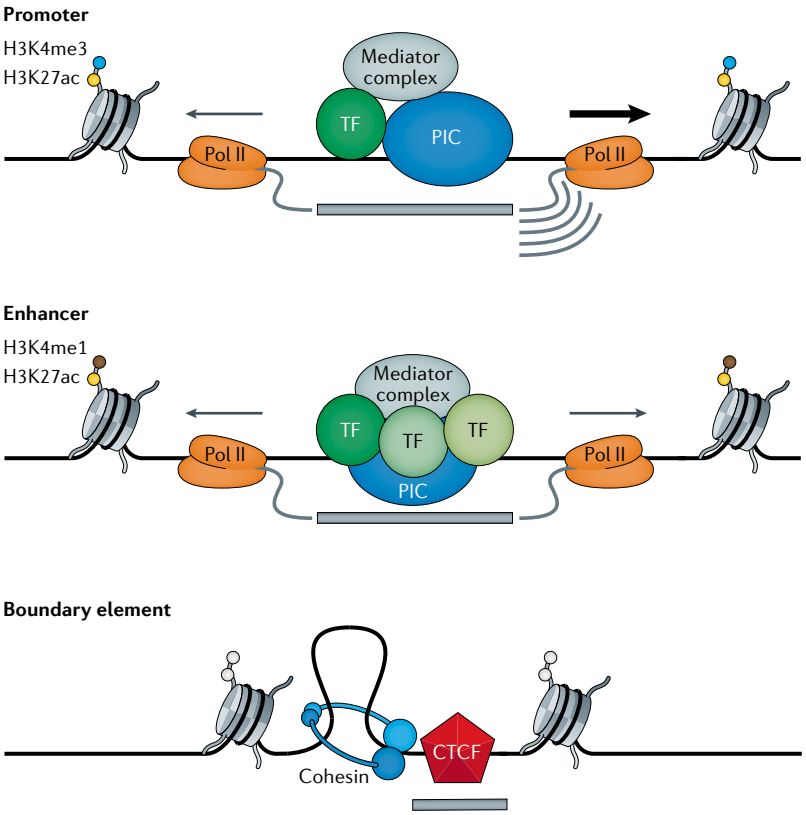


Fig. 1 | **The regulatory elements of the genome.** Overview of the three fundamental classes of regulatory elements in the genome: promoters, enhancers and boundary elements. The elements are indicated with a grey bar, and their bound factors and chromatin modifications when in an active state are shown. The transcription factors (TFs) in dark green represent ubiquitous TFs, whereas the TFs in light green represent tissue-specific and developmental stage-specific TFs. CTCF, CCCTC-binding factor; H3K4me1, histone H3 monomethylated at Lys4; H3K4me3, histone H3 trimethylated at Lys4; H3K27ac, histone H3 acetylated at Lys27; PIC, pre-initiation complex; Pol II, RNA polymerase II.

We then discuss how the regulatory elements function within their 3D genomic context, and to what extent this reflects or determines genome structure (BOX 2). Our discussion is focused mainly on euchromatin and facultative heterochromatin; the structure and function of constitutive heterochromatin have been reviewed elsewhere[14]. Although it is clear that the 3D organization of promoters, enhancers and boundary elements is critical for their function, accumulating evidence suggests that this structural context is ultimately determined by the elements themselves, their binding factors and epigenetic modifications. We thus conclude that genome structure is encoded by the sequence, linear order and activity of the regulatory elements and that many higher-order chromatin and nuclear structures are a reflection of their activity.

The regulatory elements of the genome

Active regulatory elements are located within short, nucleosome-free regions of the genome, consistent with these elements being recognized by proteins that bind DNA in a sequence-specific manner. In other words, the regulatory elements are the non-coding regions of the genome that are read and deciphered during development and differentiation. Importantly, these DNA elements are often relatively conserved in sequence and position throughout evolution, demonstrating their functional importance[15]. Although each class of regulatory elements is defined by its function in the regulation of gene expression, within each group there exists a wide range of elements, and there seems to be overlap between the groups in their functional roles[16]. The classes of regulatory elements could therefore be considered as a spectrum rather than as strict categories.

Promoters. In mammals, a core promoter (FIG. 1; TABLE 1) is defined as the minimal region of contiguous DNA sequence that is sufficient to accurately initiate basal gene expression. It includes the TSS and can extend for ~35–40 bp upstream and/or downstream of the TSS[17,18]. This small segment of DNA serves as a platform for recruiting the pre-initiation complex[19,20]. The core promoter integrates the stimulatory and repressive signals involved in regulating the recruitment of, and initiation of transcription by, RNA polymerase II. Promoters are often transcribed in both directions, although elongation usually occurs in only one direction to transcribe the associated gene. Transcripts travelling in the opposite direction are usually short, unstable and rapidly degraded by the exosome[21–23].

Despite their common purpose, the individual sequence motifs and binding sites associated with core promoters are quite diverse. The first core promoter sequence identified was the TATA box; other frequently recurring core promoter sequences have been discovered since, including the initiator element (Inr), the downstream promoter element, the downstream core element and the upstream and downstream TFIIB recognition elements[18,24]. The most common core promoter elements are simple GC-rich regions of DNA, which lie within CpG islands[25]. Different combinations of sequences within individual core promoters may in turn recruit different versions of the pre-initiation complex and its associated factors, some of which are expressed in specific cell types and developmental stages[24]. The heterogeneity of promoters is further extended by a variety of promoter proximal elements, typically located 40–250 bp upstream of the TSS and located within the same nucleosome-free region[26,27]. These compound elements may bind ubiquitous and/or tissue-specific transcription factors to modulate the activity and strength of promoters in different cellular contexts[28].

Enhancers. Individual enhancers (FIG. 1; TABLE 1) are short regions of DNA that activate transcription of the gene they regulate, independently of their position, distance or orientation relative to their target promoter[29]. Enhancers often contain multiple binding sites for tissue-specific and developmental stage-specific transcription factors. In outline, it is thought that enhancers are required to integrate the signalling and transcriptional programmes that occur throughout differentiation and development, and to communicate this information to the promoters of the genes they control[30,31]. Although some enhancers may lie within or close to their target genes, others are located far (1–1,000-kb) upstream or downstream.

Euchromatin
A relatively loosely packaged form of chromatin that is enriched in genes that are actively transcribed or poised for transcription.

Facultative heterochromatin
Regions of chromatin that are densely packaged and transcriptionally silent but may lose their condensed state and become transcriptionally active.

Constitutive heterochromatin
Regions of permanently densely packaged and transcriptionally silent chromatin that are found at specific, highly repetitive regions of the genome, such as telomeres and centromeres.

Pre-initiation complex
A large complex of proteins, including RNA polymerase II and its associated general transcription factors, which is necessary for the transcription of protein-coding genes.

The strength of individual enhancers may be related to the number and type of transcription factor-binding sites that they contain[32,33]. Bound transcription factors can in turn recruit co-activator proteins, including the mediator complex, which act on various stages of the transcription cycle to regulate gene expression[34,35]. Like promoters, enhancers also recruit RNA polymerase II and are transcribed in both directions, producing small, unstable enhancer RNAs that are rapidly degraded[36].

Enhancers vastly outnumber protein-coding genes in mammalian genomes[5,37]. It thus follows that many genes are controlled by more than one enhancer. Genes with specific expression patterns in multiple tissues and/or developmental stages are usually regulated by several distinct enhancers, which respond to transcription factors specifically expressed in the respective tissues or developmental stages[30,31,38,39]. Additionally, some genes are regulated by clusters of enhancers with similar spatio-temporal activity. Because enhancers within such clusters are often redundant, it has been suggested that the presence of multiple enhancers with similar activities confers robustness to loss-of-function mutations in individual enhancers[40–42]. It has also been proposed that clusters of enhancers can function as superenhancers with a distinct role in driving high levels of expression of critical genes that define cell identity[43]. Superenhancers have been proposed to have emergent properties resulting from synergistic cooperation between the individual elements of which they are composed[43]. Although it has been shown that individual enhancer elements in such clusters interact in higher-order structures[44–46], the nature of any functional cooperation between these elements is not yet clear. Functional studies in which components of superenhancers have been deleted have suggested that their strong effects may be explained simply by the additive effects of their individual components[32,40,47]. Further work is required, therefore, to clarify whether superenhancers are a distinct class of regulatory elements, or more simply clusters of several strong conventional enhancers[48].

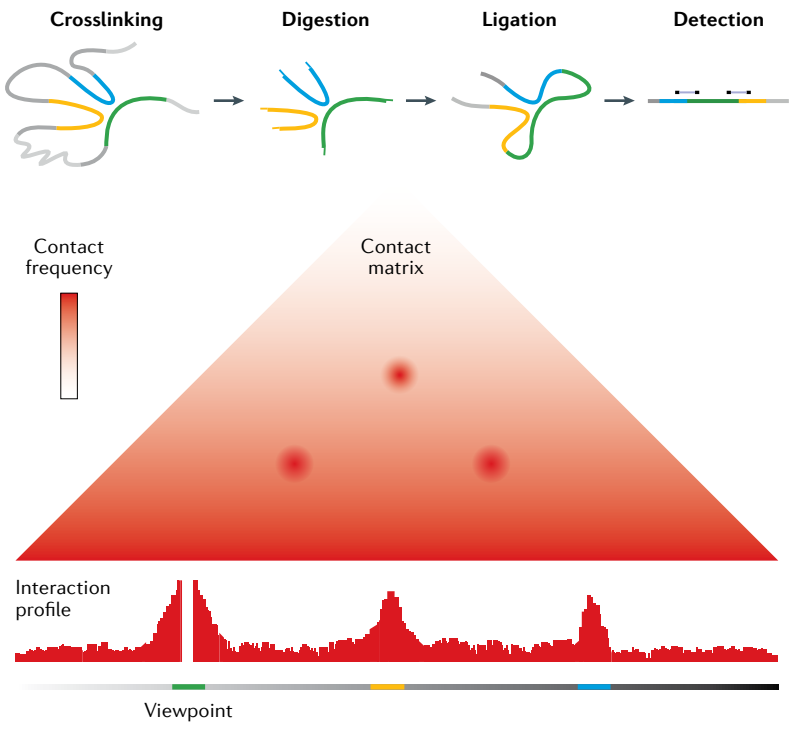
Overlap between enhancers and promoters. Although enhancers and promoters are classified as distinct entities, there is ongoing discussion about the overlap in their structure and function[16,49–51] (FIG. 1; TABLE 1). When active, both are small, nucleosome-free DNA elements that recruit large protein complexes, including transcription factors, the mediator complex and the pre-initiation complex, and both are bidirectionally transcribed. Additionally, active enhancers and promoters are marked by similar chromatin modifications. The main differences are that enhancers can act over large distances and tend to recruit tissue-specific factors, whereas promoters are usually located directly upstream of the TSS and more frequently recruit ubiquitously expressed factors. Furthermore, transcription from promoters proceeds to efficient unidirectional elongation associated with specific phosphorylation of RNA polymerase II, and high levels of trimethylation of histone H3 at Lys4 and Lys36. By contrast, enhancers direct relatively low levels of transcription, producing rapidly degraded enhancer RNAs, and are usually marked by lower levels of chromatin methylation, predominantly monomethylation of histone H3 at Lys4. These differences may, to some extent, reflect the genomic context in which the elements are located, including surrounding splicing and poly(A)

Box 1 | Approaches to study genome organization

Higher-order chromatin structures can be analysed by microscopy or assays based on proximity ligation[143], which are usually referred to as 'chromosome conformation capture (3C) techniques'[206].

The 3C methods are based on crosslinking of cells, followed by digestion of the chromatin (usually with restriction enzymes) and in situ ligation of chromatin fragments that are in close proximity in the nucleus. Quantification of the newly formed ligation junctions provides an estimate for the contact (interaction) frequency between genomic loci. The original 3C assay uses PCR to identify the contact frequency between two loci of interest in a 'one versus one' manner[206]. Subsequently, 'one versus all' methods, including 4C[202,203] and Capture-C[204,205], were developed. These methods are able to sequence all interactions with a viewpoint (bait) of interest, which are usually displayed in an interaction profile (see the figure). The development of Hi-C[81] allowed 'all versus all' investigation of chromatin interactions across the entire genome. These interactions are displayed in a contact matrix, in which the contact frequency between loci is indicated by the colour intensity at the intersection between these loci in the matrix. Because identifying interactions genome-wide requires very deep sequencing, a variety of methods to enrich the Hi-C library for interactions of interest have been developed, which produce data at higher resolution relative to sequencing depth (reviewed in[10,11,207]). Concern about potential bias in 3C approaches, in particular related to crosslinking or ligation, has led to the development of approaches without crosslinking (i3C[208]), without ligation (GAM[191] and SPRITE[197]) and without both crosslinking and ligation (DamC[209]). Importantly, these methods have generally shown good agreement with the conventional 3C approaches.

Approaches based on microscopy have the advantage that they generate data at a single cell rather than a population level and that they allow direct measurement of distances between genomic loci rather than the probability of crosslinking and ligation. Approaches based on imaging generally produce data at lower throughput and resolution compared with the 3C methods, although both throughput and resolution have increased substantially in the past few years[181,210,211]. Microscopy-based methods have different sources of potential bias, including a DNA denaturation step, but generally show good agreement with the 3C approaches[112,170,210].

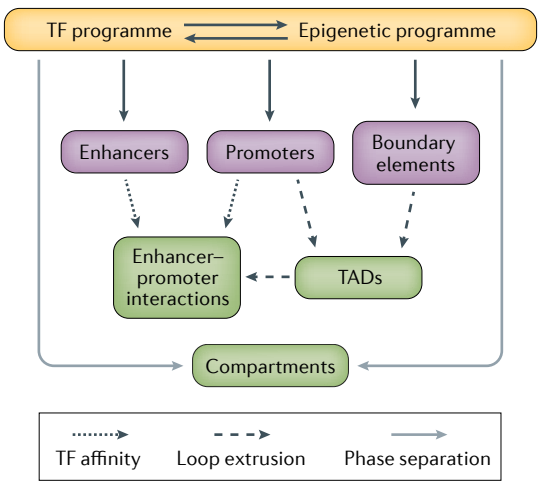


Box 2 | **The relationship between genome structure and function**

The following observations suggest that higher-order chromatin structures play an independent, primary role in regulating gene expression. The following observations suggest that the regulatory elements play a primary role in regulating gene expression and that higher-order chromatin structures are an emergent property.

- Individual chromosomes are organized in discrete chromosomal territories[76].
- Topologically associating domains (TADs) are often established early in development and are relatively invariant between cell types[168].
- Perturbing TAD structure leads to changes in gene expression[96–105].
- Forcing physical proximity between an enhancer and a promoter can result in gene activation[149,159,160].
- Some enhancer–promoter interactions seem to be established before activation of gene expression[39,105,171–174].
- Some non-membrane-bound nuclear compartments may be preformed, fixed entities[212].
- Balanced translocations are relatively common in phenotypically normal individuals[77].
- Rearrangements involving entire TADs cause no or small changes in gene expression[138].
- Compartments and subTADs differ between cell types[168] and in individual cells[95].
- Changes in TAD structure and associated changes in gene expression result from the perturbation of functional boundary elements[96–105].
- Depletion of cohesin and CCCTC-binding factor (CTCF) cause radical changes in chromatin structure but relatively little change in gene expression[62,117,119].
- In the native context of the genome, physical proximity between regulatory elements is dependent on the factors bound at these elements[147,148].
- Some enhancer–promoter interactions are formed concomitant with upregulation of gene expression[142,174,176–179].
- Many non-membrane-bound nuclear compartments are dynamic, ephemeral structures[193].
- TADs can be predicted on the basis of chromatin modifications, transcription factor (TF) binding and gene expression[200,201].

From the observations taken together, there is strong evidence for the view that higher-order chromatin structures are encoded by the regulatory elements of the genome and emerge as a reflection of their activity. We propose the following model (see the figure) for the relationship between genome structure and function. The activity of the regulatory elements of the genome is related to the TF programme and the epigenetic programme in the cell nucleus. Together, the sequence, order and activity of these regulatory elements instruct the processes underlying higher-order chromatin structures. CTCF-binding boundary elements demarcate the boundaries of cohesin-mediated loop extrusion and thereby control the positions of TAD borders; active gene promoters might act in a similar way. Within TADs, the process of loop extrusion brings enhancers and promoters into close physical proximity; this may facilitate specific enhancer–promoter interactions, which are stabilized by affinity between their bound TFs and cofactors. Dependent on their chromatin state, regions of euchromatin and heterochromatin form spatially separated compartments via a process of phase separation.



signals[52]. It has been shown that some enhancers produce long non-coding RNAs in non-coding regions of the genome, but when located within the body of a gene and surrounded by appropriate processing signals act as bone fide alternative gene promoters[53]. Similarly, it has been shown that some elements can act both as promoters for one gene and distal enhancers for another gene[54]. Overall, while we acknowledge the overlap between enhancers and promoters, we believe that the current distinction remains useful in describing the syntax of the genome.

Boundary elements. Boundary elements were first identified in *Drosophila melanogaster* (in which they are referred to as 'insulators') as genomic regions with the ability to regulate gene expression, either by preventing encroachment of heterochromatin into euchromatin and vice versa or by blocking interactions between enhancers and promoters[55,56]. In mammals, boundary elements frequently bind the zinc-finger CCCTC-binding factor (CTCF), often in association with the cohesin complex[57,58] (FIG. 1; TABLE 1). With the generation of the first high-resolution genome-wide maps of the 3D structure of mammalian genomes, it became apparent that the boundaries of structural chromatin domains are strongly enriched with CTCF-binding sites[59], mostly arranged in a convergent orientation[60]. Thus, the term 'boundary element' now usually refers to elements that have a structural role in demarcating the borders of chromatin domains. In mammalian genomes only a subset of correctly orientated CTCF-binding sites act as structural boundaries in vivo[61]. Moreover, not all structural CTCF-binding boundary elements have a functional role in regulating genome activity[62,63]. At present it is not known what distinguishes CTCF-binding sites that do or do not act as structural or functional boundaries, but their genomic context, interacting and competing proteins[64] and interactions with RNA[65,66] may all play a part.

The boundaries of structural chromatin domains are also enriched with promoters of actively transcribed genes[59,67–70], suggesting that promoters might also have a boundary function. In addition, it has been shown that when an active promoter is located between an enhancer and another, more distal, promoter, the presence of the active proximal promoter can reduce the activity of the distal gene. In some cases, it has been proposed that this effect is the result of competition between an enhancer and two or more promoters[71–73]. However, it has recently been shown by multiway chromosome conformation capture (3C) experiments and live imaging that the access of promoters to a shared enhancer is not necessarily exclusive[74,75]. It is therefore also possible that, in some cases, a proposed competition effect might be explained by the boundary function of active gene promoters.

Working together, elements with a boundary function have a critical role in regulating the specificity of interactions between enhancers and promoters. As we describe in detail in the next sections, their function is closely associated with their role in organizing the 3D architecture of the genome.

Table 1 | **Characteristics of the regulatory elements of the genome**

Characteristic	Promoter	Enhancer	Boundary element
Chromatin accessibility	+++	+++	+
Distal activation	+	+++	–
Association with CTCF	+	+	+++
Association with cohesin	+	+	+++
Boundary effect	+	?[c]	+++
Chromatin modifications			
H3K4me1	+	+++	?[a]
H3K4me2	+	+	?[a]
H3K4me3	+++	+	?[a]
H3K27ac	+++	+++	?[a]
Associated transcription factors			
Ubiquitous	+++	+	+++
Tissue-specific	+	+++	?[b]
Association with PIC	+++	+	–
Transcription			
Unidirectional	+++	–	–
Bidirectional	+	+	–

CTCF, CCCTC-binding factor; H3K4me1, histone H3 monomethylated at Lys4; H3K4me2, histone H3 dimethylated at Lys4; H3K4me3, histone H3 trimethylated at Lys4; H3K27ac, histone H3 acetylated at Lys27; PIC, pre-initiation complex. [a]At the moment it is not clear whether there are specific chromatin modifications associated with boundary elements. [b]Boundary elements are frequently bound by the ubiquitous transcription factor CTCF. It is possible that other ubiquitous and tissue-specific factors are also involved in their function; this requires further exploration. [c]It is not clear yet whether active enhancer elements could have a boundary effect.

The organization of the genome

It has been known for a long time that chromatin is not randomly distributed within the nucleus. More than a century ago, it was proposed that interphase chromosomes are organized in distinct territories in which they more frequently interact with themselves than with other chromosomes[76]. Although the organization of the genome into chromosome territories seems to be an independent mode of genome organization, which is not determined by the regulatory elements, it does not seem to be instructive for the regulation of the genome. Balanced translocations, which transfer large segments spanning several megabases from one chromosome to another, are relatively common and are usually not associated with any abnormal phenotype, unless the translocation leads to the gain, loss or rearrangement of regulatory elements surrounding the breakpoints[77,78].

Over the past two decades, developments in 3C and imaging techniques have provided a more detailed and nuanced understanding of the nuclear organization of chromosomes[8,12] (BOX 1). It is now known that during development the interphase genome becomes organized into two distinct classes of domains: compartments and topologically associating domains (TADs). It was initially thought that these domains reflect a hierarchical organization, in which compartments spanning several megabases are composed of multiple smaller TADs[79,80]. However, in the past few years it has been shown that compartments span a wide range of sizes and can be much smaller than initially thought. Moreover, accumulating evidence indicates that TADs and compartments are formed by distinct mechanisms. Rather than a hierarchical organization, these structures therefore reflect (at least) two distinct organizing principles that shape the interphase genome (FIG. 2).

Compartmentalization of the genome. Compartments were identified in the first Hi-C contact matrices as a 'chequerboard' or 'plaid' pattern spanning entire chromosomes and are formally defined as the first principal component of Hi-C data[81]. These patterns reflect the spatial separation of euchromatin and heterochromatin and distal interaction between chromatin regions bearing shared chromatin modifications (FIG. 3a). The active and inactive regions are referred to as A and B compartments, respectively. The B compartments are predominantly localized in the nuclear periphery and surrounding the nucleoli, while A compartments are in the interior of the nucleus and are sometimes associated with nuclear speckles[82]. More recent Hi-C analysis at higher resolution has revealed that A and B compartments can be further divided into at least six subcompartments comprising regions containing specific epigenetic chromatin marks and that compartments span a wide range of sizes, ranging from tens of kilobases to several megabases[60,69].

It has been proposed that the organization of chromosomes into distinct compartments is closely connected to their polymeric nature. Chromosomes can be considered as polymers composed of alternating blocks of active and inactive chromatin, with each block recruiting distinct binding factors. Polymer physics predicts that chromosomes fold so that blocks of each type cluster together, while displaying few interactions with blocks of the other type[83,84]. The resulting compartmentalization of distinct types of chromatin is thought to be driven by phase separation. A more detailed discussion of the phase separation mechanisms underlying genome compartmentalization is beyond the scope of this Review but has been given elsewhere[8,12,83–85].

Given the close link between compartmentalization and the transcriptional state of regions of chromatin, the causal and functional relationships between genome organization and transcription have been widely discussed[8,69,86,87]. Considering the recent progress in our understanding of phase separation processes underlying compartmentalization, it is possible that it is not the act of transcription per se but the epigenetic state (active versus inactive) of a region of chromatin that determines its affinity with other regions of chromatin. It therefore seems that epigenetic activity instructs the pattern of compartmentalization in the nucleus.

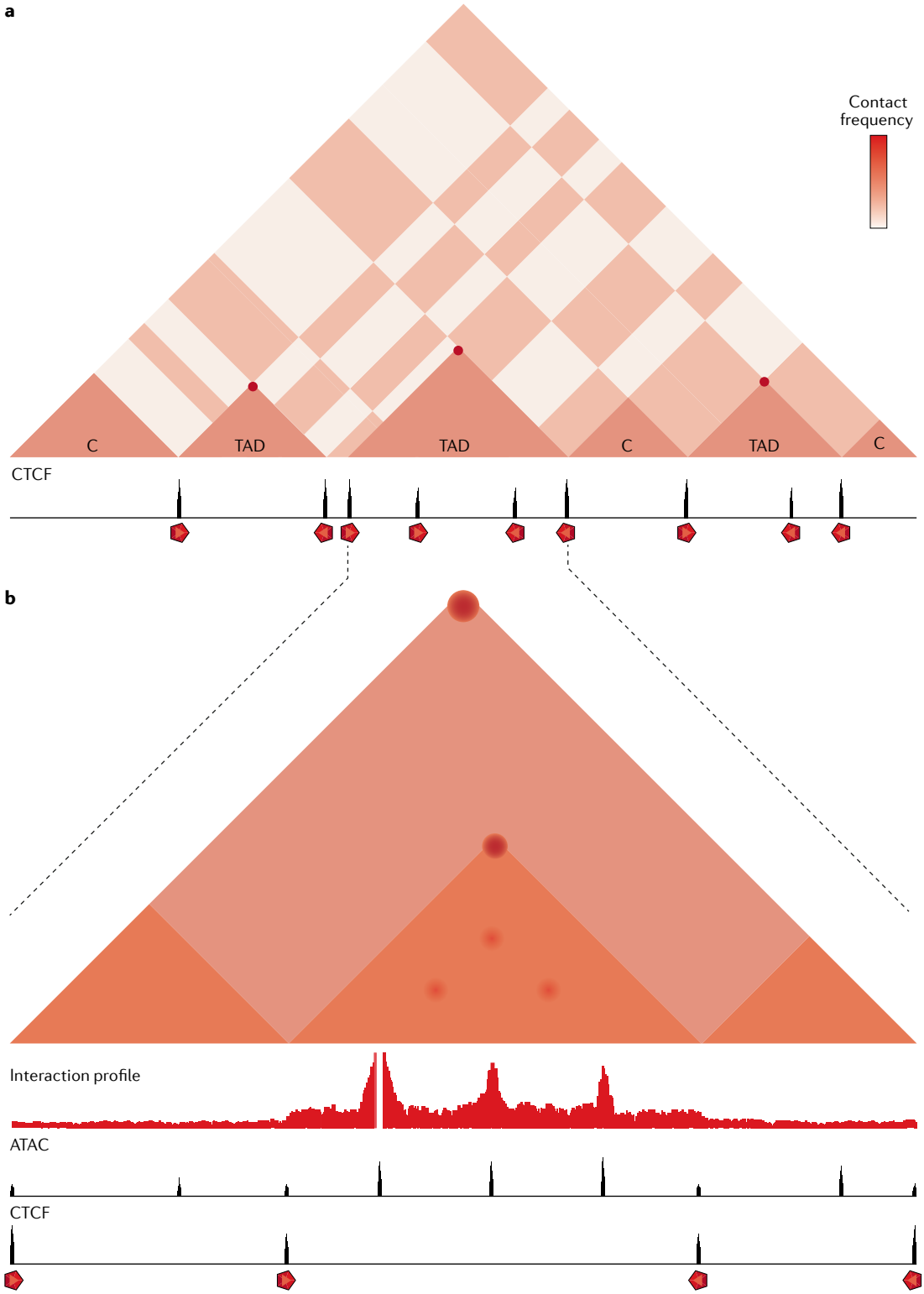
Topologically associating domains. TADs were initially identified in 3C contact matrices as chromatin domains with increased self-interactivity[59,88,89]. TADs are defined by preferential interaction of the chromatin located within them and relative depletion of interactions between chromatin located in different domains. TADs appear in contact matrices as squares or triangles along the diagonal of the matrix, depending on how the matrix is visualized, and are usually between 200 and 1,000 kb in

TATA box
Named after its conserved DNA sequence, the TATA box is a non-coding DNA sequence found in many eukaryotic core promoters that recruits the pre-initiation complex to initiate transcription.

CpG islands
Regions of the genome (~300–3,000 bp) that contain a large number of CpG dinucleotides and are associated with ~40–70% of mammalian gene promoters.

Mediator complex
A large protein complex that acts as a key transcriptional co-activator by communicating signals from transcription factors to RNA polymerase II to control its activity.

Nuclear speckles
Non-membrane-bound subdomains located in the interchromatin regions of the nucleus of mammalian cells that are enriched in splicing factors and other mRNA-processing proteins.



size[59] (FIG. 2). A number of different algorithms have been developed to identify TADs in Hi-C data[59,60,90]. Systematic TAD calling is often challenging though, as results can differ depending on the algorithm used, the parameters chosen and the quality of the experimental data[91–93]. Some TADs have corner peaks and/or stripes, which can help in their identification. However, these structures are not always present, and again their detection may depend on the depth, resolution and quality of the experimental data. Because the defining feature of a TAD is simply its self-interactivity, it is not straightforward to differentiate TADs from compartments. In combination with a lack of consistent nomenclature for the different types of chromatin domains, this has led to considerable confusion in the field. However, recent efforts to harmonize definitions have brought more clarity[12,94,95].

Phase separation
The process by which substances in a mixture become separated in two distinct phases, as occurs in a mixture of oil and water.

◀ Fig. 2 | **The organization of the genome. a** | The genome is organized into two types of domains: compartments (C) and topologically associating domains (TADs). These domains are identified in Hi-C contact matrices as regions with increased self-interactivity (triangles). Compartments and TADs reflect distinct mechanisms of genome organization and form complex, overlapping structures. Compartments reflect separation of active and inactive chromatin regions and long-range interactions between regions of the same type, visible as a chequerboard pattern. They can be distinguished as the domains whose borders overlap with this long-range chequerboard pattern. TADs are local domains resulting from loop extrusion, which can span multiple compartments (such as the left-hand and middle TADs) or can be contained within a compartment (such as the right-hand TAD). Their boundaries often overlap with convergent CCCTC-binding factor (CTCF)-binding sites (as identified by chromatin immunoprecipitation followed by sequencing), which can form relatively stable interactions and are sometimes visible as corner peaks (red spots). **b** | TADs can contain nested subTADs, which can be detected in deep, high-resolution chromosome conformation capture (3C) data. Within these domains, active regulatory elements (in open chromatin, as identified by assay for transposase-accessible chromatin using sequencing (ATAC-seq)) form specific interactions. Such interactions are difficult to identify in Hi-C data of moderate depth and resolution but can be detected in contact matrices generated from very deep Hi-C[168], Micro-C[67,68] or Tiled-C[142] data. They can also readily be identified in 'one versus all' interaction profiles from the viewpoint of a regulatory element, as generated by methods such as 4C[202,203] and Capture-C[204,205].

An important distinction between TADs and compartments is their long-range interaction patterns. As described earlier, compartments of the same type form distal interactions visible in contact matrices as an off-diagonal chequerboard pattern across and between entire chromosomes. By contrast, TADs are formed of local (less than 3 Mb) interactions along the diagonal of the contact matrix and do not form long-range contacts with other TADs. Another difference between TADs and compartments is the nature of their boundaries. In contrast to compartments, TADs have distinct boundaries, which are often bound by CTCF[59,60]. CTCF has a non-palindromic binding motif. The CTCF motifs at the boundaries of TADs are usually in a convergent orientation and form relatively strong interactions, visible as corner peaks in contact matrices[60] (FIG. 2). Numerous studies have shown that deletion or inversion of CTCF-binding sites at TAD boundaries result in disruption of TAD structures, indicating the importance of CTCF for TAD formation[96–105].

It was initially not clear how the relatively small CTCF proteins could provide insulation of much larger chromatin domains. On the basis of polymer simulations, it was proposed that a mechanism of loop extrusion could explain this and other observed properties of TAD structures[104,106]. In this process, a loop-extruding factor associates with chromatin and translocates along the fibre in opposite directions, thereby creating a progressively larger loop, until the factor is stalled at boundary elements, such as convergent CTCF-binding sites (FIG. 3b). Loop extrusion is a dynamic process, during which loops are constantly forming and falling apart in individual cells[107]. TADs should therefore not be considered as fixed structures in the nucleus. Indeed, single-cell Hi-C and chromatin tracing experiments coupled with super-resolution microscopy have shown that TAD structures vary from cell to cell[108–112]. However, averaged over millions of cells, dynamic extrusion in a region between TAD borders results in increased interactions in this region, which emerge as a TAD in contact matrices.

On the basis of its ability to translocate along chromatin and its colocalization with CTCF at domain boundaries[57,58,113–115], the multiprotein complex cohesin was suggested as a potential loop-extruding factor in the mammalian genome. In the past few years, a wealth of evidence has created strong support for a mechanism of loop extrusion dependent on cohesin and CTCF underlying TAD formation[116]. An important line of evidence is provided by genetic manipulation of cohesin, cohesin-associated factors and CTCF. It has been shown that acute depletion of the cohesin subunit RAD21 (also known as SCC1) in HCT116 cells[112,117] or HeLa cells[118] and deletion of the cohesin-associated factor NIPBL (also known as SCC2) in mouse hepatocytes[119] leads to a loss of TADs and chromatin decompaction, as predicted by loss of loop extrusion. By contrast, deletion of the cohesin release factor WAPL in HAP1 cells leads to the formation of stronger and larger TADs[120]. Acute depletion of CTCF in mouse embryonic stem cells[62] or HeLa cells[118] causes a loss of TADs but has no impact on chromatin compaction, as expected, as CTCF demarcates the extrusion boundaries but is not involved in the process of extrusion itself. Recently, the crystal structure of cohesin subunits in complex with a CTCF peptide has provided the molecular basis of the interaction between these factors. CTCF and WAPL bind the same cohesin pocket, which explains how CTCF can stabilize cohesin at TAD boundaries by preventing cohesin release mediated by WAPL[121]. Another strong line of evidence for cohesin-mediated loop extrusion has been provided by experiments involving single-molecule imaging and biochemical reconstitution[122–124]. These studies have demonstrated ATP-dependent loop extrusion by human cohesin–NIPBL complexes.

Following the discovery of TADs, analyses at higher resolution have revealed that TADs are often further organized into smaller domains. As these nested subTADs usually have domain structures and CTCF boundaries similar to those of TADs, they are likely also formed by loop extrusion. However, in comparison with TADs, subTAD boundaries exhibit weaker insulation strength and are more often cell type specific[60,94,125]. At present it is unclear what underlies these structural differences or what their functional consequences may be.

The relationship between TADs and compartments. It thus seems that the organization of the genome into compartments and TADs is mediated by two distinct mechanisms: phase separation and loop extrusion, respectively. Additional evidence for the independence of these mechanisms is provided by analysis of chromatin structures during the transition from mitosis to interphase[126–128]. 3C experiments in synchronized mammalian cell cultures have revealed that TADs emerge relatively quickly, highlighting the dynamic nature of loop extrusion, while complete re-establishment of compartmentalization is a much slower process[126,127]. Moreover, genetic manipulation of loop-extruding factors suggests that the interplay between the two independent mechanisms is antagonistic. The loss of TADs following from depletion of cohesin or NIPBL is associated with an increase in compartmentalization[117–119].

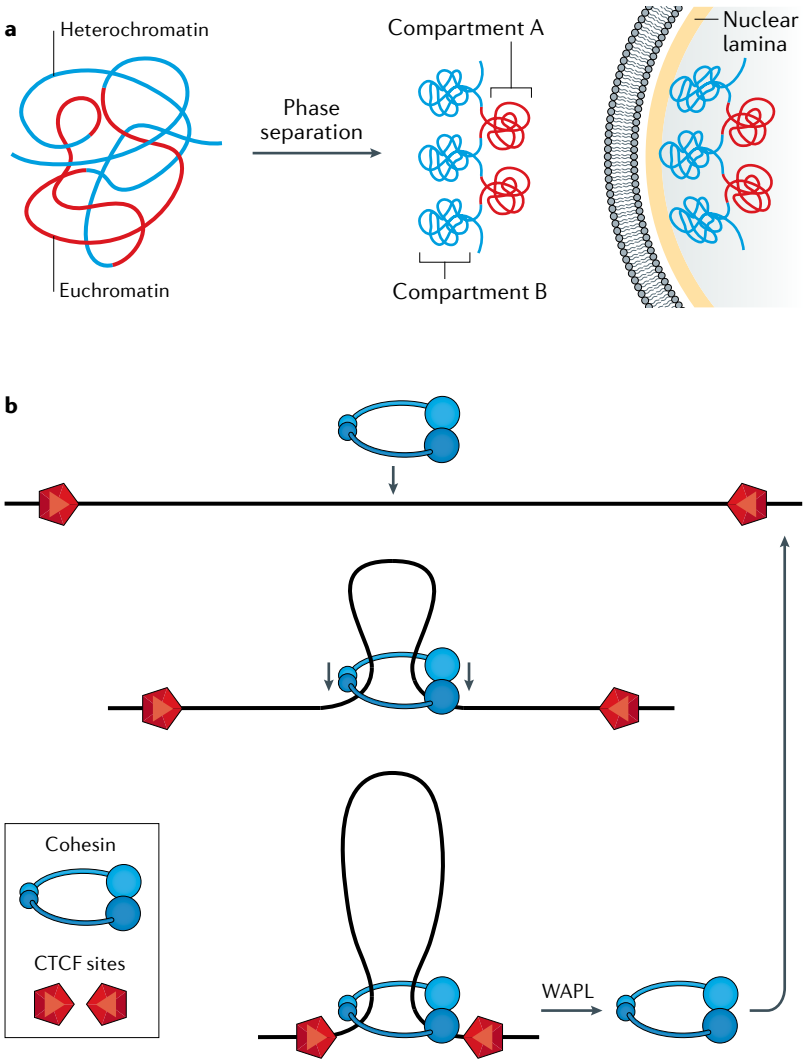


Fig. 3 | **Compartmentalization and loop extrusion. a** | Compartmentalization of the genome results from the spatial separation of euchromatin (compartment A) and heterochromatin (compartment B) and distal interactions between chromatin regions of the same type. This process is likely driven by phase separation. Heterochromatin compartments are often positioned in the nuclear periphery, whereas euchromatin compartments are usually located in the interior of the nucleus. **b** | Loop extrusion is initiated when a complex of cohesin and NIPBL (also known as SCC2; not shown) is loaded onto chromatin. As this complex translocates along the chromatin fibre, it extrudes a loop, until it is blocked by the amino terminus of CCCTC-binding factor (CTCF) at boundary elements. The cohesin complex can be released by WAPL (and its PDS5 binding partners).

Conversely, WAPL depletion results in stronger TADs and weaker compartments[118,120]. In addition, depletion of cohesin and Polycomb complexes in mouse embryonic stem cells has shown that loop extrusion by cohesin perturbs long-range interactions between regions repressed by Polycomb[129]. Together, these studies suggest that loop extrusion might be able to disrupt the intrinsic compartmentalization of the genome and pull chromatin regions into domains that are inconsistent with their chromatin state. Overall, it thus seems that the organization of the genome results from an interplay between compartmentalization and loop extrusion[130]. The architecture that emerges reflects coexisting compartments and TADs, which are not mutually exclusive and therefore result in complex overlapping and nested structures (FIG. 2).

Polycomb
The Polycomb system involves various protein complexes, including Polycomb repressive complex 1 (PRC1) and PRC2, which act as transcriptional repressors with a key role in epigenetic silencing during differentiation and development.

Regulation in the 3D genome

Regulatory elements exert their function within the context of the dynamic 3D structures into which the genome becomes organized. The relationship between these structures and genome function is a long-standing question. The organization of the genome into compartments seems to be a reflection of genome function, as it is driven by phase separation of chromosomal regions with distinct chromatin states. The structure–function relationship at the level of TADs and the enhancer–promoter interactions within them are less clear. In this section, we describe the mechanisms contributing to specific enhancer–promoter interactions and their relationship with genome function, and discuss observations addressing the potential instructive and reflective roles of genome structure in gene regulation (BOX 2).

Facilitation and insulation of enhancer–promoter communication by loop extrusion. It is generally thought that regulation of gene expression by enhancers requires close physical proximity between the enhancers and their target gene promoters[6], although there may be exceptions[131]. Loop extrusion results in the formation of chromatin structures that are thought to play an important part in the specific interaction and communication between enhancers and promoters in mammalian organisms. It has been shown that the ability of enhancers to upregulate gene expression is largely constrained to promoters located in the same TAD[132]. The insulation provided by TADs is relatively low, as the interaction frequency measured by Hi-C between loci within the same TAD is only approximately twofold higher compared with that between loci in different TADs[133,134]. This relatively low level of insulation could be explained by the fact that loop extrusion does not prevent the resulting domains from intermingling and interacting with each other. However, it is also possible that the true level of insulation is higher than measured by Hi-C due to technical limitations related to the resolution limit of Hi-C. Nonetheless, a moderate level of insulation within TADs could indicate that creating relatively insulated domains is not the only role of loop extrusion in mediating specific enhancer–promoter communication. The process of extrusion itself could also facilitate interactions between regulatory elements by bringing them into direct molecular contact. This facilitating role of loop extrusion is especially evident for regulatory elements close to TAD boundaries, which often show increased interactions throughout the entire domain, visible as stripes in contact matrices[135,136]. Facilitation of contacts by loop extrusion occurs only within TADs, as extrusion is blocked by the boundary elements of the TADs, thereby creating insulation between regulatory elements in neighbouring domains.

The process of loop extrusion may therefore contribute to both genome structure and genome function by facilitating and insulating communication between enhancers and promoters. However, the boundary elements of the genome ultimately dictate which regulatory

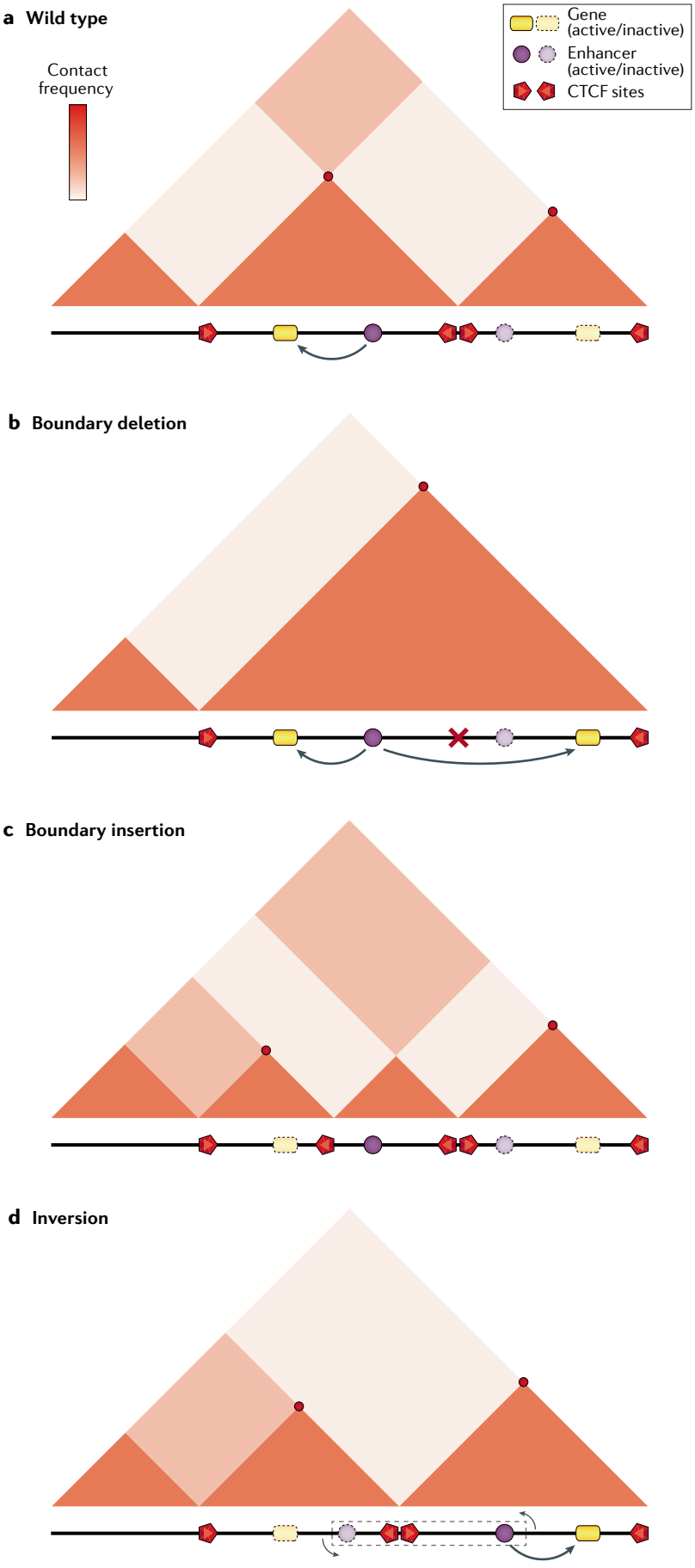


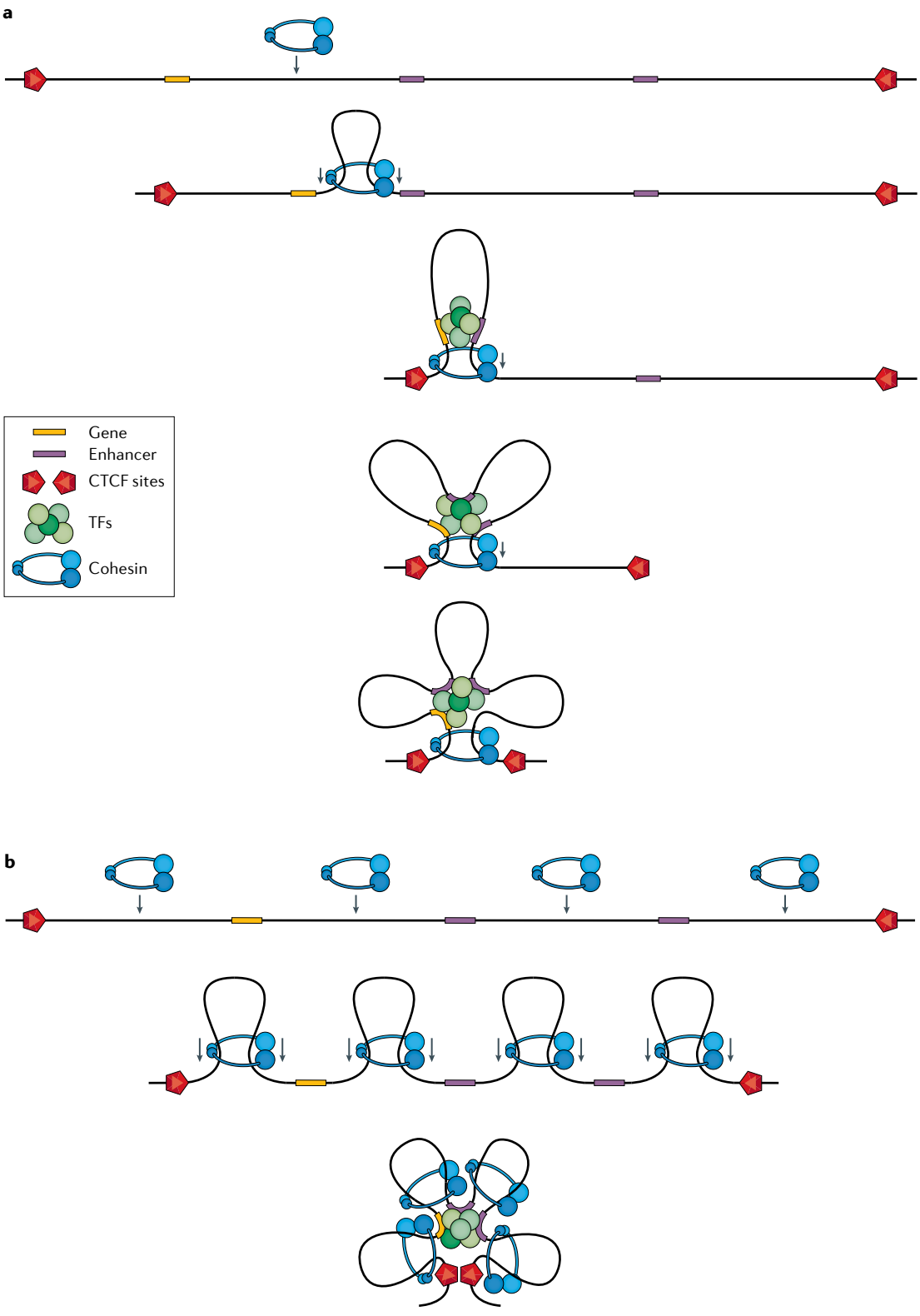
Fig. 4 | **The relationship between the linear order of the regulatory elements and the organization of the genome. a** | In a hypothetical region of the genome, an active enhancer activates an upstream gene in the middle topologically associating domain (TAD). **b** | Deletion of the downstream boundary of the middle TAD (red cross) causes two TADs to merge; the active enhancer now also activates the downstream gene. **c** | Insertion of a new boundary element within the middle TAD causes the formation of an extra TAD, with the active enhancer and the upstream gene now located in two separate TADs. The enhancer can no longer interact with and activate the upstream gene promoter. **d** | Inversion of the region highlighted by the dashed box repositions the active enhancer in a new TAD. It can no longer interact with the upstream gene promoter, but instead activates the downstream promoter. Note that in each scenario the changes in activity within the domains alter their long-range compartmentalization. CTCF, CCCTC-binding factor.

interactions are formed, as they determine the regions where contacts are facilitated and insulated. Studies involving targeted perturbations of boundary elements combined with 3C experiments have provided ample evidence for the primary role of boundary elements in demarcating TAD structures and modifying the specificity of enhancer–promoter interactions[96–105] (FIG. 4). An example of the facilitating role of loop extrusion is provided by analyses of the sonic hedgehog gene (*Shh*) locus. Combined deletion of multiple CTCF-binding sites at the TAD boundary of the *Shh* locus results in weakened interactions between the *Shh* promoter and its distal ZRS enhancer and a ~50% decrease in *Shh* expression in mouse limb buds[105]. A study of the α-globin gene locus has provided an example of the insulating role of TADs. Deletion of two CTCF-binding sites at the boundary of the α-globin gene subTAD causes ectopic interactions between the α-globin gene enhancers and upstream genes and upregulation of their expression by more than tenfold in mouse erythroid cells[98].

However, not all perturbations of CTCF-bound elements result in large changes in gene expression. It is not completely clear why some CTCF boundaries have stronger functional roles in controlling gene activity than others. However, this could be at least partly explained by genomic context and potential redundancy. Studies of the *Shh* and α-globin gene loci have highlighted that functionally important TAD boundaries often comprise multiple CTCF-binding sites to create a robust border. At the α-globin gene locus, strong effects on the activity of the upstream genes were observed on deletion of both CTCF-binding sites of the boundary, whereas deletion of one of them caused only a modest increase in expression of the upstream genes[98]. Similarly, an independent study of the *Shh* locus showed that deletion of individual CTCF-binding sites of the TAD boundaries had very minor effects on the expression pattern of *Shh*[137].

The importance of boundary elements and their relative position for mediating specific enhancer–promoter communication has also been studied by the characterization of naturally occurring and engineered structural variations in regions spanning domain boundaries. For example, it has been demonstrated

that deletions, inversions or duplications that alter the structure of the *WNT6/IHH/EPHA4/PAX3* locus can disrupt regulatory interactions and cause human limb malformations[99]. In this study, it was shown in both patient-derived fibroblasts and genetically engineered mice that the disease-relevant structural changes cause ectopic interactions between a cluster of limb enhancers, which is normally associated with *Epha4*, and other gene promoters in the locus, resulting in ectopic limb expression of these genes. Importantly, this rewiring of regulatory interactions occurred only when the structural variant disrupted a CTCF boundary. These and other experiments have demonstrated that the impact of genome rearrangements on gene misexpression and



the severity of disease phenotypes is dependent on the extent to which boundaries are inverted and repositioned relative to the enhancers and promoters whose interactions they control[99,138–141] (FIG. 4). It thus seems that gain, loss or rearrangement of the fundamental boundary elements underlie subsequent changes in chromatin structure, sometimes also leading to changes in nuclear localization, and associated changes in gene expression. This evidence provides strong support for the view that the organization of the genome represents a property emerging from the order, orientation and activity of the regulatory elements.

Mechanisms underlying enhancer–promoter interactions. In 3C contact matrices of sufficient depth and resolution, interactions between enhancers and promoters are detected as enriched contacts above the baseline of domains formed by loop extrusion and compartmentalization[67,68,142] (FIG. 2). This higher contact frequency indicates that additional mechanisms contribute to the formation of enhancer–promoter interactions.

Physical interactions between regulatory elements were described long before the detailed 3D organization of the interphase genome was known[143,144]. The mouse β-globin gene locus was among the first mammalian gene clusters in which relatively long-range enhancer–promoter interactions were characterized in detail[145,146]. It was suggested that these interactions are driven by affinity between active regulatory elements, mediated by transcription factors bound to these elements. This theory was later confirmed by several studies that demonstrated that the enhancer–promoter interactions at the β-globin gene locus are dependent on the erythroid transcription factors GATA1 (REF.[147]) and KLF1 (REF.[148]) and the more widely expressed 'looping' factor LDB1 (REFS[149–152]). Several other transcriptional cofactors with key roles in mediating enhancer–promoter interactions have been described[153], including YY1 (REFS[154,155]) and the mediator complex[125,156], although the structural contribution of the mediator complex to enhancer–promoter communication has recently been disputed[157,158]. Overall, it seems that protein–protein interactions between the tissue-specific and ubiquitous factors bound at regulatory elements play an important role in mediating their interactions.

It has been shown that engineering forced contacts between an enhancer and a promoter can result in gene activation in cell lines[149,159,160], which could be interpreted as evidence for an instructive role of genome structure. However, in the native context, the structure of the genome seems to be dependent on the factors bound at promoters, enhancers and boundary elements, implying that genome structure is reflective of the sequence, activity and linear order of the regulatory elements.

Enhancers and promoters often form interactions with multiple other regulatory elements. It has been shown that these elements can form complex, higher-order, hub-like structures, in which multiple enhancers and promoters colocalize simultaneously[44–46]. Although it has been suggested that there is a degree of inherent specificity between enhancers and promoters mediated by the compatibility between their bound factors[28,161], it is likely that enhancer–promoter communication would not be sufficiently specific if it were dependent on the affinity between bound transcription factors alone. It has been shown that on deletion of a CTCF boundary upstream of the α-globin gene cluster, gene promoters from the upstream domain join the hub containing the α-globin genes and their enhancers and become highly upregulated in this active environment[74]. This observation is in line with the studies described above, which have shown that disruption of loop extrusion boundaries can result in promiscuous enhancer activity.

Loop extrusion mediated by cohesin and halted at boundary elements therefore likely contributes to specificity between affinity-mediated interactions between regulatory elements. It is possible that cohesin also plays a more direct role in mediating enhancer–promoter interactions. It has been shown that a subset of cohesin molecules that do not overlap with CTCF colocalize with the mediator complex[156] and other transcription factors[162]. This finding could indicate that cohesin is transiently halted at regulatory elements bound by these factors during the process of loop extrusion, which could potentially result in strengthening of these interactions. Another interpretation of their colocalization is that cohesin might be loaded at these active elements.

Despite extensive investigation, the precise role of loop extrusion in regulating enhancer–promoter interactions and gene activity is not clear. In contrast to the strong effects of perturbations of some TAD borders on the activity of nearby genes, genome-wide perturbations of both cohesin and CTCF have unexpectedly modest effects on gene expression, despite causing an almost complete loss of TAD structures[62,117,119]. One suggested explanation for these modest effects is that cohesin and CTCF are especially important for genes regulated by distal enhancers[117], as regulatory elements separated by large genomic distances could be more dependent on loop extrusion to come into close proximity. It is also possible that loop extrusion is more important for the establishment of enhancer–promoter

◀ Fig. 5 | **Formation of enhancer–promoter interactions by loop extrusion and affinity.** Shown is a schematic overview of how a combination of loop extrusion and affinity between proteins bound at regulatory elements could contribute to the formation of specific enhancer–promoter interactions. As the cohesin complex translocates along the chromatin fibre and extrudes a loop, it brings regulatory elements into close proximity. The cohesin complex is halted when it approaches the amino terminus of CCCTC-binding factor (CTCF) located at boundary elements, which explains the formation of interactions between CTCF-binding sites at domain boundaries and enriched intradomain interactions. Within these domains, higher-order hub structures comprising interactions between multiple enhancers and promoters are formed. It is likely that these structures result from a combination of loop extrusion and protein–protein interactions between the factors bound at enhancers and promoters, although the precise contribution of both processes is still unclear. As shown in panel **a**, it is possible that contacts between regulatory elements are initially facilitated by loop extrusion and subsequently stabilized by affinity between the proteins bound at these elements. As illustrated in panel **b**, it is also possible that loop extrusion has a more direct role in mediating enhancer–promoter interactions if the transcription factors (TFs) and cofactors bound at enhancers and promoters transiently halt cohesin translocation. For clarity a single cohesin complex is shown in panel **a** and multiple cohesin complexes are shown in panel **b**. At the moment it is not clear at what location and with what frequency cohesin is loaded onto chromatin. However, this is not critical for the models; the model in panel **a** could also work with multiple simultaneous loading events, whereas the model in panel **b** could also work with sequential loading events.

interactions than for their maintenance. In support of this hypothesis, it has been shown that cohesin depletion affects expression of inducible genes in mouse myeloid cells triggered by inflammatory stimuli[163]. Regardless, the observation that drastic changes in genome structure caused by depletion of cohesin and CTCF are not associated with catastrophic changes in gene expression argues against a major instructive role of these structures in regulating gene expression.

Despite outstanding questions, the picture that emerges is that both loop extrusion directed by boundary elements and affinity-mediated contacts between enhancers and promoters contribute to regulatory chromatin interactions. This implies that the formation of specific chromatin structures is dependent on the activity of the fundamental regulatory elements of the genome. It is possible that loop extrusion functions to bring enhancers and promoters located between sets of boundary elements into contact and that such contacts are subsequently stabilized by protein–protein interactions between the factors bound at these elements (FIG. 5). Such a model is consistent with polymer simulations[130,164,165].

The establishment of genome architecture during differentiation and development. The relationship between genome structure and function can also be addressed by examining how both change during differentiation and development[166]. Although some TAD structures change during differentiation, TAD borders are overall relatively stable[125,167–170]. By contrast, enhancer–promoter interactions are often specific for the cell types in which the gene they control is expressed. However, it is unclear when specific interactions between enhancers and promoters within TADs are formed and what their precise relationship is with gene activation. There are several examples of loci in which interactions between enhancers and promoters seem to be established before gene expression[39,105,171–174]. It has been suggested that such preconfigured loci might function to prime genes for transcriptional activation[175], thus arguing for a primary instructive role of genome structure for gene regulation. However, there are also several cases of de novo formation of enhancer–promoter interactions during differentiation, which are more directly associated with gene activation[174,176–179]. For example, analysis of chromatin interactions at several erythroid gene loci, including the α-globin gene locus, during mouse erythroid differentiation at high temporal and spatial resolution, showed that interactions between enhancers and promoters are already formed before maximum gene activation, but at a relatively low frequency. During differentiation, as the activity of the regulatory elements increases, these enhancer–promoter interactions gradually strengthen, which correlates with a progressive increase in gene expression[142]. This finding argues that the activity of the regulatory elements is instructive for specific enhancer–promoter interactions.

Although there might be different ways in which enhancer–promoter interactions are related to gene activation, it is also possible that some of the differences in the models reflect the point in differentiation at which they have been analysed and the resolution of the analysis. It will be important to analyse more gene loci at high spatial and temporal resolution to determine exactly when and how potentially preconfigured structures are formed.

Overall, the temporal pattern of the formation of TADs and enhancer–promoter interactions during development seems to match the activity of the regulatory elements that underlie the formation of these structures. The tissue-invariant factors occupying boundary elements contribute to the formation of TADs, which are relatively stable during development. The tissue-specific transcription factors that bind to enhancers and promoters mediate their specific interactions, which occur only in the tissues in which these factors are expressed and the regulatory elements are active.

Non-membrane-bound nuclear bodies. As genomic analyses are revealing more detail about the higher-order organization of the genome, recent developments in microscopy are starting to visualize dynamic nuclear structures at a molecular level[112,180–185]. These include non-membrane-bound subcompartments, which are regions of chromatin where specific combinations of DNA, RNA and proteins interact at increased concentrations with respect to other regions of the nucleus. These components are recruited to undertake various nuclear processes, including DNA replication, DNA damage repair, transcription, repression and RNA processing. As part of these processes, it is thought that the relevant molecules may interact and concentrate via high-affinity oligomerization or by forming numerous weak multivalent interactions associated with phase separation[186–189]. Such aggregative processes may bring together widely distributed regions of the genome with a similar function to form nuclear subcompartments for specific processes. For example, the A and B compartments preferentially associate with other compartments of similar identities with respect to chromatin modifications and gene expression[81]. Within A compartments, more specific interactions reflect shared localization of co-regulated genes in nuclear subcompartments[190,191], such as transcriptional hubs[192,193] (previously referred to as 'transcription factories'[194]) and nuclear speckles[195–197], whereas silenced heterochromatic regions in B compartments colocalize at the periphery of nucleoli[197] or at the nuclear lamina[198]. Furthermore, regions targeted by the Polycomb repressive complexes form contacts within Polycomb bodies[199]. At this point it is still speculative, but it is possible that (some of) these structures result from self-assembly into dynamic molecular condensates, in which high concentrations of the relevant molecules may allow increased efficiency of the processes occurring.

Conclusion and future perspectives

We have discussed how many aspects of the 3D structure of the genome emerge as a consequence of the distribution of promoters, enhancers and boundary elements and the processes they control, particularly transcriptional and epigenetic regulation and the translocation of cohesin along chromatin. It is much less clear that the higher-order structure of chromatin and the nuclear

Polycomb bodies
Foci of Polycomb group proteins in the nucleus that are involved in both genome organization and repression of gene expression.

Molecular condensates
Non-membrane-bound subcompartments in the nucleoplasm or cytoplasm that are strongly enriched in or depleted of specific proteins or nucleic acids.

substructures that emerge play an instructive role in gene regulation over and above what is dictated by the regulatory elements (BOX 2).

As discussed, the organization of the genome into A and B compartments seems to result from self-association and phase separation of similarly modified chromatin. This process is determined by DNA-binding proteins and their activities, which are delimited by the regulatory elements. As yet there is no way to prevent the formation of these self-assembling compartments without perturbing the nuclear processes by which they form and therefore no simple experiment to test their additional functional role, if any. The organization of chromosomes into TADs and subTADs is most readily explained by loop extrusion delimited by boundary elements, which creates domains of interactions relevant for gene regulation. Abrogation of these higher-order conformations by removal of either cohesin or CTCF seems to have a surprisingly small effect on gene expression. This finding suggests that these structures may be of relatively little functional importance in themselves but rather provide a very informative reflection of the activity and organization of regulatory elements. Supporting this view, it is of interest that it has recently been suggested that 3D genome structure, as observed in 3C matrices, can be accurately predicted only on the basis of the distribution of regulatory elements and chromatin modifications by machine learning[200,201].

In future work, it will be important to investigate in greater detail how compartmentalization, loop extrusion, enhancer–promoter interactions and transcription occur in real time in single cells. Furthermore, it will be of interest to get a better understanding of the biochemistry and dynamics of the proteins interacting with the regulatory elements driving these processes within subnuclear compartments and molecular condensates. We anticipate that integration of super-resolution imaging and single-cell sequencing approaches will provide answers to these and other outstanding questions in the field, which will lead to a better understanding of the principles underlying gene regulation, genome organization and their relationship with each other.

Published online: 24 November 2020

1. Cavalli, G. & Heard, E. Advances in epigenetics link genetics to the environment and disease. *Nature* **571**, 489–499 (2019).
2. Pang, B. & Snyder, M. P. Systematic identification of silencers in human cells. *Nat. Genet.* **52**, 254–263 (2020).
3. Ngan, C. Y. et al. Chromatin interaction analyses elucidate the roles of PRC2-bound silencers in mouse development. *Nat. Genet.* **52**, 264–272 (2020).
4. Kellis, M. et al. Defining functional DNA elements in the human genome. *Proc. Natl Acad. Sci. USA* **111**, 6131 (2014).
5. Dunham, I. et al. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
6. Furlong, E. E. M. & Levine, M. Developmental enhancers and chromosome topology. *Science* **361**, 1341–1345 (2018).
7. Schoenfelder, S. & Fraser, P. Long-range enhancer–promoter contacts in gene expression control. *Nat. Rev. Genet.* **20**, 437–455 (2019).
8. Rowley, M. J. & Corces, V. G. Organizational principles of 3D genome architecture. *Nat. Rev. Genet.* **19**, 789–800 (2018).
9. Stadhouders, R., Filion, G. J. & Graf, T. Transcription factors and 3D genome conformation in cell-fate decisions. *Nature* **569**, 345–354 (2019).
10. Kempfer, R. & Pombo, A. Methods for mapping 3D chromosome architecture. *Nat. Rev. Genet.* **21**, 207–226 (2020).
11. McCord, R. P., Kaplan, N. & Giorgetti, L. Chromosome conformation capture and beyond: toward an integrative view of chromosome structure and function. *Mol. Cell* **77**, 688–708 (2020).
12. Mirny, L. A., Imakaev, M. & Abdennur, N. Two major mechanisms of chromosome organization. *Curr. Opin. Cell Biol.* **58**, 142–152 (2019).
13. Robson, M. I., Ringel, A. R. & Mundlos, S. Regulatory landscaping: how enhancer-promoter communication is sculpted in 3D. *Mol. Cell* **74**, 1110–1122 (2019).
14. Janssen, A., Colmenares, S. U. & Karpen, G. H. Heterochromatin: guardian of the genome. *Annu. Rev. Cell Dev. Biol.* **34**, 265–288 (2018).
15. Berthelot, C., Villar, D., Horvath, J. E., Odom, D. T. & Flicek, P. Complexity and conservation of regulatory landscapes underlie evolutionary resilience of mammalian gene expression. *Nat. Ecol. Evol.* **2**, 152–163 (2018).
16. Andersson, R. & Sandelin, A. Determinants of enhancer and promoter activities of regulatory elements. *Nat. Rev. Genet.* **21**, 71–87 (2020).
17. Smale, S. T. & Kadonaga, J. T. The RNA polymerase II core promoter. *Annu. Rev. Biochem.* **72**, 449–479 (2003).
18. Baumann, M., Pontiller, J. & Ernst, W. Structure and basal transcription complex of RNA polymerase II core promoters in the mammalian genome: an overview. *Mol. Biotechnol.* **45**, 241–247 (2010).
19. Sainsbury, S., Bernecky, C. & Cramer, P. Structural basis of transcription initiation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* **16**, 129–143 (2015).
20. Cramer, P. Organization and regulation of gene transcription. *Nature* **573**, 45–54 (2019).
21. Core, L. J., Waterfall, J. J. & Lis, J. T. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**, 1845 (2008).
22. Seila, A. C. et al. Divergent transcription from active promoters. *Science* **322**, 1849 (2008).
23. Preker, P. et al. RNA exosome depletion reveals transcription upstream of active human promoters. *Science* **322**, 1851 (2008).
24. Smale, S. T. Core promoters: active contributors to combinatorial gene regulation. *Genes Dev.* **15**, 2503–2508 (2001).
25. Cooper, S. J., Trinklein, N. D., Anton, E. D., Nguyen, L. & Myers, R. M. Comprehensive analysis of transcriptional promoter structure and function in 1% of the human genome. *Genome Res.* **16**, 1–10 (2006).
26. Su, W., Jackson, S., Tjian, R. & Echols, H. DNA looping between sites for transcriptional activation: self-association of DNA-bound Sp1. *Genes Dev.* **5**, 820–826 (1991).
27. Calhoun, V. C., Stathopoulos, A. & Levine, M. Promoter-proximal tethering elements regulate enhancer-promoter specificity in the Drosophila Antennapedia complex. *Proc. Natl Acad. Sci. USA* **99**, 9243 (2002).
28. Zabidi, M. A. et al. Enhancer–core-promoter specificity separates developmental and housekeeping gene regulation. *Nature* **518**, 556–559 (2015).
29. Banerji, J., Rusconi, S. & Schaffner, W. Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell* **27**, 299–308 (1981).
30. Long, H. K., Prescott, S. L. & Wysocka, J. Ever-changing landscapes: transcriptional enhancers in development and evolution. *Cell* **167**, 1170–1187 (2016).
31. Field, A. & Adelman, K. Evaluating enhancer function and transcription. Annu. Rev. Biochem. https://doi.org/10.1146/annurev-biochem-011420-095916 (2020).
32. Hay, D. et al. Genetic dissection of the α-globin super-enhancer in vivo. *Nat. Genet.* **48**, 895–903 (2016).
33. Siersbæk, R. et al. Transcription factor cooperativity in early adipogenic hotspots and super-enhancers. *Cell Rep.* **7**, 1443–1455 (2014).
34. Allen, B. L. & Taatjes, D. J. The mediator complex: a central integrator of transcription. *Nat. Rev. Mol. Cell Biol.* **16**, 155–166 (2015).
35. Soutourina, J. Transcription regulation by the mediator complex. *Nat. Rev. Mol. Cell Biol.* **19**, 262–274 (2018).
36. Andersson, R. et al. An atlas of active enhancers across human cell types and tissues. *Nature* **507**, 455–461 (2014).
37. Shen, Y. et al. A map of the cis-regulatory sequences in the mouse genome. *Nature* **488**, 116–120 (2012).
38. Anderson, E. & Hill, R. E. Long range regulation of the sonic hedgehog gene. *Curr. Opin. Genet. Dev.* **27**, 54–59 (2014).
39. Montavon, T. et al. A regulatory archipelago controls Hox genes transcription in digits. *Cell* **147**, 1132–1145 (2011).
40. Osterwalder, M. et al. Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature* **554**, 239–243 (2018).
41. Perry, M. W., Boettiger, A. N., Bothma, J. P. & Levine, M. Shadow enhancers foster robustness of Drosophila gastrulation. *Curr. Biol.* **20**, 1562–1567 (2010).
42. Frankel, N. et al. Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature* **466**, 490–493 (2010).
43. Whyte, W. A. et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
44. Oudelaar, A. M. et al. Single-allele chromatin interactions identify regulatory hubs in dynamic compartmentalized domains. *Nat. Genet.* **50**, 1744–1751 (2018).
45. Allahyar, A. et al. Enhancer hubs and loop collisions identified from single-allele topologies. *Nat. Genet.* **50**, 1151–1160 (2018).
46. Ing-Simmons, E. et al. Spatial enhancer clustering and regulation of enhancer-proximal genes by cohesin. *Genome Res.* **25**, 504–513 (2015).
47. Moorthy, S. D. et al. Enhancers and super-enhancers have an equivalent regulatory role in embryonic stem cells through regulation of single or multiple genes. *Genome Res.* **27**, 246–258 (2017).
48. Pott, S. & Lieb, J. D. What are super-enhancers? *Nat. Genet.* **47**, 8–12 (2015).
49. Henriques, T. et al. Widespread transcriptional pausing and elongation control at enhancers. *Genes Dev.* **32**, 26–41 (2018).
50. Mikhaylichenko, O. et al. The degree of enhancer or promoter activity is reflected by the levels and directionality of eRNA transcription. *Genes Dev.* **32**, 42–57 (2018).
51. Nguyen, T. A. et al. High-throughput functional comparison of promoter and enhancer activities. *Genome Res.* **26**, 1023–1033 (2016).
52. Almada, A. E., Wu, X., Kriz, A. J., Burge, C. B. & Sharp, P. A. Promoter directionality is controlled by U1 snRNP and polyadenylation signals. *Nature* **499**, 360–363 (2013).
53. Kowalczyk, M. S. et al. Intragenic enhancers act as alternative promoters. *Mol. Cell* **45**, 447–458 (2012).

54. Dao, L. T. M. et al. Genome-wide characterization of mammalian promoters with distal enhancer functions. *Nat. Genet.* **49**, 1073–1081 (2017).
55. Bell, A. C., West, A. G. & Felsenfeld, G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98**, 387–396 (1999).
56. Kellum, R. & Schedl, P. A position-effect assay for boundaries of higher order chromosomal domains. *Cell* **64**, 941–950 (1991).
57. Parelho, V. et al. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132**, 422–433 (2008).
58. Wendt, K. S. et al. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451**, 796–801 (2008).
59. Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
60. Rao, S. S. P. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).
61. Merkenschlager, M. & Nora, E. P. CTCF and cohesin in genome folding and transcriptional gene regulation. *Annu. Rev. Genomics Hum. Genet.* **17**, 17–43 (2016).
62. Nora, E. P. et al. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* **169**, 930–944.e922 (2017).
63. Huang, H. et al. CTCF mediates dosage and sequence-context-dependent transcriptional insulation through formation of local chromatin domains. Preprint at *bioRxiv* https://doi.org/10.1101/2020.07.07.192526 (2020).
64. Kaaij, L. J. T., Mohn, F., van der Weide, R. H., de Wit, E. & Bühler, M. The ChAHP complex counteracts chromatin looping at CTCF sites that emerged from SINE expansions in mouse. *Cell* **178**, 1437–1451. e1414 (2019).
65. Hansen, A. S. et al. Distinct classes of chromatin loops revealed by deletion of an RNA-binding region in CTCF. *Mol. Cell* **76**, 395–411.e313 (2019).
66. Saldaña-Meyer, R. et al. RNA interactions are essential for CTCF-mediated genome organization. *Mol. Cell* **76**, 412–422.e415 (2019).
67. Hsieh, T.-H. S. et al. Resolving the 3D landscape of transcription-linked mammalian chromatin folding. *Mol. Cell* **78**, 539–553.e8 (2020).
68. Krietenstein, N. et al. Ultrastructural details of mammalian chromosome architecture. *Mol. Cell* **78**, 554–565.e7, e557 (2020).
69. Rowley, M. J. et al. Evolutionarily conserved principles predict 3D chromatin organization. *Mol. Cell* **67**, 837–852.e837 (2017).
70. Harrold, C. L. et al. A functional overlap between actively transcribed genes and chromatin boundary elements. Preprint at *bioRxiv* https://doi.org/10.1101/2020.07.01.182089 (2020).
71. De Gobbi, M. et al. A regulatory SNP causes a human genetic disease by creating a new transcriptional promoter. *Science* **312**, 1215–1217 (2006).
72. Cho, S. W. et al. Promoter of lncRNA Gene PVT1 Is a tumor-suppressor DNA boundary element. *Cell* **173**, 1398–1412.e1322 (2018).
73. Enver, T. et al. Developmental regulation of human fetal-to-adult globin gene switching in transgenic mice. *Nature* **344**, 309–313 (1990).
74. Oudelaar, A. M. et al. A revised model for promoter competition based on multi-way chromatin interactions at the α-globin locus. *Nat. Commun.* **10**, 1–8 (2019).
75. Fukaya, T., Lim, B. & Levine, M. Enhancer control of transcriptional bursting. *Cell* **166**, 358–368 (2016).
76. Cremer, T. & Cremer, C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat. Rev. Genet.* **2**, 292–301 (2001).
77. Baptista, J. et al. Molecular cytogenetic analyses of breakpoints in apparently balanced reciprocal translocations carried by phenotypically normal individuals. *Eur. J. Hum. Genet.* **13**, 1205–1212 (2005).
78. Ghavi-Helm, Y. et al. Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression. *Nat. Genet.* **51**, 1272–1282 (2019).
79. Bonev, B. & Cavalli, G. Organization and function of the 3D genome. *Nat. Rev. Genet.* **17**, 772–772 (2016).
80. Gibcus, J. H. & Dekker, J. The hierarchy of the 3D genome. *Mol. Cell* **49**, 773–782 (2013).
81. Lieberman-Aiden, E. et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).
82. Bickmore, W. A. & van Steensel, B. Genome architecture: domain organization of interphase chromosomes. *Cell* **152**, 1270–1284 (2013).
83. Esposito, A. et al. Models of polymer physics for the architecture of the cell nucleus. *WIREs Syst. Biol. Med.* **11**, e1444 (2019).
84. Hildebrand, E. M. & Dekker, J. Mechanisms and functions of chromosome compartmentalization. *Trends Biochem. Sci.* **45**, 385–396 (2020).
85. Feodorova, Y., Falk, M., Mirny, L. A. & Solovei, I. Viewing nuclear architecture through the eyes of nocturnal mammals. *Trends Cell Biol.* **30**, 276–289 (2020).
86. van Steensel, B. & Furlong, E. E. M. The role of transcription in shaping the spatial organization of the genome. *Nat. Rev. Mol. Cell Biol.* **20**, 290 (2019).
87. Rada-Iglesias, A., Grosveld, F. G. & Papantonis, A. Forces driving the three-dimensional folding of eukaryotic genomes. *Mol. Syst. Biol.* **14**, e8214 (2018).
88. Nora, E. P. et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012).
89. Sexton, T. et al. Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell* **148**, 458–472 (2012).
90. Crane, E. et al. Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature* **523**, 240–244 (2015).
91. Dali, R. & Blanchette, M. A critical assessment of topologically associating domain prediction tools. *Nucleic Acids Res.* **45**, 2994–3005 (2017).
92. Zufferey, M., Tavernari, D., Oricchio, E. & Ciriello, G. Comparison of computational methods for the identification of topologically associating domains. *Genome Biol.* **19**, 217 (2018).
93. Forcato, M. et al. Comparison of computational methods for Hi-C data analysis. *Nat. Methods* **14**, 679–685 (2017).
94. Beagan, J. A. & Phillips-Cremins, J. E. On the existence and functionality of topologically associating domains. *Nat. Genet.* **52**, 8–16 (2020).
95. de Wit, E. TADs as the caller calls them. *J. Mol. Biol.* **432**, 638–642 (2020).
96. Dowen, J. M. et al. Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* **159**, 374–387 (2014).
97. Gómez-Marín, C. et al. Evolutionary comparison reveals that diverging CTCF sites are signatures of ancestral topological associating domains borders. *Proc. Natl Acad. Sci. USA* **112**, 7542 (2015).
98. Hanssen, L. L. P. et al. Tissue-specific CTCF–cohesin-mediated chromatin architecture delimits enhancer interactions and function in vivo. *Nat. Cell Biol.* **19**, 952–961 (2017).
99. Lupiáñez, D. G. et al. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**, 1012–1025 (2015).
100. Narendra, V. et al. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science* **347**, 1017–1021 (2015).
101. Symmons, O. et al. The *Shh* topological domain facilitates the action of remote enhancers by reducing the effects of genomic distances. *Dev. Cell* **39**, 529–543 (2016).
102. Guo, Y. et al. CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell* **162**, 900–910 (2015).
103. de Wit, E. et al. CTCF binding polarity determines chromatin looping. *Mol. Cell* **60**, 676–684 (2015).
104. Sanborn, A. L. et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl Acad. Sci. USA* **112**, E6456–E6465 (2015).
105. Paliou, C. et al. Preformed chromatin topology assists transcriptional robustness of Shh during limb development. *Proc. Natl Acad. Sci. USA* **116**, 12390–12399 (2019).
106. Fudenberg, G. et al. Formation of chromosomal domains by loop extrusion. *Cell Rep.* **15**, 2038–2049 (2016).
107. Hansen, A. S., Pustova, I., Cattoglio, C., Tjian, R. & Darzacq, X. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *eLife* **6**, e25776 (2017).
108. Stevens, T. J. et al. 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature* **544**, 59–64 (2017).
109. Nagano, T. et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **502**, 59–64 (2013).
110. Flyamer, I. M. et al. Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature* **544**, 110–114 (2017).
111. Gassler, J. et al. A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. *EMBO J.* **36**, 3600–3618 (2017).
112. Bintu, B. et al. Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science* **362**, eaau1783 (2018).
113. Hadjur, S. et al. Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature* **460**, 410–413 (2009).
114. Rubio, E. D. et al. CTCF physically links cohesin to chromatin. *Proc. Natl Acad. Sci. USA* **105**, 8309 (2008).
115. Stedman, W. et al. Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. *EMBO J.* **27**, 654–666 (2008).
116. Fudenberg, G., Abdennur, N., Imakaev, M., Goloborodko, A. & Mirny, L. A. Emerging evidence of chromosome folding by loop extrusion. *Cold Spring Harb. Symp. Quant. Biol.* **82**, 45–55 (2017).
117. Rao, S. S. P. et al. Cohesin loss eliminates all loop domains. *Cell* **171**, 305–320.e324 (2017).
118. Wutz, G. et al. Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J.* **36**, 3573–3599 (2017).
119. Schwarzer, W. et al. Two independent modes of chromatin organization revealed by cohesin removal. *Nature* **551**, 51–56 (2017).
120. Haarhuis, J. H. I. et al. The cohesin release factor WAPL restricts chromatin loop extension. *Cell* **169**, 693–707.e614 (2017).
121. Li, Y. et al. The structural basis for cohesin–CTCF-anchored loops. *Nature* **578**, 472–476 (2020).
122. Davidson, I. F. et al. DNA loop extrusion by human cohesin. *Science* **366**, 1338 (2019).
123. Kim, Y., Shi, Z., Zhang, H., Finkelstein, I. J. & Yu, H. Human cohesin compacts DNA by loop extrusion. *Science* **366**, 1345 (2019).
124. Golfier, S., Quail, T., Kimura, H. & Brugués, J. Cohesin and condensin extrude DNA loops in a cell cycle-dependent manner. *eLife* **9**, e53885 (2020).
125. Phillips-Cremins, J. E. et al. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* **153**, 1281–1295 (2013).
126. Abramo, K. et al. A chromosome folding intermediate at the condensin-to-cohesin transition during telophase. *Nat. Cell Biol.* **21**, 1393–1402 (2019).
127. Zhang, H. et al. Chromatin structure dynamics during the mitosis-to-G1 phase transition. *Nature* **576**, 158–162 (2019).
128. Nagano, T. et al. Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature* **547**, 61–67 (2017).
129. Rhodes, J. D. P. et al. Cohesin disrupts polycomb-dependent chromosome interactions in embryonic stem cells. *Cell Rep.* **30**, 820–835.e810 (2020).
130. Nuebler, J., Fudenberg, G., Imakaev, M., Abdennur, N. & Mirny, L. A. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc. Natl Acad. Sci. USA* **115**, E6697 (2018).
131. Benabdallah, N. S. et al. Decreased enhancer-promoter proximity accompanying enhancer activation. *Mol. Cell* **76**, 473–484.e477 (2019).
132. Symmons, O. et al. Functional and topological characteristics of mammalian regulatory domains. *Genome Res.* **24**, 390–400 (2014).
133. Dekker, J. & Mirny, L. The 3D genome as moderator of chromosomal communication. *Cell* **164**, 1110–1121 (2016).
134. Chang, L.-H., Ghosh, S. & Noordermeer, D. TADs and their borders: free movement or building a wall? *J. Mol. Biol.* **432**, 643–652 (2020).
135. Vian, L. et al. The energetics and physiological impact of cohesin extrusion. *Cell* **173**, 1165–1178.e1120 (2018).
136. Barrington, C. et al. Enhancer accessibility and CTCF occupancy underlie asymmetric TAD architecture and cell type specific genome topology. *Nat. Commun.* **10**, 2908 (2019).
137. Williamson, I. et al. Developmentally regulated Shh expression is robust to TAD perturbations. *Development* **146**, dev179523 (2019).
138. Franke, M. et al. Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature* **538**, 265–269 (2016).
139. Kragesteen, B. K. et al. Dynamic 3D chromatin architecture contributes to enhancer specificity and limb morphogenesis. *Nat. Genet.* **50**, 1463–1473 (2018).

140. Despang, A. et al. Functional dissection of the Sox9–Kcnj2 locus identifies nonessential and instructive roles of TAD architecture. *Nat. Genet.* **51**, 1263–1271 (2019).
141. Kraft, K. et al. Serial genomic inversions induce tissue-specific architectural stripes, gene misexpression and congenital malformations. *Nat. Cell Biol.* **21**, 305–310 (2019).
142. Oudelaar, A. M. et al. Dynamics of the 4D genome during in vivo lineage specification and differentiation. *Nat. Commun.* **11**, 2722 (2020).
143. Cullen, K. E., Kladde, M. P. & Seyfred, M. A. Interaction between transcription regulatory regions of prolactin chromatin. *Science* **261**, 203 (1993).
144. Ptashne, M. Gene regulation by proteins acting nearby and at a distance. *Nature* **322**, 697–701 (1986).
145. Tolhuis, B., Palstra, R.-J., Splinter, E., Grosveld, F. & de Laat, W. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol. Cell* **10**, 1453–1465 (2002).
146. Carter, D., Chakalova, L., Osborne, C. S., Dai, Y.-F. & Fraser, P. Long-range chromatin regulatory interactions in vivo. *Nat. Genet.* **32**, 623–626 (2002).
147. Vakoc, C. R. et al. Proximity among distant regulatory elements at the β-globin locus requires GATA-1 and FOG-1. *Mol. Cell* **17**, 453–462 (2005).
148. Drissen, R. et al. The active spatial organization of the β-globin locus requires the transcription factor EKLF. *Genes Dev.* **18**, 2485–2490 (2004).
149. Deng, W. et al. Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* **149**, 1233–1244 (2012).
150. Song, S.-H., Hou, C. & Dean, A. A positive role for NLI/Ldb1 in long-range β-globin locus control region function. *Mol. Cell* **28**, 810–822 (2007).
151. Deng, W. et al. Reactivation of developmentally silenced globin genes by forced chromatin looping. *Cell* **158**, 849–860 (2014).
152. Krivega, I., Dale, R. K. & Dean, A. Role of LDB1 in the transition from chromatin looping to transcription activation. *Genes Dev.* **28**, 1278–1290 (2014).
153. Kim, S. & Shendure, J. Mechanisms of interplay between transcription factors and the 3D genome. *Mol. Cell* **76**, 306–319 (2019).
154. Weintraub, A. S. et al. YY1 is a structural regulator of enhancer-promoter loops. *Cell* **171**, 1573–1588. e1528 (2017).
155. Beagan, J. A. et al. YY1 and CTCF orchestrate a 3D chromatin looping switch during early neural lineage commitment. *Genome Res.* **27**, 1139–1152 (2017).
156. Kagey, M. H. et al. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430–435 (2010).
157. El Khattabi, L. et al. A pliable mediator acts as a functional rather than an architectural bridge between promoters and enhancers. *Cell* **178**, 1145–1158. e1120 (2019).
158. Jaeger, M. G. et al. Selective mediator dependence of cell-type-specifying transcription. *Nat. Genet.* **52**, 719–727 (2020).
159. Morgan, S. L. et al. Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping. *Nat. Commun.* **8**, 15993 (2017).
160. Kim, J. H. et al. LADL: light-activated dynamic looping for endogenous gene expression control. *Nat. Methods* **16**, 633–639 (2019).
161. Haberle, V. et al. Transcriptional cofactors display specificity for distinct types of core promoters. *Nature* **570**, 122–126 (2019).
162. Schmidt, D. et al. A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Res.* **20**, 578–588 (2010).
163. Cuartero, S. et al. Control of inducible gene expression links cohesin to hematopoietic progenitor self-renewal and differentiation. *Nat. Immunol.* **19**, 932–941 (2018).
164. Chiariello, A. M. et al. A dynamic folded hairpin conformation is associated with α-globin activation in erythroid cells. *Cell Rep.* **30**, 2125–2135.e2125 (2020).
165. Buckle, A., Brackley, C. A., Boyle, S., Marenduzzo, D. & Gilbert, N. Polymer simulations of heteromorphic chromatin predict the 3D folding of complex genomic loci. *Mol. Cell* **72**, 786–797.e711 (2018).
166. Zheng, H. & Xie, W. The role of 3D genome organization in development and cell differentiation. *Nat. Rev. Mol. Cell Biol.* **34**, 903–916 (2019).
167. Dixon, J. R. et al. Chromatin architecture reorganization during stem cell differentiation. *Nature* **518**, 331–336 (2015).
168. Bonev, B. et al. Multiscale 3D genome rewiring during mouse neural development. *Cell* **171**, 557–572.e524 (2017).
169. Stadhouders, R. et al. Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. *Nat. Genet.* **50**, 238–249 (2018).
170. Brown, J. M. et al. A tissue-specific self-interacting chromatin domain forms independently of enhancer-promoter interactions. *Nat. Commun.* **9**, 376 (2018).
171. Jin, F. et al. A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* **503**, 290–294 (2013).
172. Ghavi-Helm, Y. et al. Enhancer loops appear stable during development and are associated with paused polymerase. *Nature* **513**, 89–100 (2014).
173. Andrey, G. et al. A switch between topological domains underlies HoxD genes collinearity in mouse limbs. *Science* **340**, 1234167 (2013).
174. Rubin, A. J. et al. Lineage-specific dynamic and pre-established enhancer–promoter contacts cooperate in terminal differentiation. *Nat. Genet.* **49**, 1522–1528 (2017).
175. de Laat, W. & Duboule, D. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* **502**, 499–506 (2013).
176. Freire-Pritchett, P. et al. Global reorganisation of cis-regulatory units upon lineage commitment of human embryonic stem cells. *eLife* **6**, e21926 (2017).
177. Palstra, R.-J. et al. The beta-globin nuclear compartment in development and erythroid differentiation. *Nat. Genet.* **35**, 190–194 (2003).
178. Vernimmen, D., Gobbi, M. D., Sloane-Stanley, J. A., Wood, W. G. & Higgs, D. R. Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. *EMBO J.* **26**, 2041–2051 (2007).
179. Williamson, I., Lettice, L. A., Hill, R. E. & Bickmore, W. A. Shh and ZRS enhancer colocalisation is specific to the zone of polarising activity. *Development* **143**, 2994–3001 (2016).
180. Boettiger, A. N. et al. Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* **529**, 418–422 (2016).
181. Mateo, L. J. et al. Visualizing DNA folding and RNA in embryos at single-cell resolution. *Nature* **568**, 49–54 (2019).
182. Wang, S. et al. Spatial organization of chromatin domains and compartments in single chromosomes. *Science* **353**, 598 (2016).
183. Cardozo Gizzi, A. M. et al. Microscopy-based chromosome conformation capture enables simultaneous visualization of genome organization and transcription in intact organisms. *Mol. Cell* **74**, 212–222.e215 (2019).
184. Szabo, Q. et al. TADs are 3D structural units of higher-order chromosome organization in Drosophila. *Sci. Adv.* **4**, eaar8082 (2018).
185. Fabre, P. J. et al. Large scale genomic reorganization of topological domains at the HoxD locus. *Genome Biol.* **18**, 149 (2017).
186. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).
187. Alberti, S. Phase separation in biology. *Curr. Biol.* **27**, R1097–R1102 (2017).
188. Boeynaems, S. et al. Protein phase separation: a new phase in cell biology. *Trends Cell Biol.* **28**, 420–435 (2018).
189. Sabari, B. R., Dall'Agnese, A. & Young, R. A. Biomolecular condensates in the nucleus. *Trends Biochem. Sci.* https://doi.org/10.1016/j.tibs.2020.06.007 (2020).
190. Schoenfelder, S. et al. Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nat. Genet.* **42**, 53–61 (2010).
191. Beagrie, R. A. et al. Complex multi-enhancer contacts captured by genome architecture mapping. *Nature* **543**, 519–524 (2017).
192. Cho, W.-K. et al. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **361**, 412 (2018).
193. Cisse, I. I. et al. Real-time dynamics of RNA polymerase II clustering in live human cells. *Science* **341**, 664 (2013).
194. Jackson, D. A., Hassan, A. B., Errington, R. J. & Cook, P. R. Visualization of focal sites of transcription within human nuclei. *EMBO J.* **12**, 1059–1065 (1993).
195. Brown, J. M. et al. Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. *J. Cell Biol.* **182**, 1083–1097 (2008).
196. Chen, Y. et al. Mapping 3D genome organization relative to nuclear compartments using TSA-Seq as a cytological ruler. *J. Cell Biol.* **217**, 4025–4048 (2018).
197. Quinodoz, S. A. et al. Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. *Cell* **174**, 744–757.e724 (2018).
198. Guelen, L. et al. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453**, 948–951 (2008).
199. Denholtz, M. et al. Long-range chromatin contacts in embryonic stem cells reveal a role for pluripotency factors and polycomb proteins in genome organization. *Cell Stem Cell* **13**, 602–616 (2013).
200. Fudenberg, G., Kelley, D. R. & Pollard, K. S. Predicting 3D genome folding from DNA sequence with Akita. *Nat. Methods* https://doi.org/10.1038/s41592-020-0958-x (2020).
201. Schwessinger, R. et al. DeepC: predicting 3D genome folding using megabase-scale transfer learning. Nat. Methods https://doi.org/10.1038/s41592-020-0960-3 (2020).
202. Zhao, Z. et al. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat. Genet.* **38**, 1341–1347 (2006).
203. Simonis, M. et al. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat. Genet.* **38**, 1348–1354 (2006).
204. Davies, J. O. J. et al. Multiplexed analysis of chromosome conformation at vastly improved sensitivity. *Nat. Methods* **86**, 1202–1210 (2015).
205. Hughes, J. R. et al. Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nat. Genet.* **46**, 205–212 (2014).
206. Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *Science* **295**, 1306–1311 (2002).
207. Übelmesser, N. & Papantonis, A. Technologies to study spatial genome organization: beyond 3C. *Brief. Funct. Genomics* **18**, 395–401 (2019).
208. Brant, L. et al. Exploiting native forces to capture chromosome conformation in mammalian cell nuclei. *Mol. Syst. Biol.* **12**, 891 (2016).
209. Redolfi, J. et al. DamC reveals principles of chromatin folding in vivo without crosslinking and ligation. *Nat. Struct. Mol. Biol.* **26**, 471–480 (2019).
210. Finn, E. H. et al. Extensive heterogeneity and intrinsic variation in spatial genome organization. *Cell* **176**, 1502–1515.e1510 (2019).
211. Nguyen, H. Q. et al. 3D mapping and accelerated super-resolution imaging of the human genome using in situ sequencing. *Nat. Methods* **17**, 822–832 (2020).
212. Osborne, C. S. et al. Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat. Genet.* **36**, 1065–1071 (2004).

Acknowledgements
The authors thank R. A. Beagrie, J. M. Brown and M. T. Kassouf for insightful comments on the manuscript. The authors apologize to colleagues whose important studies they were unable to cite due to space constraints. This work was supported by the Max Planck Society (A.M.O.) and the UK Medical Research Council (D.H.; MR/T014067/1).

Author contributions
The authors contributed equally to all aspects of the article.

Competing interests
The authors declare no competing interests.

Peer review information
Nature Reviews Genetics thanks A. Papantonis, J. Wysocka and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.