


OPINION

Organizational principles of 3D genome architecture

M. Jordan Rowley and Victor G. Corces 

Abstract | Studies of 3D chromatin organization have suggested that chromosomes are hierarchically organized into large compartments composed of smaller domains called topologically associating domains (TADs). Recent evidence suggests that compartments are smaller than previously thought and that the transcriptional or chromatin state is responsible for interactions leading to the formation of small compartmental domains in all organisms. In vertebrates, CTCF forms loop domains, probably via an extrusion process involving cohesin. CTCF loops cooperate with compartmental domains to establish the 3D organization of the genome. The continuous extrusion of the chromatin fibre by cohesin may also be responsible for the establishment of enhancer–promoter interactions and stochastic aspects of the transcription process. These observations suggest that the 3D organization of the genome is an emergent property of chromatin and its components, and thus may not be only a determinant but also a consequence of its function.

Chromatin is organized within the 3D nuclear space, efficiently packaging the genome while allowing proper expression and replication of the genetic material. The relative position of specific loci in the nucleus of individual cells within a population can be visualized using microscopy-based techniques, thus allowing for the understanding of cell-to-cell variation in the arrangement of the chromatin fibre at individual loci. Molecular approaches such as Hi-C can be used to map all interactions between distant loci in the genome, but they require the use of millions of cells and, therefore, provide a view of genome 3D organization that represents an ensemble of the individual cells present in the population¹. High-throughput microscopy and single-cell Hi-C are beginning to bridge information obtained using these two approaches^{2–7}.

Here, we discuss recent findings suggesting a departure from established dogma in our view of the mechanisms by which 3D chromatin organization is established and its relationship to the regulation of transcription. Recent observations indicate the existence of two independent but partially related organizational principles governing the

formation and maintenance of 3D chromatin organization: small compartmental domains that form as a consequence of the transcription and chromatin state and cohesin-mediated CTCF loops. In the sections below, we first discuss a commonly held model of chromatin organization and then present recent work indicating the need to revise this model. We describe common methods that have been used to computationally identify features of Hi-C maps as a way to demonstrate how some computational parameters, particularly bin size, can affect the interpretation of these data. We discuss evidence suggesting the existence of small compartmental domains, which are established as a consequence of interactions between proteins involved in transcription activation or silencing in each domain. We then analyse evidence in support of cohesin-mediated extrusion as the mechanism underlying the establishment of CTCF loops. Finally, we bring these two sources of information together into a unifying view of 3D genome organization, suggesting that transcription and architecture are closely interdependent and influence each other. Rather than a hierarchical top-down view of nuclear architecture in which 3D chromatin

organization determines gene expression, we suggest a balance between compartmental domains and CTCF loops. These observations lead to a new understanding of the causal relationship between transcription and the 3D arrangement of the genome in the nuclear space.

Features of chromatin organization

Contact maps of Hi-C data showed a genome-wide view of interactions between all sequences in the mammalian genome for the first time in 2009 (REF.⁸). These maps display a plaid pattern of chromatin interactions over distances as far as the length of a chromosome. These interactions are a manifestation of the segregation of the genome into two compartments, named A and B, defined by the eigenvector or first component of a principal component analysis (PCA)⁸ (FIG. 1a). Although the eigenvector gives information on the A or B state, the magnitude of this characteristic and the region of the linear genome to which this state applies, the published literature generally uses the term compartment to refer to both the domains and their interactions manifested as the plaid pattern observed in Hi-C heat maps. Sequences in the A compartment generally contain transcribed genes and active histone modifications, although some regions in A compartments are not transcribed. Analogously, the B compartment contains inactive genes with histone modifications associated with a transcriptionally repressed state⁸ but also some transcribed genes (FIG. 1a). Owing to the high cost of sequencing and the inherent short-range interaction bias of Hi-C, initial maps had low coverage across the 2D genome matrix, with ~10 million paired reads⁹. Thus, sequences were binned in 1 Mb bins for the purpose of identifying compartments using PCA and, consequently, compartments were identified as being multi-Mb in size⁸. As sequencing costs decreased, Hi-C data sets became richer, with 200–300 million paired reads, allowing the partitioning of data for analysis into ~40 kb bins. Using this smaller bin size, computational algorithms measuring the directionality of interactions in the genome identified topologically associating domains (TADs) as structures in the 0.2–1.0 Mb range^{9–13} (FIG. 1b). Whereas A and B

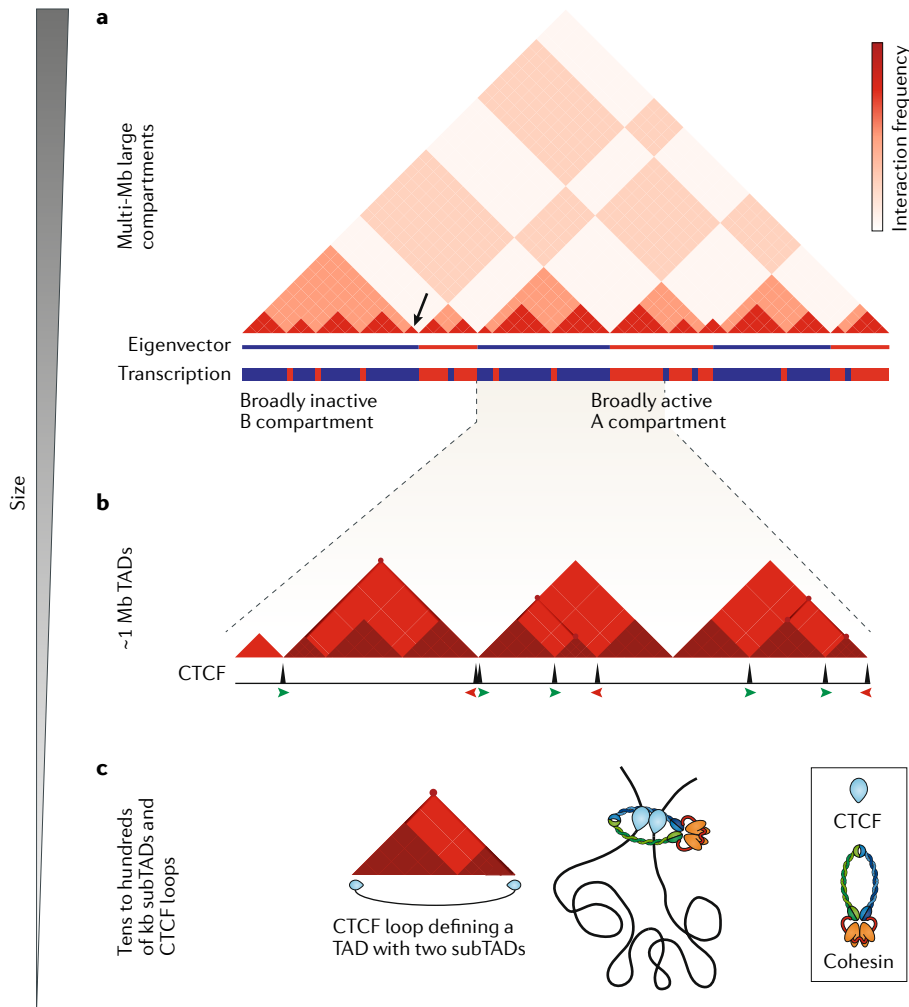


Fig. 1 | Current model of chromatin organization. The hierarchical model of chromatin organization suggests that different sized features contribute to each other's formation. In this model, compartments are large multi-megabase structures of the 3D genome, whereas topologically associating domains (TADs) are substructures inside compartments. **a** | The interaction signals (varying intensities of red) from low-resolution Hi-C data partitioned into megabase-sized bins are shown. The panel represents a cartoon version of an actual Hi-C heat map. The eigenvector describes the first component of the principal component analysis and identifies A (red) and B (blue) compartments, which correlate with mostly transcriptionally active and inactive regions of the genome, respectively⁸. **b** | TADs are smaller regions of the genome identified with higher-resolution Hi-C data partitioned into ~40 kb bins using an algorithm to detect changes in the directionality of interactions^{9–12}. The panel shows a small section of the genome corresponding to one B compartment and half of an A compartment in the diagram above. TADs contain smaller subTADs characterized by higher interaction frequencies (darker shade of red) and CTCF loops detected as strong punctate signals corresponding to strong interactions between CTCF sites. The direction of CTCF sites is indicated by the orientations and red and green colours of the arrowheads. Note that only some TADs coincide with CTCF loops, and CTCF is present at only the borders of some TADs. Only some CTCF loops are detected at this resolution. **c** | The structure of a TAD as detected at ~40 kb resolution, containing two subTADs and flanked by CTCF–cohesin sites forming a loop, is shown.

compartments correspond to domains that interact preferentially with sequences in other A or B compartment regions, respectively, TADs correspond to sequences that interact preferentially with themselves rather than with other regions of the genome. TADs are separated by boundaries that are enriched in CTCF binding sites and highly transcribed genes. It is important to consider

that not all TADs defined computationally at ~40 kb resolution using algorithms that measure a directionality index are flanked by CTCF sites^{9–13} (FIG. 1b), although a subset of TADs are defined by CTCF loops (FIG. 1c). The relative sizes of these features have led to a hierarchical model of chromatin organization, in which compartments are composed of several TADs^{14,15} (FIG. 1a).

The first high-resolution Hi-C data set of a mammalian genome contained ~5 billion paired reads, making it possible to bin reads at 5 kb resolution. The use of the Arrowhead algorithm allowed the identification of contact domains, which are smaller in size than TADs. A subset of these contact domains arises as a consequence of point-to-point interactions between two sequences bound by CTCF. These loops are relatively stable, as they can be observed in high-resolution Hi-C data as strong punctate signals at the summits of some domains¹⁶ (FIG. 2a). Given the population nature of Hi-C data, these strong spots can be interpreted as interactions occurring in all the cells of the population or as very stable interactions taking place in a subpopulation of cells. In addition to these CTCF loop domains, a second type of contact domain, originally named ‘ordinary domains’, is characterized by the presence of specific histone modifications and is not flanked by CTCF¹⁶ (FIG. 2a). Identification of compartments using PCA and high-resolution Hi-C data binned at 10–50 kb suggests that compartments are smaller than previously thought, as small as a single active or inactive locus¹⁷ (FIG. 2a; compare eigenvector with domains above, represented by red triangles). Therefore, the ordinary domains are likely to correspond to small compartments formed by the segregation of active and inactive chromatin^{16,17}. In the rest of the manuscript, we will use the term ‘compartmental domains’^{17,18} to refer to domains identified by PCA using high-resolution Hi-C data and bin sizes in the range of a few kb, whereas we will use the term compartments to refer to the original domains identified using bins in the hundreds of kilobases to 1 Mb range. Compartmental domains, as is the case for compartments, can be classified into active A and inactive B domains, are smaller than TADs and are present inside, between or overlapping CTCF loops in mammals^{17,18} (FIG. 2a). CTCF loops can sometimes encompass two compartmental domains with different transcriptional states, increasing interactions between them (FIG. 2b). In other cases, CTCF loops contain only sequences in the same transcriptional state and decrease interactions between two adjacent compartmental domains (FIG. 2b). These observations suggest that, instead of being composed of hierarchically related large compartments and smaller TADs, chromosomes are organized into similarly sized compartmental domains and CTCF loops. High-resolution Hi-C maps also allowed the identification of compartmental domains in *Drosophila melanogaster*, where A

compartmental domains have a median size of just 15 kb (REF.¹⁷). Their size corresponds to the median size of transcriptionally active blocks of chromatin, a finding that suggests a relationship between compartmental domain formation and chromatin state¹⁷. Compartmental domains are also found in many other organisms such as *Arabidopsis thaliana*, *Oryza sativa*, *Zea mays*, *Caenorhabditis elegans*, *Neurospora crassa* and *Plasmodium falciparum* and are likely to explain most aspects of 3D chromatin organization in these organisms^{17,19}.

3D organization and transcription

The close correlation between A and B compartmental domains and the transcriptionally active or inactive state of chromatin suggests a possible causal relationship between the two. Indeed, the correspondence between 3D chromatin organization and transcriptional state is sufficiently precise to accurately predict Hi-C maps in many different organisms using global run-on sequencing (GRO-seq), RNA sequencing (RNA-seq) or histone modifications^{17,20–22}. When discussing evidence supporting this relationship, we distinguish between transcription per se and the transcriptional state, that is, proteins or histone modifications normally associated with expressed or silenced genes, even if the genes in a transcriptionally active state are not actually transcribed. Several studies have tried to establish a possible role for transcription in chromosome organization using chemical inhibition of the initiation or elongation of RNA polymerase II (RNAPII). Some drugs used to inhibit transcription also affect RNAPII levels at the promoter, whereas others do not. Therefore, when interpreting results from this type of experiment, it is important to consider whether only transcription has been affected or whether the presence of proteins in the transcription complex — which may be responsible for mediating interactions between A compartmental domains — has been affected too. In the bacteria *Caulobacter crescentus* (also known as *Caulobacter vibroides*) and *Bacillus subtilis*, inhibition of transcription using rifampicin results in a loss of contact domains^{23,24}. However, similar experiments in eukaryotes have led to more nuanced observations. Inhibition of transcription initiation in *D. melanogaster* cells using triptolide results in a reduction of interactions inside compartmental domains as well as a reduction in the plaid pattern representing contacts between compartmental domains^{17,25,26}. In contrast to bacteria, inhibition of

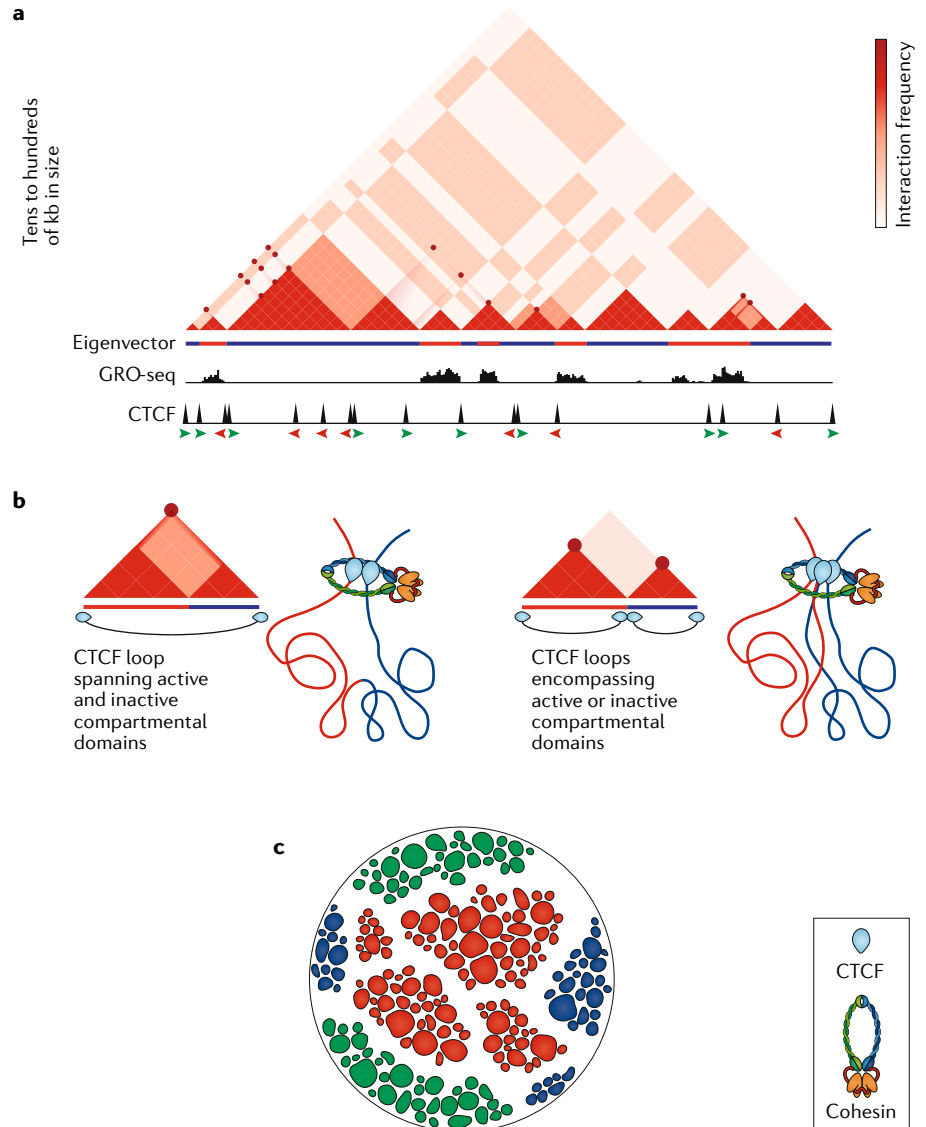


Fig. 2 | New model of chromatin organization. An alternative model of chromatin organization incorporates recent findings obtained with very-high-resolution data partitioned in 1–5 kb bins^{16,17}. **a** | The cartoon corresponds to the domain marked with a small, black arrow in FIG. 1a, and it is a representation of the actual Hi-C heat map, emphasizing the complexity of interactions present in a region that appears as a uniform minute topologically associating domain (TAD) in low-resolution data. The eigenvector obtained by binning the data at 5–20 kb allows the identification of compartmental domains¹⁷, which accurately correspond to the active or inactive transcriptional state determined by global run-on sequencing (GRO-seq). Punctate signals represent CTCF loops between CTCF binding sites in convergent orientation. The direction of CTCF sites is indicated by the orientations and red and green colours of the arrowheads. ‘Ordinary domains’, which are not spanned by CTCF loops, for example, domains 4–6 from the right, are likely to correspond to compartmental domains^{16,17}. **b** | Some CTCF loops encompass active and inactive compartmental domains, increasing interactions between these two domains that would normally not take place (left). Other CTCF loops encompass individual compartmental domains, and the formation of the loop decreases interactions between two adjacent domains (right). Therefore, the presence of CTCF loops modulates interactions among compartmental domains. **c** | Segregation of chromatin states in the nucleus may occur as a consequence of the presence of different classes of multivalent proteins that mediate class-specific interactions to create different phases, which result in droplets of distinct chromatin states within the nucleus. In the cartoon, red represents proteins and histone modifications present at genes or regulatory sequences in a transcriptionally active state, blue represents histone H3 lysine 27 trimethylation (H3K27me3) and Polycomb group proteins, and green represents H3K9me3 and heterochromatin protein 1α (HP1α).

transcription in *D. melanogaster* does not eliminate compartmental domains entirely. One possible explanation is that the maintenance of these domains is dependent on the presence of proteins related to transcription rather than on the transcription process. Under the conditions used to inhibit transcription with triptolide, a substantial amount of RNAPII, and perhaps other components of the transcription complex, remains at the promoter²⁷. Indeed, the degree of RNAPII loss after treatment correlates with the degree to which compartmental domains are reduced¹⁷. Furthermore, the *D. melanogaster* heat shock response, which results in the downregulation of most genes, has a stronger effect on RNAPII levels than triptolide treatment and causes a more pronounced loss of compartmental domains^{17,25}.

Studies of the relationship between transcription and the formation of compartmental domains have also taken

advantage of the gradual establishment of normal transcription during early embryonic development. In *D. melanogaster*, embryos undergo 12 nuclear divisions before global transcription can be detected by standard methods. During nuclear cycle 12 (nc12), a few genes are transcribed but most are inactive. These few transcribed genes correspond to the few compartmental domains present at this stage²⁶. After genome-wide transcriptional activation during nc13 and nc14, new domains form around the newly transcribed loci²⁶. These observations support a relationship between transcriptional activation and compartmental domain formation during *D. melanogaster* development. To test this relationship, α -amanitin or triptolide were used to inhibit transcription at an early embryonic stage in order to prevent normal genome-wide transcriptional activation. This inhibition resulted in decreased compartmental domain formation, suggesting that transcription may be

responsible for organizing chromatin into domains during *D. melanogaster* embryonic development²⁶. However, compartmental domains were not entirely eliminated, and their intensity correlates with the amount of RNAPII left at gene promoters after transcription inhibition¹⁷. These results support the conclusion that proteins related to the transcriptional state, rather than transcription itself, may be responsible for the establishment of compartmental domains.

Similar experiments have been carried out during vertebrate embryonic development. In contrast to *D. melanogaster*, compartments are detected early in zebrafish development before the genome is transcribed, and they are lost upon genome-wide transcriptional activation²⁸. However, results from Hi-C experiments performed with mouse embryos at different stages of embryonic development suggest that the appearance of contact domains correlates with the time of transcription activation, with weak interactions observed at the two-cell stage that strengthen in the eight-cell embryo^{29,30}. Treatment with α -amanitin starting at the zygotic pronuclear 4 (PN4) stage and continuing for 2 days, a time in which embryos would normally mature to the eight-cell stage, instead results in embryos arrested at the two-cell stage³⁰. Hi-C maps of these arrested embryos show slightly stronger domains than those present in untreated two-cell stage embryos but weaker domains than the normal eight-cell stage. When compared with normal eight-cell stage embryos, the results suggest an effect of transcription inhibition on compartmental domain organization. However, as the treated cells are arrested at the two-cell stage but their age is that of control eight-cell embryos, this result is difficult to interpret, as it is unclear to which stage they should be compared. In a similar study, transcription was inhibited using α -amanitin beginning at the PN3 stage and lasting for 20 or 45 hours, when embryos would normally proceed to late two-cell or eight-cell stages, respectively. In both cases, embryos were arrested at the two-cell stage²⁹. Hi-C maps of these embryos show that, although transcriptional inhibition did not completely stop the progression of chromatin organization during embryonic development, many features of chromatin organization do appear less pronounced (see Du et al. 2017 Extended Data 8b and compare Extended Data 8d with Extended Data 4e)²⁹. These results may be interpreted to suggest that transcription may not play a large role in mammalian chromatin

Glossary

ChIA-PET

Chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) utilizes chromatin immunoprecipitation followed by proximity ligation to identify chromatin interactions between loci bound by a protein of interest.

ChIP–exo and CHIP–nexus

Chromatin immunoprecipitation (ChIP) followed by exonuclease digestion (ChIP–exo) is a technique that is used in place of standard chromatin immunoprecipitation followed by sequencing (ChIP–seq) to identify protein binding sites at higher resolution. The higher resolution is achieved because exonuclease treatment trims stretches of flanking DNA that are not directly bound by the protein of interest. ChIP–nexus utilizes a different library preparation strategy to reportedly improve signal compared with ChIP–exo.

Compartmental domains

Domains in Hi-C data that are not formed by a CTCF loop and are formed instead by the segregation of active and inactive chromatin.

CTCF loops

Point-to-point interactions between loci that coincide with CTCF and cohesin occupancy and often contain CTCF motifs in convergent orientation. These appear as bright punctae corresponding to high-frequency interactions in Hi-C contact maps.

Directionality index

A common method of computationally identifying topologically associating domain (TAD) borders. A directionality is calculated for each binned genomic locus to describe the preference of interaction signal with bins on the right (positive directionality) or with bins on the left (negative directionality). TAD borders are defined at transitions between negative and positive directionality.

Gene loop

A loop formed by interactions between the transcription start site and the transcription termination site.

Global run-on sequencing

(GRO-seq). A method involving isolation of nascent transcripts and high-throughput sequencing to study active transcription genome-wide.

Hi-C

A method using proximity ligation and high-throughput sequencing to identify all interactions taking place throughout the genome.

Loop extrusion

A model in which chromatin is pulled through the cohesin or condensin ring to form loops.

Oligopaint

A method of labelling DNA using short fluorescently labelled oligonucleotides for high-resolution imaging of chromatin.

Ordinary domains

Domains observed in Hi-C data that are not spanned by a CTCF loop. They are probably the same as compartmental domains.

STORM

(Stochastic optical reconstruction microscopy). Super-resolution imaging using individual photo-switchable fluorophores.

Transcriptional states

The state of a locus based on the presence of chromatin-bound proteins or covalent histone modifications that correlate with gene silencing or active transcription.

Transcription factory

A distinct nuclear location where RNA polymerase II (RNAPII) accumulates on the basis of the observation that components of the transcription complex can be detected as discrete foci by microscopy. The transcription factory hypothesis suggests that genes are recruited to these nuclear locations in order to be transcribed.

organization. However, both CTCF loops and compartmental domains contribute to the establishment of 3D architecture, and inhibiting transcription may affect only compartmental domains. Thus the presumably unchanged CTCF loops after transcriptional inhibition may not allow the detection of changes in compartmental domains owing to the relatively low Hi-C sequencing depth used in these experiments^{29,30}. Although technically difficult, the definitive answer to these questions will require an analysis of the distribution of CTCF and cohesin during early mammalian embryogenesis, as well as the effect of transcription inhibitors on the actual transcription rate and the levels of RNAPII and other transcription factors at the promoters of genes. This will distinguish between a requirement for transcription per se versus a requirement for factors in the transcription complex that mediate interactions leading to the formation of compartmental domains.

Additional insights into the relationship between the formation of compartmental domains and transcription come from the analysis of 3D chromatin organization in the mature oocyte and sperm, the two cells whose genomes will contribute to the one-cell zygote. The mature oocyte is arrested in metaphase II of meiosis, and its genome is not transcribed. As expected from previous analyses of 3D chromatin organization in mitotic chromosomes, Hi-C maps reveal an absence of long-range intra-chromosomal interactions in the nucleus of mature oocytes^{29–31}. However, although also transcriptionally inactive, sperm chromosomes are organized in the 3D space in a manner similar to that of embryonic stem or somatic cells, with clear compartmental domains and CTCF loops^{29,30,32,33}. One explanation for this observation is that 3D organization is established in round spermatids, which are transcriptionally active, and it is maintained in sperm. Alternatively, sperm retain transcription factors and nucleosomes with specific histone modifications at a subset of particular sites, including transcription start sites (TSSs) and distal intergenic sites presumed to be regulatory sequences³². It is possible that proteins present at these sites mediate interactions that result in the formation of compartmental domains. It is unclear whether these domains are maintained in the zygote or whether they disappear as protamines present in the sperm are replaced by nucleosomes and then are re-established as new transcription is initiated in two-cell embryos.

Formation of compartmental domains

Compartmental domains in *D. melanogaster* and mammals are composed of one or more adjacent genes in the same transcriptional or chromatin state^{17,18}. These domains are the result of interactions among sequences located within the domain. Most frequent among these interactions are those taking place between the start and the termination sites of transcribed genes¹⁷. These interactions cause the formation of gene loops, resulting in the accumulation of proteins involved in transcription at a single site, including transcription factors bound to enhancers, the transcription complex at the promoter and proteins involved in splicing and transcription termination. In *D. melanogaster*, ~15 different architectural proteins, including CTCF, may also contribute to this local increase in protein concentration^{17,34}. These proteins, which were originally identified on the basis of their insulating effects on enhancer–promoter interactions, do not form stable loops as CTCF does in vertebrates^{17,35}. Instead, *D. melanogaster* architectural proteins and associated RNAs bind to genomic sites containing multiple binding motifs and located in close proximity to promoters, perhaps contributing to an increase in the local protein concentration at these sites. Interestingly, some of these proteins are modified by sumoylation or parylation, which could amplify their ability to interact with other proteins able to bind SUMO or poly(ADP-ribose) (PAR)³⁴. It is therefore tempting to speculate that insulator bodies, which can be visualized in the nucleus as nuclear bodies or membraneless organelles with antibodies to any of the *D. melanogaster* architectural proteins, are the result of inter-compartmental domain contacts mediated by cooperative interactions among architectural and other proteins involved in the transcription process^{36–40}.

The plaid pattern observed in Hi-C heat maps from a population of cells is likely to be a combination of individual interaction patterns present in each cell of the population. Inter-compartmental domain interactions are stochastic, and their frequency or stability may depend on the number, affinity and interaction ability of the proteins involved, which determine the cooperativity of the interactions³⁸. It has been recently proposed that cooperative interactions among large numbers of multivalent transcription factors can be explained by a phase separation model^{36,37}. On the basis of this model, active and inactive regions of the genome

— A and B compartmental domains, respectively — containing two different sets of multivalent proteins may be able to interact with members of their own class, forming two different phases that preclude inter A–B compartmental contacts. For example, it was recently shown that phase-separated heterochromatin protein 1 α (HP1 α)-mediated heterochromatin droplets are formed in vitro and can be detected in vivo³⁸. Similarly, phase-separated droplets of active chromatin might be formed by proteins that contain intrinsically disordered domains, such as those commonly found in transcription factors and RNAPII^{39,40}. Indeed, it was recently shown that RNAPII forms phase-separated droplets⁴¹. Interestingly, RNAPII droplets were dependent on the hyperphosphorylation of the RNAPII carboxy-terminal domain, indicating that transcriptional activation could promote phase separation of active chromatin⁴¹. Phase separation of chromatin into droplets could regulate functional aspects of compartmental domain interactions. For example, droplet formation may increase the concentration of transcription factors and RNAPII at active chromatin, analogous to the transcription factory hypothesis⁴². A model of phase-separated chromatin (FIG. 2c) would entail the constant fusion and fission of chromatin droplets, suggesting that compartmental domains are involved in dynamic interactions. The dynamics of such droplet activity in the cellular population could explain why active compartmental domains appear to interact with every other active locus across the length of the chromosome in Hi-C heat maps^{8,16}.

CTCF loop domains from extrusion

In addition to compartmental domains and their associated long-range intra-chromosomal interactions, high-resolution Hi-C maps of mammalian cells show thousands of intense, highly localized, punctate signals that correspond to loops anchored at CTCF sites¹⁶ (FIG. 2b). Ninety-two per cent of CTCF loops identified by Hi-C, or 65% identified by CTCF ChIA-PET, occur between motifs in a convergent forward–reverse orientation^{16,43} (FIG. 3a). Most other CTCF loops are formed between motifs oriented in forward–forward or reverse–reverse orientations but are comparatively weak in Hi-C experiments, making them more easily detectable by CTCF ChIA-PET^{16,43}. Therefore, both Hi-C and CTCF ChIA-PET show that CTCF loops preferentially occur between motifs

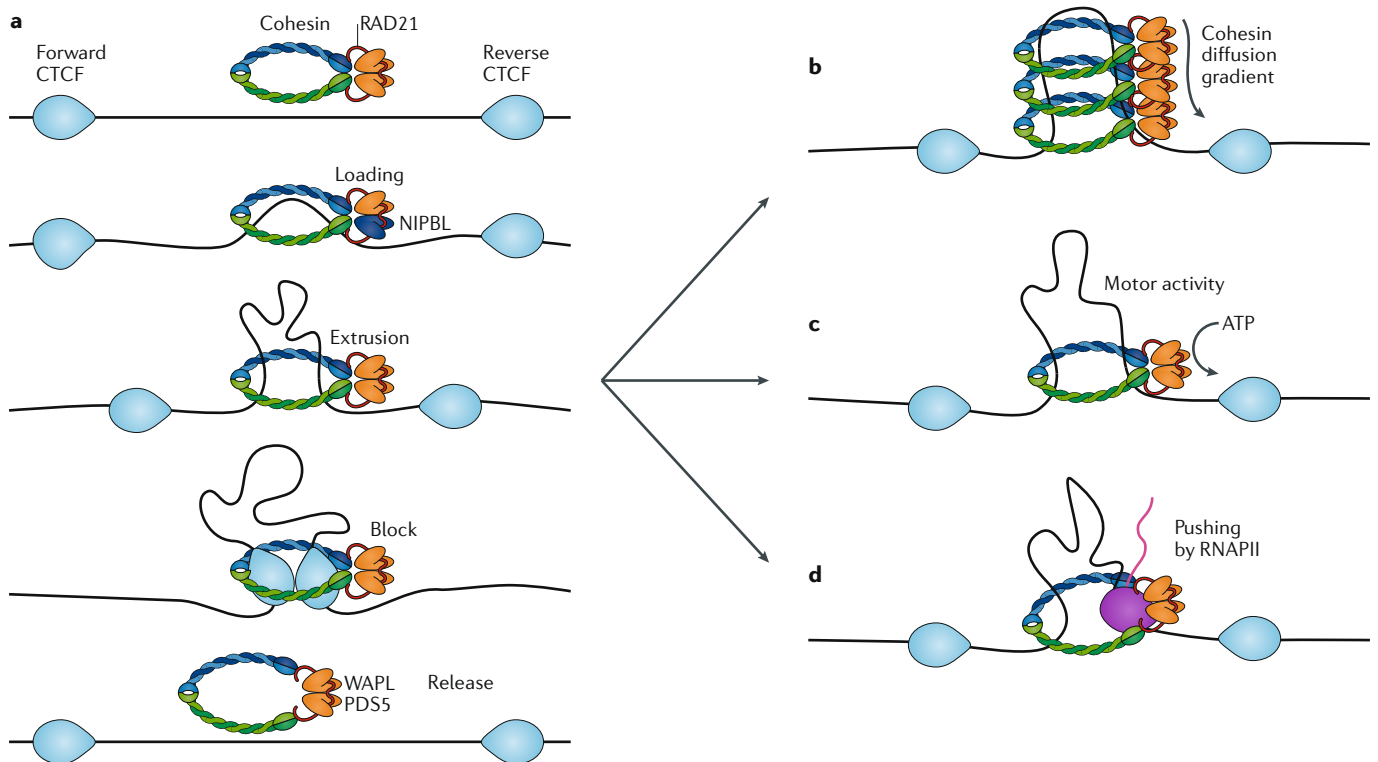


Fig. 3 | Mechanisms of loop extrusion. a | General model of loop extrusion. The extrusion process involves cohesin composed of structural maintenance of chromosomes (SMC) proteins SMC1 and SMC3 and RAD21; cohesin is loaded onto chromatin via NIPBL⁹⁴. Extrusion is blocked at CTCF sites arranged in a convergent head-to-head orientation^{45–49}. Some proportion of cohesin is released throughout this process by the activity of WAPL and PDS5 (REF. ⁹⁴). **b** | Extrusion via cohesin diffusion.

Extrusion may occur by constant loading of cohesin resulting in a diffusion gradient⁷⁰. **c** | Extrusion via cohesin motor activity. An alternative explanation for extrusion is that the process is driven by the motor activity of cohesin via ATP hydrolysis^{54,72}. **d** | Extrusion via pushing of cohesin by RNA polymerase II (RNAPII). Other factors able to move along chromatin, such as RNAPII (purple), may help cohesin to extrude DNA^{50,51,73–76}.

in convergent orientation, with a strong bias against the opposite orientation. The importance of this orientation preference has been demonstrated by CRISPR-mediated inversion of CTCF motifs at individual loci, which alters the corresponding loop domain and allows the formation of new enhancer–promoter interactions⁴⁴. The consistency in the preference for convergent CTCF motifs at loop anchors is also underscored by the ability to accurately predict changes in CTCF loops after CRISPR-mediated inversion or deletion of CTCF motifs⁴⁵. These findings have important implications to explain how CTCF bound at a specific site finds its partner site to form a loop. If DNA-bound CTCF is able to diffuse unrestrictedly in the 3D space to encounter the second loop anchor at tens or hundreds of kilobases away, the orientation of the CTCF-bound motif should be irrelevant to the establishment of CTCF–CTCF interactions and the formation of a loop. Experimental observations indicating a preference for motifs in a convergent orientation serving as loop anchors suggest that CTCF molecules encounter each other

in a dimension-restricted space. This can be accounted for if the formation of CTCF loops takes place via an extrusion process mediated by the cohesin ring^{45–49}.

The loop extrusion model suggests that structural maintenance of chromosomes (SMC) proteins, as part of cohesin or condensin, progressively extrude chromatin until blocked by CTCF bound to a properly oriented site^{45–49} (FIG. 3a). Several groups have investigated the mechanics of SMC-mediated loop formation, and multiple observations suggest that cohesin and condensin rings can topologically entrap and move along the DNA until meeting an obstacle that blocks this movement. For example, in vitro experiments have shown that cohesin can diffuse along anchored DNA, a process that can be blocked by CTCF^{50–52}. It was also shown that condensin can translocate a second piece of DNA relative to the first, and more recently Ganji et al. were able to visualize condensin-mediated loop extrusion in vitro^{53,54}. Condensin-mediated extrusion of naked DNA can occur via a single ring, but it is not known whether

this also happens with chromatin templates or whether a single cohesin ring can also extrude DNA in vitro⁵³. One cohesin ring can capture two separate pieces of DNA but only if the second piece of DNA is single-stranded⁵⁵. This provides some evidence of cohesin-mediated extrusion through a single ring, but more work is needed to fully understand the process on a chromatin template in vivo. In testing the extrusion model, polymer physics simulations have suggested that cohesin starts randomly in the genome and that the extrusion process is continuous until blocked by CTCF when approaching from the 3' end of the motif, that is, inside the loop⁴⁶ (FIG. 3a). In support of this, results from ChIP–exo and ChIP–nexus experiments show that cohesin is enriched at the 3' end of CTCF motifs^{43,56}. Additionally, whereas lack of cohesin results in loss of CTCF loop domains^{18,57,58}, depletion of the cohesin release factor WAPL causes cohesin to remain bound to chromatin for longer periods, resulting in the formation of larger CTCF loops^{58–60}. This suggests a processive mechanism of loop formation, such that

the size of CTCF loops corresponds to the amount of time that cohesin is able to extrude chromatin before encountering an obstacle that stops extrusion.

How CTCF provides a barrier to extrusion is unknown, but it may be related to CTCF-induced conformational changes to chromatin. CTCF binding repositions nucleosomes⁶¹ and was recently shown to cause large changes to naked DNA *in vitro*⁶². Experiments using atomic force microscopy showed that DNA wraps around the bound CTCF protein, forming CTCF-centric circles ~67–80 nm in diameter⁶². These CTCF–DNA circles are larger than the ~20 nm proteins that are able to block cohesin sliding *in vitro*^{50,51}. However, this fails to explain the unidirectional blockage of extrusion coincident with CTCF motif orientation or the fact that the presence of CTCF and cohesin alone cannot fully explain the formation of loops between specific sites in the genome because many CTCF peaks detected by chromatin immunoprecipitation followed by sequencing (ChIP–seq) do not form loops even when in convergent orientation^{16,63,64}. Additionally, many loops appear to change during differentiation, often without changes in CTCF binding^{65,66}. For example, experiments using Hi-C resulted in the identification of 184 loops gained and 33 loops lost during macrophage differentiation without alterations in CTCF occupancy⁶⁵. Interestingly, these loops are enriched in the AP-1 motif, suggesting that transcription factors may play a role in regulating loop formation between CTCF-bound loci⁶⁵. Thus, although much of the focus on the establishment of CTCF loops has centred on this protein, future work should also focus on the possible role of transcription factors in CTCF loop formation.

Mechanisms of loop extrusion

The mechanistic details of the process by which loops are extruded by the cohesin ring are now beginning to be understood. Analysis of 3D chromatin organization of metaphase chromosomes suggests that all compartmental domains, their interactions and CTCF-mediated loops disappear as condensin-mediated loops progress to condense chromosomes during mitosis^{31,67}. This indicates that cells need to reconstruct the 3D organization of their genomes when they exit mitosis. As compartmental domains appear to be a consequence of the transcriptional or chromatin state of the genome, the activation of transcription at the M/G1 boundary should be sufficient to restore this aspect of 3D organization after the loss of condensin-mediated

chromosome condensation. Alternatively, the memory of chromatin state via histone modifications or other proteins that remain bound during mitosis may be sufficient to restore compartmental domains in the absence of actual transcription. However, re-establishment of CTCF loops will require cohesin-mediated extrusion, which presumably will necessitate large amounts of ATP and will have to be done rapidly. Recent experiments examining the re-establishment of CTCF loops after cohesin removal have shed light on the issue of how rapidly cohesin can extrude to restore CTCF loop domains. Lieberman Aiden and colleagues used auxin-mediated degradation of the RAD21 subunit of cohesin to examine this process. After removal of RAD21 for a 6-hour period, all CTCF loops are eliminated, indicating that CTCF cannot stabilize loops without cohesin involvement. This supports a model whereby loops and the associated domains are dynamic features that may require constant extrusion^{46,68}. Restoring cohesin by removal of auxin results in the re-establishment of CTCF loops as large as 900 kb within 40 minutes¹⁸. If loop extrusion was to use one cohesin complex per loop, this suggests that the speed of extrusion by cohesin is at least 375 bp per second or that each of the two topologically entrapped sections of DNA are pulled through cohesin rings at 188 bp per second. This is likely to be a conservative estimate, as it does not account for the time required for cohesin to re-accumulate in the nucleus and to be loaded onto chromatin before extrusion can begin again. Analysis of chromosome organization in *B. subtilis* using Hi-C at several different time points also gives insights into the speed of the extrusion process. These experiments allowed the visualization of the progressive ‘zip-up’ of DNA from the replication origin by an SMC protein complex^{23,69}. The estimate of the rate of extrusion in bacteria on the basis of these experiments is ~850 bp per second²³. A second study using real-time imaging of loop extrusion *in vitro* found that condensin extrudes at ~600 bp per second⁵³. Thus, although the actual extrusion rate in mammals is unknown, it is likely to be somewhere between 374 and 850 bp per second, suggesting that extrusion is a sufficiently fast process to account for the need to rapidly form loops at different stages of the cell cycle.

The mechanisms by which cohesin can translocate along the DNA at such speeds are not known, but several models have been put forward to explain this phenomenon, including diffusion,

motor activity and pushing by other macromolecular assemblies. Diffusion of cohesin by Brownian motion along the 10 nm chromatin fibre may explain the forces controlling the extrusion process (FIG. 3b). Simulations of this model have suggested that diffusion may reach high speeds depending on the concentration of cohesin loaded⁷⁰. In agreement with this model, analysis of loops re-established after cohesin removal and restoration shows that those containing more of the cohesin loader NIPBL recover sooner than loops containing lower levels of this protein¹⁸. It is tempting to speculate that the greater amount of cohesin loaded on chromatin results in an increased diffusion rate and extrusion of the loops. However, the correlation between recovery time and the number of NIPBL binding sites may also be explained by a higher probability of cohesin loading at an early stage in the recovery period, or NIPBL may improve loop enlargement in other ways. Thus, whether cohesin diffusion can explain loop extrusion without additional help is unclear and requires additional experimental evidence.

SMC complexes possess ATPase activity, and it is possible that they act as their own motors to translocate chromatin using energy from ATP (FIG. 3c). The condensin complex was shown to move *in vitro* in a randomly chosen but single direction⁵⁴. This unidirectional movement requires the ATPase domain, indicating that energy consumption is necessary for motor activity⁵⁴. It was calculated that condensin moves 30 bp per ATP hydrolysed⁵⁴. The average CTCF loop size in humans is 180 kb (REF.¹⁶), which would require ~6,000 molecules of ATP per loop, assuming that *in vitro* studies are an accurate representation of *in vivo* function. When considering an upper estimate of 50,000 CTCF loops in the genome^{16,44}, forming all of them once would require an average of $\sim 3 \times 10^8$ molecules of ATP. This energy requirement is approximately in line with other energy estimates suggesting the need for 4×10^5 molecules of ATP per second for the movement of a cell or 2×10^7 molecules of ATP per second for the duplication of its proteome once every 24 hours⁷¹. Therefore, the energy requirement to form CTCF loops by extrusion powered by the cohesin ATPase activity is consistent with other energy requirements of the cell. However, the estimate of energy needed would increase substantially when considering the idea that loops are continuously extruding and/or have multiple extrusion complexes per loop⁴⁶. It should also be noted that these measurements of

motor activity were performed *in vitro* on naked DNA using the condensin complex from yeast, and it is possible that mammalian cohesin may have different rates of processing *in vivo*. Indeed, there is evidence that proteins bound to DNA slow down or block cohesin movement *in vitro*^{50,51}. Determining whether cohesin acts as its own motor and if the cell devotes the required energy for loop extrusion is a crucial issue addressed by recent results suggesting that ATP is required for CTCF loop formation⁷². In these experiments, cells were initially depleted of cohesin to eliminate CTCF loops, and then cohesin was allowed to recover under normal conditions or after depletion of ATP. Cells under normal conditions were able to re-establish CTCF loops, whereas those depleted of ATP were not, indicating that CTCF loop formation requires ATP, probably during the extrusion process⁷².

Cohesin may also be pushed along chromatin by unknown translocating factors, one of which may be RNAPII (FIG. 3d). Movement of cohesin as a consequence of the transcription process has been shown *in vitro*^{50,51}. Also, induction of transcription *in vivo* relocates cohesin to the 3' end of convergently oriented genes in yeast^{73,74}. In bacteria, transcription affects progress of the SMC complex as it zips up the chromosome^{23,69}. Changing the orientation of transcription to be contrary to the zip-up direction antagonizes the progress of SMC and influences interactions^{23,69}. Thus, RNAPII may push the SMC complex and thereby influence chromatin organization. In mammals, cohesin is also thought to be pushed by transcription^{51,75}. Deletion of the cohesin release factor WAPL in combination with deletion of CTCF results in cohesin relocation to the 3' end of genes, similar to what was shown in yeast⁷⁵. Cohesin relocation by transcription could be due to direct pushing by RNAPII⁷³ or could occur indirectly through chromatin supercoiling⁷⁶. In support of this, it was recently found that inhibition of transcription elongation by flavopiridol results in a moderate decrease in CTCF looping, although this effect is not as strong as that of ATP depletion⁷². In spite of this evidence, loop extrusion via RNAPII thrusting fails to explain how loops form in inactive regions of the genome. Additionally, the speed of loop extrusion, which is at least 374 bp per second, as discussed above, does not fit with current estimates that place RNAPII elongation rates at 9–90 bp per second^{18,77}. These issues question the feasibility of transcription as the driver of loop extrusion. Alternatively, instead of

pushing cohesin, the slow elongation rate may suggest that RNA polymerase may interfere with cohesin movement and thereby slow down the extrusion process over transcriptionally active regions.

Although it is unclear which of these models best describes loop extrusion (FIG. 3), analyses directed towards understanding the translocation speed, energy consumption and relationship between transcription and SMC complex movement will be informative in deciding among them. Although there is some evidence supporting each model, each has its own limitations, and it is possible that a combination of these mechanisms underlies the extrusion process. This process, if continuous and random, gives rise to CTCF loops at locations where extrusion is stopped by this protein while bringing together sequences located within the same or different compartmental domains. Forces underlying the extrusion process must coexist with those responsible for interactions among compartmental domains in the same transcriptional state. Therefore, these two types of domains must influence each other, a matter that we discuss next.

Interplay of 3D organizational drivers

Complete depletion of CTCF results in embryonic lethality in mice⁷⁸, but the use of auxin-mediated degradation approaches has allowed several recent studies to examine its role in 3D chromatin organization (FIG. 4). After depletion, CTCF loops disappear, yet interactions among and within compartmental domains remain^{58,63,64} (FIG. 4b). Indeed, CTCF loops and compartmental domains seem to form independently of each other, supporting a model where chromosomes are organized by two distinct but partially interdependent features, compartmental domains and CTCF loops.

Like CTCF, cohesin is also important for loop formation. Loss of the NIPBL–MAU2 cohesin loader (FIG. 3a) reduces cohesin loading onto DNA^{37,59}. This does not affect CTCF binding to CTCF sites but results in a widespread loss of CTCF loops. In contrast to CTCF loss, NIPBL–MAU2 deficiency results in stronger segregation between A and B compartmental domains^{57,59}. In similar studies, depletion of the RAD21 subunit of cohesin (FIG. 3a) or deletion of other subunits causes a widespread loss of CTCF loops^{18,58,60}. Unlike CTCF depletion, but similar to the loss of NIPBL, the depletion of RAD21 results in increased segregation of active and inactive regions into compartmental domains^{18,58}. This is manifested by more defined squares in the Hi-C checkerboard pattern (FIG. 4c). The different results

obtained after depletion of CTCF or cohesin suggest that cohesin mediates interactions other than those involved in CTCF–CTCF contacts. In CTCF-depleted cells, cohesin is likely to be able to extrude randomly and thereby may prevent complete segregation of compartmental domains. Indeed, inducible deletion of CTCF results in a widespread relocation of cohesin away from CTCF motifs⁷⁵. A few recent studies also examined what happens to chromatin organization when cohesin is blocked from disassociating from chromatin by depleting the cohesin release factors WAPL or PDS5 (FIG. 3a). Loss of either of these two proteins causes no dramatic change to existing CTCF loops, but new loops are formed spanning larger distances than those present in wild-type cells^{58–60}. This suggests that the residency time of cohesin determines loop size. Thus, extrusion may be only partially blocked, or perhaps stalled, by each convergently oriented CTCF site, and extrusion may continue until the DNA is released by WAPL or PDS5. Interestingly, the increased residency time of cohesin after WAPL or PDS5 depletion results in an increase in short-range contacts and a decrease in long-range interactions between compartmental domains^{58,59} (FIG. 4d). Without CTCF, the length and genomic location of extruded regions will depend upon the relative ratio of loading and unloading by NIPBL and WAPL, respectively. Loading and unloading of cohesin are thought to occur randomly, which would cause the formation of random loops in each individual cell. Because there is no precise extrusion stopping point without CTCF, these loops would appear as random signal when averaged across the cell population. When NIPBL or RAD21 are depleted, the loss of extrusion-mediated local interactions could allow compartmental interactions to more easily occur over long distances. Inversely, when WAPL or PDS5 are depleted, increased local interactions as a result of continuous extrusion may reduce the ability of long-range inter-compartmental domain interactions to take place (FIG. 4d).

3D organization and gene regulation

Previous studies indicate that CTCF and cohesin play two functionally distinct roles: to help enhancers find their cognate promoters and to restrict enhancer–promoter interactions between sequences located inside and outside CTCF loops^{79,80}. These two apparently opposite functions of CTCF can be explained on the basis of the requirement of extrusion to form

CTCF loops. CTCF represents a barrier to loop extrusion, and the formation of a loop domain results in increased interactions within the loop (FIG. 5a). The involvement of cohesin in the extrusion process explains its requirement for interactions not only between the anchors but also within loop domains^{18,57–59}. Therefore, enhancers and promoters that lie in the interior of the loop may interact more frequently with each other than in the absence of a loop. The extrusion process also explains why enhancers and promoters that lie on opposite sides of a loop anchor are less likely to interact (FIG. 5a). However, the extrusion process that results in the formation of CTCF loops does not completely preclude interactions between sequences on opposite sides of a loop anchor. Active compartmental domains located inside a loop can partially escape CTCF loops and interact with sequences in other active domains located outside of the CTCF loop¹⁷ (FIG. 5a). Importantly, because cohesin may move past CTCF anchors at a low frequency⁶⁸, this could explain why long-range interactions between compartmental domains are able to escape the constraints of CTCF loops.

These observations suggest a crucial role for CTCF in gene expression. However, the effects of CTCF removal on transcription can be quite variable, depending on the situation. Complete loss of CTCF is lethal during embryonic development, whereas haploinsufficiency results in intellectual disability, microcephaly and growth retardation^{78,81,82}. Heterozygous CTCF-knockout mice show a high incidence of tumours, and mutation of specific CTCF binding sites correlates with various cancers in humans^{83,84}. Accordingly, changes in CTCF looping at specific genomic sites have effects on the expression of nearby genes. For example, deletion of CTCF motifs at the *HoxA* locus results in increased interactions between active regions and genes that are normally repressed in motor neurons. In agreement with a model where the formation of CTCF loop domains increases interactions between enhancers and promoters, this increase in interactions corresponds to a large increase in gene expression⁸⁵. Similarly, CRISPR-mediated inversion of individual CTCF sites at the protocadherin alpha (*Pcdha*) locus results in the loss of interactions between the HS5-1 enhancer and *Pcdha* promoters, with a corresponding decrease in gene expression. This inversion also results in a gain of interactions between the HS5-1 enhancer

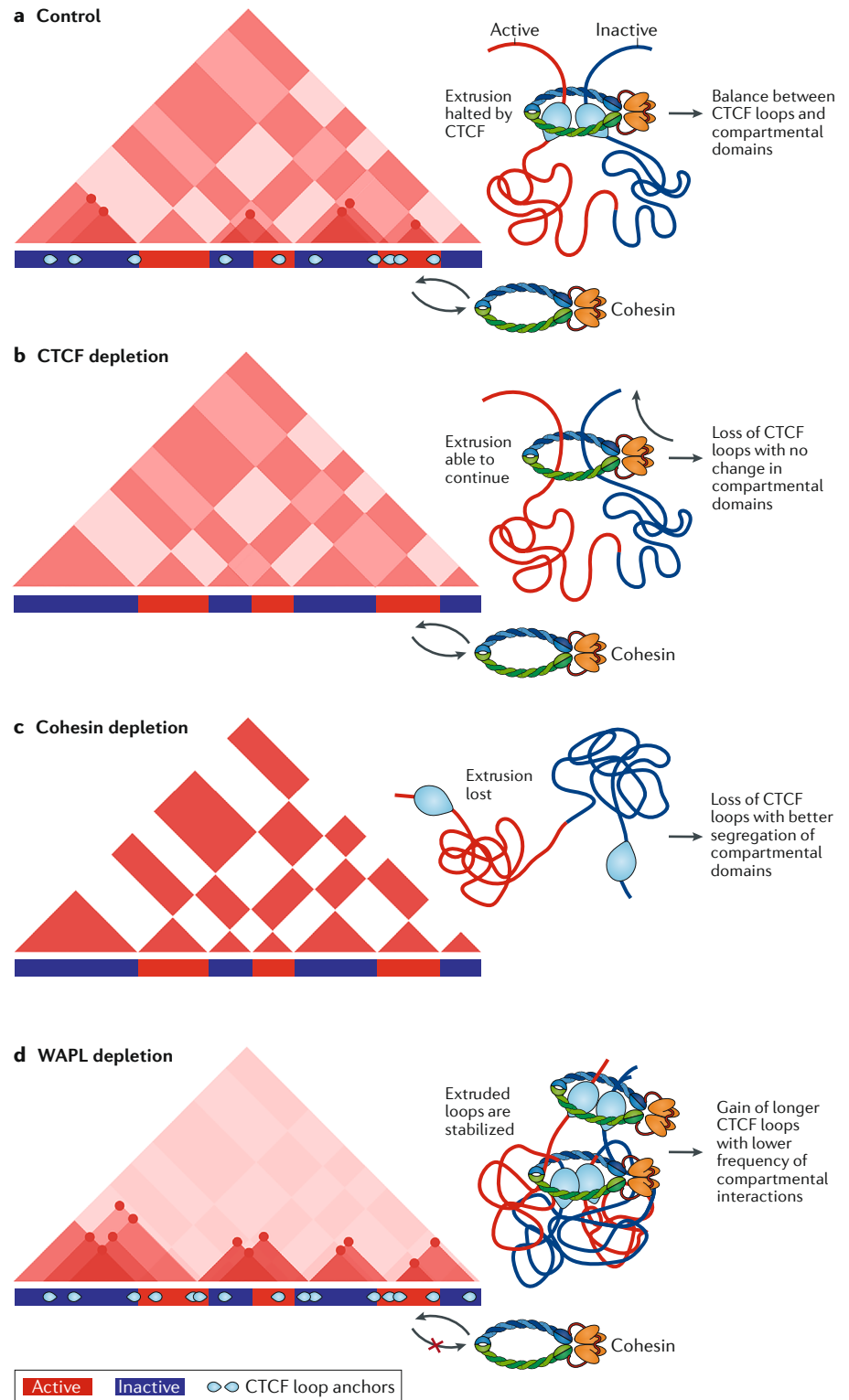


Fig. 4 | Effects of CTCF, cohesin or WAPL depletion on 3D chromatin organization. **a** | Chromatin is organized in the 3D nuclear space by CTCF loops and compartmental domains. Some CTCF loops restrict the ability of active (red) and inactive (blue) regions to segregate into compartmental domains, whereas others increase the frequency of interactions between two adjacent active and inactive domains (right). **b** | Depletion of CTCF results in a loss of CTCF loops but no change in compartmental interactions, probably because cohesin is able to randomly continue extruding chromatin^{58,63,64}. **c** | Depletion of cohesin results in loss of CTCF loops, more distinct compartmental domains and stronger inter-compartmental interactions^{18,57–60}. **d** | Depletion of WAPL results in gain of longer CTCF loops and decreases interactions among compartmental domains^{58–60}.

and genes in the protocadherin beta (*Pcdhb*) locus. However, instead of a gain in gene expression corresponding to the increased interactions with an active enhancer, *Pcdhb* genes display decreased gene expression⁴⁴. In a different study, CRISPR inversion of individual CTCF sites results in changes in interactions but mild changes (~1.5-fold to 2.5-fold) in gene expression⁸⁶. In the context of these findings, it is surprising that general depletion of CTCF or cohesin in various cell types under cultured conditions has very small immediate consequences on transcription, although cells depleted of CTCF die after 4 days in culture⁶³. For example, knockdown of CTCF in HEK293T cells results in only 161 differentially expressed genes⁸⁷. This small effect may have been influenced by the residual amount of CTCF remaining after

knockdown. However, nearly complete depletion of CTCF via auxin degradation results in only 370 differentially expressed genes, of which only 43 show at least a fivefold change in expression⁶³. A separate study in which CTCF loops were eliminated by RAD21 degradation found only two genes with at least a fivefold change in gene expression¹⁸. These findings suggest that gene expression is often surprisingly resilient to acute changes in CTCF loops or cohesin-mediated extrusion. Reconciling the drastic consequences of CTCF depletion on phenotypes in living organisms compared with the minor immediate changes to gene expression observed in cultured cell lines will be essential in understanding the role of this protein in transcription.

In addition to CTCF, other proteins present on chromatin may also affect

progression of the cohesin ring and thereby influence chromatin interactions. Experiments measuring translocation of cohesin in vitro find that the process is hindered by DNA-bound proteins, including nucleosomes^{50,51}. The degree of interference with translocation is directly related to the size of the protein or complex, probably owing to the difficulty in passing through the cohesin ring⁵⁰. This suggests that large complexes of transcription factors could present barriers to loop extrusion where the cohesin ring slows down but not to the same extent as CTCF sites. This may partially explain enrichment of interactions within the interior of CTCF loop domains. Although the cohesin ring may form a somewhat stable interaction with CTCF, the extrusion process is likely to be recurrent, with several rings constantly extruding along the same loop⁴⁶. This concept has important mechanistic consequences for gene expression. Continuous extrusion results in frequent interactions between sequences within the loop, contributing to the intensity of Hi-C signal in the interior of loop domains. Although these interactions are in principle random, transient retention of cohesin at sites of large protein complexes may help bring these sequences together and increase their interaction frequency. For example, large protein complexes bound to enhancers or promoters such as Mediator, or protein complexes bound to histone modifications, may help enhancers and promoters located within the loop contact each other. Additionally, retention of the extrusion complex at enhancers may depend on specific subunits of cohesin. For example, it was recently shown that cohesin present at CTCF loop anchors can contain either the SA1 or SA2 subunits, whereas cohesin complexes at non-CTCF sites contain almost exclusively SA2⁸⁸. The existence of non-CTCF cohesin sites, which are often at enhancers, suggests that different populations of cohesin may have distinct retention properties during extrusion. Brief cohesin retention, or extrusion stalling, at enhancers may explain the enhancer-dependent bursting of transcription activation, as genes are activated each time cohesin-mediated extrusion brings promoters in contact with enhancers⁸⁹ (FIG. 5b). However, this model fails to explain why acute depletion of RAD21 only has a minor effect on gene expression. Recently, it was found that many enhancer–promoter interactions are dependent on Yin Yang 1 (YY1)^{90,91}. These enhancers are enriched with both

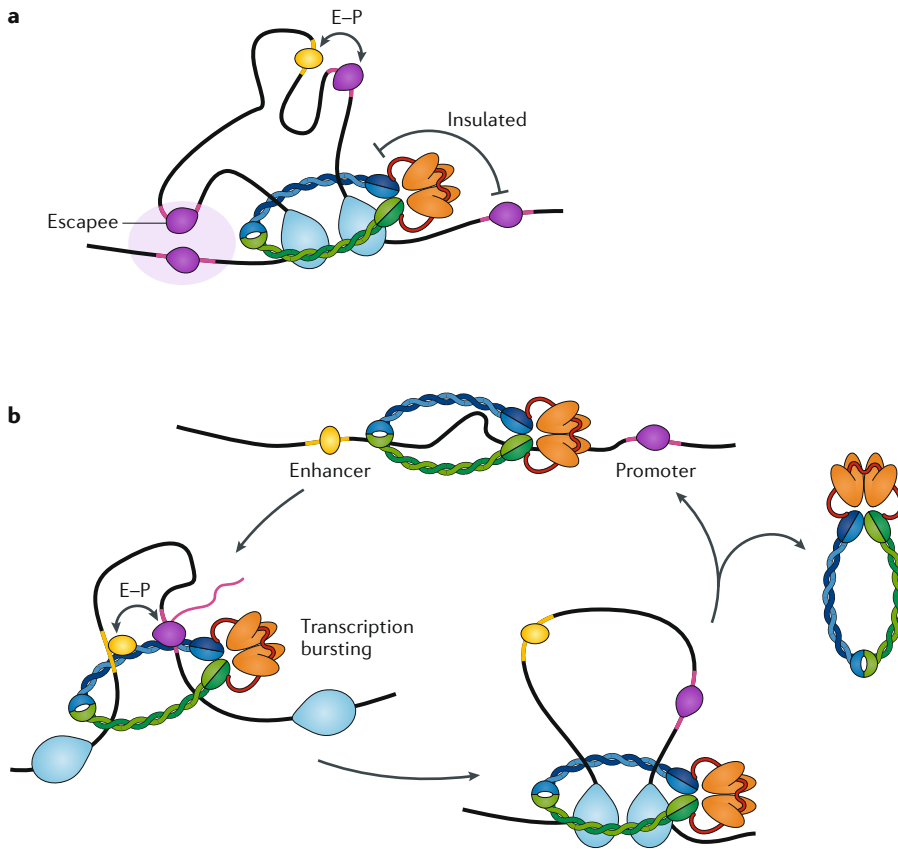


Fig. 5 | CTCF loops and enhancer–promoter interactions. a | CTCF loops establish domains in which sequences can interact more frequently. These contacts are thought to help promote enhancer–promoter (E–P) interactions when inside the domain but help insulate against those outside the domain^{79,80}. Enhancers are shown in yellow bound by a yellow transcription factor, and promoters are shown in pink bound by a purple RNA polymerase II (RNAPII). However, examples of genes that escape the CTCF domain and interact with adjacent sequences can be observed in Hi-C data¹⁷. It is likely that these ‘escapee’ genes interact with promoters or regulatory sequences within A compartmental domains (large light purple oval). **b** | A speculative model of transcriptional activation is shown. In this model, genes are inactive when extrusion has not begun (top) and are activated once extrusion brings together enhancers and promoters (left). Gene activity is lost once extrusion moves past the enhancer or promoter (right) but will be re-established during each extrusion event. Regular extrusion events causing gene activation at discrete times may explain transcriptional bursting.

YY1 and cohesin, and it is thought that YY1 may partially block cohesin-mediated loop extrusion to form enhancer–promoter interactions^{88–92}. Interestingly, YY1 is able to dimerize and enhance DNA interactions in vitro without requiring cohesin⁹⁰. Therefore, it is possible that cohesin-mediated extrusion may initiate enhancer–promoter interactions that are then partially stabilized by YY1 dimerization. This may explain the minor effect that acute cohesin degradation has on gene expression. A deeper exploration of this idea and the characterization of other proteins that are involved in blocking loop extrusion or in stabilizing enhancer–promoter interactions will be important in future work.

Interactions among A or B compartmental domains may also help stabilize active or silenced transcriptional states. Contacts among A compartments presumably take place through interactions among multivalent proteins present at enhancers and promoters, as well as RNAs and components of the splicing and termination machinery. These interactions may contribute to the co-regulation of genes bound by similar transcription factors and to an increase in the local concentration of the transcription machinery, resulting in the formation of structures similar to transcription factories. Interestingly, these structures are not stable, as they must be disrupted by the continuous extrusion via the cohesin complex, which increases interactions within loops while decreasing contacts between loops and compartmental domains.

Conclusion and future perspectives

Results discussed here suggest that the genomes of all organisms examined to date are organized into compartmental domains. These domains may represent the most basic form of 3D chromatin architecture, and they are established as a consequence of interactions among protein complexes associated with DNA sequences on the basis of their transcriptional or chromatin state. In vertebrates, an additional level of organization is established as a result of the extrusion process mediated by cohesin and perhaps also condensin. Stalling of extrusion by CTCF leads to the formation of stable loops that regulate enhancer–promoter interactions. It is possible that a subset of CTCF loops are common to all cells and are maintained in the germ line and early embryogenesis, in which case the resulting 3D architecture imposed by these loops could be

considered to regulate transcription rather than being a consequence of this process. However, CTCF is present at many sites in the genome that lack the CTCF motif, where it may be recruited by transcription factors, in which case the organization imposed by these loops would be a consequence of transcription. In addition to further exploring the relationship between 3D genome organization and transcription, improvement in the following three technical areas may lead to important advances the field.


Shrinking genome organization.

Improvements to the Hi-C methodology and a lower cost of sequencing have enhanced the resolution at which chromatin organization can be visualized. Continued improvements in Hi-C resolution may allow the understanding of the contribution of single genes or regulatory sequences to 3D chromatin organization. Hi-C maps also suggest the existence of simultaneous interactions among multiple loci in the genome, but this may be a consequence of the cell populations used to obtain most Hi-C data sets. Technical innovations that allow the sequencing of long reads should afford the visualization of possible multi-loop structures contributing to 3D organization and the understanding of their functional importance.

Visualizing loop extrusion. The extrusion process has not been directly measured or visualized for cohesin in mammalian cells. Full acceptance of this model will depend on obtaining direct evidence for the extrusion process in the context of CTCF loops in vivo. Although work in vitro and in bacterial systems using similar SMC complexes is promising, further work with mammalian cohesin on actual chromatin will be an important step forward.

Population versus single-cell chromatin organization. Information from thousands of single-cell Hi-C maps has been used to track the dynamics of CTCF loops and compartmental domains. The results suggest that features of chromatin organization may vary substantially between individual nuclei⁶. However, single-cell Hi-C studies have been limited by the coverage, and therefore resolution, achievable for these Hi-C maps^{2,6,93}. Microscopy methods such as Oligopaint with STORM imaging are approaching the resolution at which domain structures can be visualized in single cells⁴. Thus, a combination of genomics

and microscopy approaches may be useful in examining single-cell chromatin organization.

M. Jordan Rowley and Victor G. Corces *

Department of Biology, Emory University, Atlanta, GA, USA.

*e-mail: vgcorces@gmail.com

<https://doi.org/10.1038/s41576-018-0060-8>

Published online: 26 October 2018

- Fraser, J., Williamson, I., Bickmore, W. A. & Dostie, J. An overview of genome organization and how we got there: from FISH to Hi-C. *Microbiol. Mol. Biol. Rev.* **79**, 347–372 (2015).
- Stevens, T. J. et al. 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature* **544**, 59–64 (2017).
- Nagano, T. et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **502**, 59–64 (2013).
- Beliveau, B. J. et al. Single-molecule super-resolution imaging of chromosomes and in situ haplotype visualization using Oligopaint FISH probes. *Nat. Commun.* **6**, 7147 (2015).
- Ni, Y. et al. Super-resolution imaging of a 2.5 kb non-repetitive DNA in situ in the nuclear genome using molecular beacon probes. *eLife* **6**, e21660 (2017).
- Nagano, T. et al. Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature* **547**, 61–67 (2017).
- Pegoraro, G. & Misteli, T. High-throughput imaging for the discovery of cellular mechanisms of disease. *Trends Genet.* **33**, 604–615 (2017).
- Lieberman-Aiden, E. et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).
- Lajoie, B. R., Dekker, J. & Kaplan, N. The hitchhiker's guide to Hi-C analysis: practical guidelines. *Methods* **72**, 65–75 (2015).
- Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
- Sexton, T. et al. Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148**, 458–472 (2012).
- Hou, C., Li, L., Qin, Z. S. & Corces, V. G. Gene density, transcription, and insulators contribute to the partition of the *Drosophila* genome into physical domains. *Mol. Cell* **48**, 471–484 (2012).
- Nora, E. P. et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012).
- Bonev, B. & Cavalli, G. Organization and function of the 3D genome. *Nat. Rev. Genet.* **17**, 661–678 (2016).
- Schmitt, A. D., Hu, M. & Ren, B. Genome-wide mapping and analysis of chromosome architecture. *Nat. Rev. Mol. Cell. Biol.* **17**, 743–755 (2016).
- 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).
- Rowley, M. J. et al. Evolutionarily Conserved Principles Predict 3D Chromatin Organization. *Mol. Cell* **67**, 837–852 (2017).
- Rao, S. et al. Cohesin loss eliminates all loop domains, leading to links among superenhancers and downregulation of nearby genes. Preprint at *BioRxiv* <https://doi.org/10.1101/159782> (2017).
- Dong, P. et al. 3D chromatin architecture of large plant genomes determined by local A/B compartments. *Mol. Plant* **10**, 1497–1509 (2017).
- Haddad, N., Jost, D. & Vaillant, C. Perspectives: using polymer modeling to understand the formation and function of nuclear compartments. *Chromosome Res.* **25**, 35–50 (2017).
- Huang, J., Marco, E., Pinello, L. & Yuan, G.-C. Predicting chromatin organization using histone marks. *Genome Biol.* **16**, 162 (2015).
- Di Pierro, M., Cheng, R. R., Lieberman Aiden, E., Wolyne, P. G. & Onuchic, J. N. De novo prediction of human chromosome structures: epigenetic marking patterns encode genome architecture. *Proc. Natl Acad. Sci. USA* **114**, 12126–12131 (2017).
- Wang, X., Brandão, H. B., Le, T. B. K., Laub, M. T. & Rudner, D. Z. *Bacillus subtilis* SMC complexes

- juxtacpose chromosome arms as they travel from origin to terminus. *Science* **355**, 524–527 (2017).
24. Le, T. B. K., Imakaev, M. V., Mirny, L. A. & Laub, M. T. High-resolution mapping of the spatial organization of a bacterial chromosome. *Science* **342**, 731–734 (2013).
 25. Li, L. et al. Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing. *Mol. Cell* **58**, 216–231 (2015).
 26. Hug, C. B., Grimaldi, A. G., Kruse, K. & Vaquerizas, J. M. Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription. *Cell* **169**, 216–228 (2017).
 27. Bensaude, O. Inhibiting eukaryotic transcription. Which compound to choose? How to evaluate its activity?: Which compound to choose? How to evaluate its activity? *Transcription* **2**, 103–108 (2011).
 28. Kaaij, L. J. T., van der Weide, R. H., Ketting, R. F. & de Wit, E. Systemic loss and gain of chromatin architecture throughout zebrafish development. *Cell Rep.* **24**, 1–10 (2018).
 29. Du, Z. et al. Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature* **547**, 232–235 (2017).
 30. Ke, Y. et al. 3D chromatin structures of mature gametes and structural reprogramming during mammalian embryogenesis. *Cell* **170**, 367–381 (2017).
 31. Naumova, N. et al. Organization of the mitotic chromosome. *Science* **342**, 948–953 (2013).
 32. Jung, Y. H. et al. Chromatin states in mouse sperm correlate with embryonic and adult regulatory landscapes. *Cell Rep.* **18**, 1366–1382 (2017).
 33. Battulin, N. et al. Comparison of the three-dimensional organization of sperm and fibroblast genomes using the Hi-C approach. *Genome Biol.* **16**, 77 (2015).
 34. Cubeñas-Potts, C. & Corces, V. G. Architectural proteins, transcription, and the three-dimensional organization of the genome. *FEBS Lett.* **589**, 2923–2930 (2015).
 35. Cubeñas-Potts, C. et al. Different enhancer classes in *Drosophila* bind distinct architectural proteins and mediate unique chromatin interactions and 3D architecture. *Nucleic Acids Res.* **45**, 1714–1730 (2016).
 36. Harlen, K. M. & Churchman, L. S. The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. *Nat. Rev. Mol. Cell Biol.* **18**, 263–273 (2017).
 37. Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K. & Sharp, P. A. A. Phase separation model for transcriptional control. *Cell* **169**, 13–23 (2017).
 38. Larson, A. G. et al. Liquid droplet formation by HP1 α suggests a role for phase separation in heterochromatin. *Nature* **547**, 236–240 (2017).
 39. Lin, Y., Currie, S. L. & Rosen, M. K. Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs. *J. Biol. Chem.* **292**, 19110–19120 (2017).
 40. van der Lee, R. et al. Classification of intrinsically disordered regions and proteins. *Chem. Rev.* **114**, 6589–6631 (2014).
 41. Lu, H. et al. Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. *Nature* **558**, 318–323 (2018).
 42. Brackley, C. A., Johnson, J., Kelly, S., Cook, P. R. & Marenduzzo, D. Simulated binding of transcription factors to active and inactive regions folds human chromosomes into loops, rosettes and topological domains. *Nucleic Acids Res.* **44**, 3503–3512 (2016).
 43. Tang, Z. et al. CTCF-mediated human 3D genome architecture reveals chromatin topology for transcription. *Cell* **163**, 1611–1627 (2015).
 44. Guo, Y. et al. CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell* **162**, 900–910 (2015).
 45. 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).
 46. Fudenberg, G. et al. Formation of chromosomal domains by loop extrusion. *Cell Rep.* **15**, 2038–2049 (2016).
 47. Nichols, M. H. & Corces, V. G. A. CTCF code for 3D genome architecture. *Cell* **162**, 703–705 (2015).
 48. Nasmyth, K. Disseminating the genome: joining, resolving, and separating sister chromatids during mitosis and meiosis. *Annu. Rev. Genet.* **35**, 673–745 (2001).
 49. Alipour, E. & Marko, J. F. Self-organization of domain structures by DNA-loop-extruding enzymes. *Nucleic Acids Res.* **40**, 11202–11212 (2012).
 50. Stigler, J., Camdere, G. Ö., Koshland, D. E. & Greene, E. C. Single-molecule imaging reveals a collapsed conformational state for DNA-bound cohesin. *Cell Rep.* **15**, 988–998 (2016).
 51. Davidson, I. F. et al. Rapid movement and transcriptional re-localization of human cohesin on DNA. *EMBO J.* **35**, 2671–2685 (2016).
 52. Kanke, M., Tahara, E., Huis In't Veld, P. J. & Nishiyama, T. Cohesin acetylation and Wapl-Pds5 oppositely regulate translocation of cohesin along DNA. *EMBO J.* **35**, 2686–2698 (2016).
 53. Ganji, M. et al. Real-time imaging of DNA loop extrusion by condensin. *Science* **360**, 102–105 (2018).
 54. Terakawa, T. et al. The condensin complex is a mechanochemical motor that translocates along DNA. *Science* **358**, 672–676 (2017).
 55. Murayama, Y., Samora, C. P., Kurokawa, Y., Iwasaki, H. & Uhlmann, F. Establishment of DNA-DNA interactions by the cohesin ring. *Cell* **172**, 465–477 (2018).
 56. Nagy, G. et al. Motif oriented high-resolution analysis of ChIP-seq data reveals the topological order of CTCF and cohesin proteins on DNA. *BMC Genomics* **17**, 637 (2016).
 57. Schwarzer, W. et al. Two independent modes of chromatin organization revealed by cohesin removal. *Nature* **551**, 51–56 (2017).
 58. 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).
 59. Haarhuis, J. H. I. et al. The cohesin release factor WAPL restricts chromatin loop extension. *Cell* **169**, 693–707 (2017).
 60. Gassler, J. et al. A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. *EMBO J.* **36**, 3600–3618 (2017).
 61. Fu, Y., Sinha, M., Peterson, C. L. & Weng, Z. The insulator binding protein CTCF positions 20 nucleosomes around its binding sites across the human genome. *PLoS Genet.* **4**, e1000138 (2008).
 62. Mawhinney, M. T. et al. CTCF-induced circular DNA complexes observed by atomic force microscopy. *J. Mol. Biol.* **430**, 759–776 (2018).
 63. Nora, E. P. et al. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* **169**, 930–944 (2017).
 64. Kubo, N. et al. Preservation of chromatin organization after acute loss of CTCF in mouse embryonic stem cells. Preprint at *BioRxiv*. <https://doi.org/10.1101/118737> (2017).
 65. Phanstiel, D. H. et al. Static and dynamic DNA loops form AP-1-bound activation hubs during macrophage development. *Mol. Cell* **67**, 1037–1048 (2017).
 66. Bonev, B. et al. Multiscale 3D genome rewiring during mouse neural development. *Cell* **171**, 557–572 (2017).
 67. Gibcus, J. H. et al. A pathway for mitotic chromosome formation. *Science* **359**, eaao6135 (2018).
 68. 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).
 69. Tran, N. T., Laub, M. T. & Le, T. B. K. SMC progressively aligns chromosomal arms in *Caulobacter crescentus* but is antagonized by convergent transcription. *Cell Rep.* **20**, 2057–2071 (2017).
 70. Brackley, C. A. et al. Nonequilibrium chromosome looping via molecular slip links. *Phys. Rev. Lett.* **119**, 138101 (2017).
 71. Flamholz, A., Phillips, R. & Milo, R. The quantified cell. *Mol. Biol. Cell* **25**, 3497–3500 (2014).
 72. Vian, L. et al. The energetics and physiological impact of cohesin extrusion. *Cell* **173**, 1165–1178 (2018).
 73. Ocampo-Hafalla, M., Muñoz, S., Samora, C. P. & Uhlmann, F. Evidence for cohesin sliding along budding yeast chromosomes. *Open Biol.* **6**, 150178 (2016).
 74. Bausch, C. et al. Transcription alters chromosomal locations of cohesin in *Saccharomyces cerevisiae*. *Mol. Cell Biol.* **27**, 8522–8532 (2007).
 75. Busslinger, G. A. et al. Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature* **544**, 503–507 (2017).
 76. Racko, D., Benedetti, F., Dorier, J. & Stasiak, A. Transcription-induced supercoiling as the driving force of chromatin loop extrusion during formation of TADs in interphase chromosomes. *Nucleic Acids Res.* **46**, 1648–1660 (2017).
 77. Jonkers, I. & Lis, J. T. Getting up to speed with transcription elongation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* **16**, 167–177 (2015).
 78. Moore, J. M. et al. Loss of maternal CTCF is associated with peri-implantation lethality of Cctf null embryos. *PLoS ONE* **7**, e34915 (2012).
 79. Downen, J. M. et al. Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* **159**, 374–387 (2014).
 80. Hnisz, D., Day, D. S. & Young, R. A. Insulated neighborhoods: structural and functional units of mammalian gene control. *Cell* **167**, 1188–1200 (2016).
 81. Wan, L.-B. et al. Maternal depletion of CTCF reveals multiple functions during oocyte and preimplantation embryo development. *Development* **135**, 2729–2738 (2008).
 82. Gregor, A. et al. De novo mutations in the genome organizer CTCF cause intellectual disability. *Am. J. Hum. Genet.* **93**, 124–131 (2013).
 83. Kemp, C. J. et al. CTCF haploinsufficiency destabilizes DNA methylation and predisposes to cancer. *Cell Rep.* **7**, 1020–1029 (2014).
 84. Katainen, R. et al. CTCF/cohesin-binding sites are frequently mutated in cancer. *Nat. Genet.* **47**, 818–821 (2015).
 85. Narendra, V. et al. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science* **347**, 1017–1021 (2015).
 86. de Wit, E. et al. CTCF binding polarity determines chromatin looping. *Mol. Cell* **60**, 676–684 (2015).
 87. Zuin, J. et al. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc. Natl Acad. Sci. USA* **111**, 996–1001 (2014).
 88. Kojic, A. et al. Distinct roles of cohesin-SA1 and cohesin-SA2 in 3D chromosome organization. *Nat. Struct. Mol. Biol.* **25**, 496–504 (2018).
 89. Fukaya, T., Lim, B. & Levine, M. Enhancer control of transcriptional bursting. *Cell* **166**, 358–368 (2016).
 90. Weintraub, A. S. et al. YY1 is a structural regulator of enhancer-promoter loops. *Cell* **171**, 1573–1588 (2017).
 91. 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).
 92. Pan, X. et al. YY1 controls Ig κ repertoire and B cell development, and localizes with condensin on the Ig κ locus. *EMBO J.* **32**, 1168–1182 (2013).
 93. Flyamer, I. M. et al. Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature* **544**, 110–114 (2017).
 94. Gligoris, T. & Löwe, J. Structural insights into ring formation of cohesin and related SMC complexes. *Trends Cell Biol.* **26**, 680–693 (2016).

Acknowledgements

Work in the authors' laboratory is supported by US Public Health Service Award R01 GM035463 (V.G.C.) and Pathway to Independence Award K99/R00 GM127671 (M.J.R.) from the US National Institutes of Health (NIH). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Author contributions

M.J.R. was involved in researching data for the article. Both authors were responsible for discussing content, writing the article and reviewing and/or editing the manuscript before submission.

Competing interests

The authors declare no competing interests.

Publisher's note

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