Briefings in Functional Genomics, 19(2), 2020, 128–138

doi: 10.1093/bfgp/elz028
Advance Access Publication Date: 6 February 2020
Review paper

Order and disorder: abnormal 3D chromatin organization in human disease

Chiara Anania and Darío G. Lupiáñez

Corresponding author: Darío G. Lupiáñez. E-mail: Dario.Lupianez@mdc-berlin.de

Abstract

A precise three-dimensional (3D) organization of chromatin is central to achieve the intricate transcriptional patterns that are required to form complex organisms. Growing evidence supports an important role of 3D chromatin architecture in development and delineates its alterations as prominent causes of disease. In this review, we discuss emerging concepts on the fundamental forces shaping genomes in space and on how their disruption can lead to pathogenic phenotypes. We describe the molecular mechanisms underlying a wide range of diseases, from the systemic effects of coding mutations on 3D architectural factors, to the more tissue-specific phenotypes resulting from genetic and epigenetic modifications at specific loci. Understanding the connection between the 3D organization of the genome and its underlying biological function will allow a better interpretation of human pathogenesis.

Key words: 3D chromatin organization; epigenetics; disease; long-range gene regulation

Introduction

The spatial organization of eukaryotic chromatin represents a crucial step that connects linear genomic information with its biological function [1]. As such, 3D genomic architecture organization is linked to a wide range of cellular processes, such as DNA replication, maintenance and transcription. Far from random, chromatin organization follows certain principles that reflect the necessity to connect physically distant functional regions of the genome, such as regulatory elements and promoters, to execute precise transcriptional programs [2]. Recent discoveries highlight the importance of this structural order in orchestrating embryonic development and its disruption as a driver of human disease [3, 4].

Advances in microscopy-based approaches and the development of chromosome conformation capture (3C) techniques, including Hi-C, enabled the study of chromatin interactions at an unprecedented resolution [5, 6], revealing a hierarchical organization that operates at multiple levels and scales. In vertebrates, chromosomes tend to occupy distinct nuclear territories, reflecting positional preferences that correlate with the increased frequency observed for recurrent translocational events between certain genomic locations [7, 8]. At a sub-chromosomal level, however, chromatin is segregated into two distinct multi-megabase compartments [9]. On the one hand, A compartments are characterized by open chromatin, enriched on active histone marks as well as on genes displaying high levels of transcription. On the other hand, B compartments usually contain inactive genes and are linked to repressive histone modifications. The described associations denote the dynamic nature of chromatin compartmentalization, a property also highlighted by the significant degree of 'compartment switching' observed upon cellular differentiation or reprogramming that is linked to gene expression changes at corresponding loci [10–15]. At a smaller scale, chromatin is organized into topologically associated domains (TADs), sub-megabase regions with increased self-interaction frequency, but relatively isolated from other neighboring domains [16, 17]. TAD insulation occurs

Chiara Anania is a PhD student at the Berlin Institute for Medical System Biology. She is interested in understanding the interplay between chromatin architecture and gene regulation in development and disease.
Darío G. Lupiáñez is a group leader at the Berlin Institute for Medical Systems Biology. His research focuses on the study of gene regulation in development, disease and evolution, with a special emphasis on 3D chromatin organization.

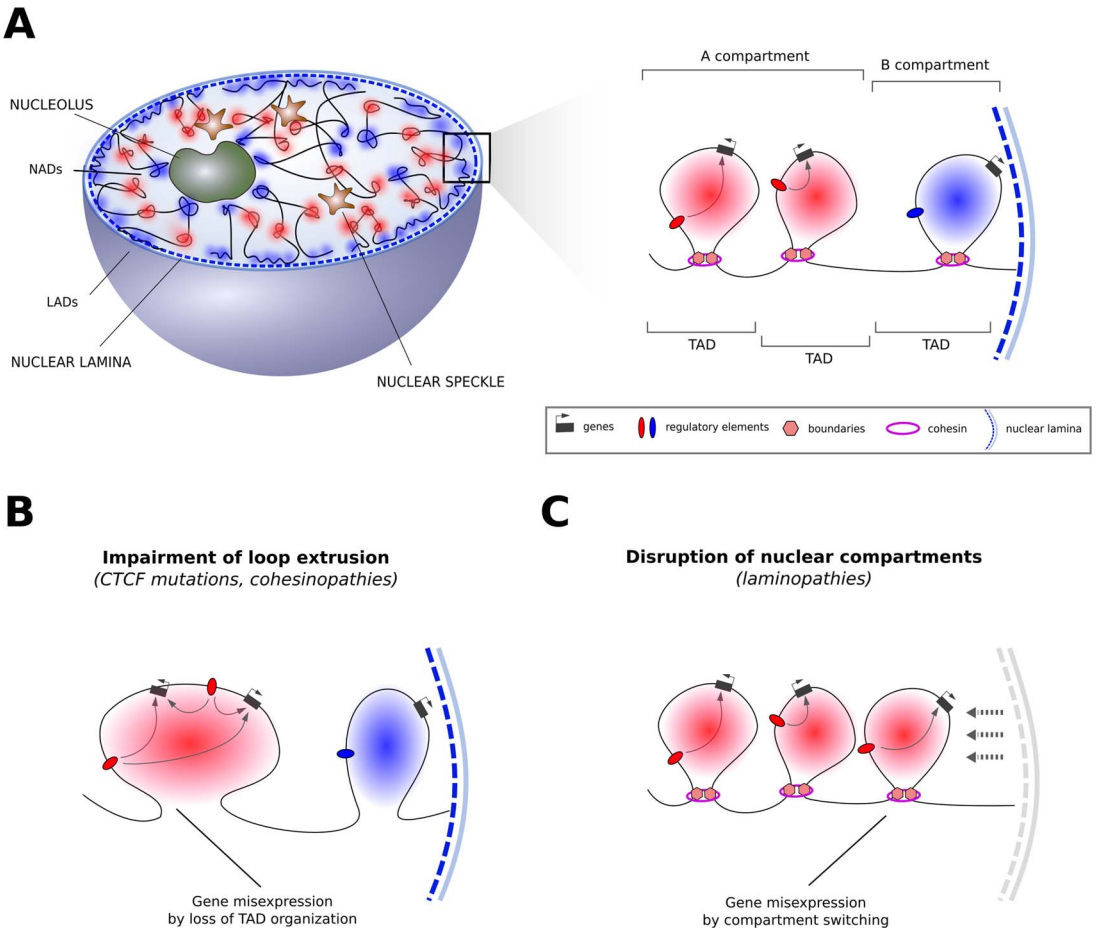

Figure 1. Factors organizing chromatin in the 3D space and effects of associated mutations. Within the nucleus, chromatin segregates into transcriptionally active, A compartments (red) or inactive, B compartments (blue), that also associate to specialized locations such as the NL, nuclear speckles or the nucleolus. Chromatin is further subdivided into TADs, mainly formed by loop extrusion and delimited by boundaries that are frequently associated to CTCF and the cohesin complex. The disruption of individual components of the loop extrusion machinery (CTCF or the cohesin complex) can induce gene misexpression by a global loss of TAD organization. In contrast, compartments are largely preserved. The disruption of nuclear compartments, like the NL can induce gene misexpression by compartment switching. In this case, TAD organization is overall preserved.

predominantly at boundary regions, which are enriched in housekeeping genes, tRNA and, most notably, in the CCCTC-binding factor (CTCF). In contrast to compartments, the position of TADs and boundaries is largely conserved between cell types and across species, delineating these organizational domains as a conserved scaffold that partitions vertebrate genomes [11, 16]. Overall, TADs correlate extensively with large domains that display a marked regulatory potential [18], suggestive of their capability to constrain but also facilitate enhancer-promoter interactions. An additional organizational level can be observed in the form of chromatin loops, defined as point-to-point interactions connecting genomic loci. Most chromatin loops are mediated by CTCF binding [12] and frequently delimit TAD boundaries, although they are also positioned within them. As occurs with TADs, CTCF-mediated loops generally display a constitutive nature, with a high degree of conservation between cell types and during differentiation. In contrast, CTCF-independent loops are more dynamic and associated to interactions that relate directly to the transcriptional process, such as enhancer-promoter or Polycomb-mediated contacts [10, 12]. In both cases, chromatin loops facilitate the further division of TADs into smaller and nested substructures denominated sub-TADs, with more dynamic and cell-to-cell variability than TADs [12, 19–22]. Understanding the precise interplay between these distinct hierarchical orders of chromatin organization is essential to interpret genomic function and to evaluate its role in cell homeostasis and developmental processes.

In this review, we discuss current concepts in 3D chromatin organization and contextualize them to illustrate the pathological effects of abnormal genome folding. We focus on mutations affecting essential factors of spatial genome architecture, highlighting the widespread consequences of coding mutations and the more locus-specific effects resulting from the alteration of genomic-binding sites. Furthermore, we describe how epigenetic alterations, such as DNA methylation, or how recently discovered mechanisms involving transposable elements (TEs) can also impact spatial genomic folding and cause aberrant phenotypes.

Breaking the rules: mutations in 3D architectural factors

Growing evidence delineates a model where, at least, two independent mechanisms organize vertebrate genomes in the nuclear space (Figure 1A). First, chromatin is segregated into A/B compartments reflecting the clustering of regions with similar epigenetic status, and suggesting an active involvement of the

transcriptional process in spatial genomic organization [23]. This notion is further supported by the existence of specialized regions within the nucleus, such as transcriptionally active nuclear speckles, or repressive environments like nucleolus-associating domains (NADs) or lamina-associating domains (LADs) [16, 24]. However, recent functional evidence suggests that is not transcription *per se*, but likely the components of the transcriptional machinery what underlies this type of organization [10, 23, 25–29].

A second organizational layer is found at the level of TADs and loops, which formation can be well explained by a chromatin-extrusion process [30, 31]. In this model, a loop-extruding factor, the ring-shaped cohesin complex, extrude interphase chromatin until encountering a physical barrier, in most cases chromatin-bound CTCFs, where the complex will be stalled until its dissociation. An important feature of the loop extrusion model is the orientation of CTCF-binding motifs, which is preferentially convergent between looped pairs [12, 32]. At TAD boundaries, CTCF sites with divergent orientations are frequently clustered. This evolutionary-conserved signature appears to facilitate the insulator effect observed at boundaries, by projecting interactions in opposite directions [32, 33].

Novel protein degradation methods revealed that the depletion of either CTCF or cohesin complex components cause a global impairment on the formation of TAD and loops, as observed in corresponding Hi-C maps. However, segregation into A/B compartments is largely unaffected or even reinforced [34–38]. This effect has been also denoted using an orthogonal approach that visualizes accessible chromatin using super-resolution microscopy (3D ATAC-PALM) [preprint: 39]. In this study, CTCF or cohesin depletion increased the clustering of accessible chromatin, which would be representative of A compartments. In both approaches, the observed effects in chromatin organization were largely reversed upon recovery of the protein levels, after degron removal. Altogether, these findings support the independence of both organizational forces, with compartment segregation as a default mechanism of chromatin organization where loop extrusion is superimposed. In these functional studies, however, the authors report that the impact of altered loop extrusion is followed by moderate transcriptional effects, with just a few hundreds of genes displaying significant changes [35, 36], thus raising concerns on the relevance of TADs and chromatin loops in controlling gene expression. It would be important to note that these studies, performed in cultured cell lines, only provide a limited view of the potential effects that loop-extrusion impairment might cause on an entire organism. The huge variety of cell and tissue types that arise during development might expose non TAD-isolated genes to a large number of regulatory elements with marked spatio-temporal activity patterns, thereby increasing exponentially the chances to induce misexpression leading to abnormal phenotypes. *In vivo* studies depleting cohesin or CTCF conditionally support this notion and demonstrate their important role in developmental processes such as inflammatory response [40] or remote memory [41]. In these cases, the observed phenotypes appear to be mediated by a selective dysregulation of genes associated with those processes. Therefore, a more detailed characterization of the effects of CTCF/cohesin depletion on distinct cell/tissue types and across multiple developmental time points would be necessary to precisely determine the impact that altered 3D spatial organization can exert on gene regulation. Nevertheless, mutations in the factors controlling loop extrusion, as well as in compartmentalization, can lead to a wide variety of pathogenic phenotypes that appear to be directly related to abnormal 3D chromatin folding, as described in the following sections.

Impairment of loop extrusion

As exposed above, the cohesin complex and CTCF have been identified as key factors controlling TAD and loop formation, through active chromatin extrusion [30, 31]. The impairment of this process has been shown to cause dramatic and widespread effects in chromatin organization, providing molecular clues to understand the pathogenic effect of mutations affecting the loop-extrusion machinery.

CTCF is a multi-functional zinc-finger protein shown to be involved in an important number of cellular process such as transcriptional control, DNA splicing, recombination or replication [42, 43]. CTCF is essential for mammalian development, as homozygous knockout mice embryos decay at the blastocyst stage upon exhaustion of the maternal pool of the protein [44]. The observed lethality, in agreement with the multi-functional nature of this factor [43], has historically limited the study of the function of CTCF in organizing chromatin. However, as exposed previously, the utilization of protein degradation systems provided a novel approach to overcome this important issue [34, 35]. The application of these methods manifested that, upon CTCF loss, TADs and CTCF-mediated loops disappear in a dosage-dependent manner, but segregation into A/B compartments is largely preserved. These results not only demonstrate the critical function of CTCF in 3D genomic organization, but also the independent action of transcription-based compartmentalization and loop-extrusion forces.

Developmental defects can also be observed in patients with CTCF haploinsufficiency, which display phenotypes such as intellectual disability, microcephaly or growth retardation. Noteworthy, gene misregulation occurs at gene promoters that are predominantly engaged with active enhancers through CTCF-mediated interactions, suggesting that the observed effects might be explained by an involvement of CTCF in organizing genomes spatially [45]. Furthermore, mutations on CTCF zinc-fingers residues, potentially affecting chromatin binding, have been frequently associated to distinct cancer types such as endometrium, breast, prostate and Wilms' tumor [46–48]. This striking association is also reflected on the higher incidence level of cancer-related somatic mutations occurring at CTCF/cohesin binding sites [49]. A lower expression of CTCF, as well as of cohesin, caused by unknown mutations, was also found in childhood acute lymphoblastic leukemia (ALL). As expected, the reduction on CTCF expression led to a general genome-wide dysregulation of gene expression mainly due to global decrease on TAD insulation, again providing mechanistic clues on how the loss of CTCF can induce disease [50].

The other main component of the loop-extrusion machinery, the cohesin complex, belongs to the group of the structural maintenance of chromosome (SMC) complexes. SMCs display a ring-shaped structure that encircles DNA and mediate a number of biological processes, such as sister-chromatid segregation, DNA repair or transcriptional control [51]. As occurs for CTCF, protein depletion of different subunits of the complex, like RAD21 or SMC3, or of the chromatin-loading factor NIPBL results in a global disappearance of TADs and CTCF-mediated loops [34, 37]. However, these effects are also accompanied by a reinforcement of A/B compartmentalization, suggesting that chromatin extrusion can still occur in the absence of CTCF, but not of cohesin. Interestingly, depletions on the chromatin releasing factor WAPL or PDS5, which increase the persistence

of the complex on chromatin, have a reduced effect on existing CTCF-mediated loops and TADs, although new chromatin loops are formed at larger distances [34, 38]. These results manifest the dynamic nature of loop extrusion that is highly dependent on the residence time of both CTCF and cohesin on chromatin, thus providing a framework to interpret the effects of pathogenic human mutations associated to the complex.

Such mutations are causative of a group of rare developmental and intellectual disabilities denominated cohesinopathies. The most prominent subgroup of cohesinopathies is the Cornelia de Lange syndrome (CdLS), associated to facial features, hirsutism, various ophthalmologic abnormalities, limb defects, gastroesophageal dysfunction, growth retardation and neurodevelopmental delay [52]. The etiology of up to 65% of the cases can be attributed to mutations on NIPBL, or on the two cohesin subunits SMC1A and SMC3. While the exact pathomechanism for the syndrome remains to be elucidated, it is plausible that an impairment of the complex structure might affect 3D chromatin organization, causing widespread gene dysregulation [53]. This precise effect was described on a genome-wide expression analysis in human lymphoblastoid cell line (LCL) derived from CdLS patients carrying diverse mutations on the complex subunits [54]. Overall, the severity of the disease increased with the degree of gene dysregulation, which occurred around cohesin-bound regions. Those effects were suggestive of a potential role of cohesin as an insulator, an activity that was specifically disrupted in CdLS probands.

In summary, the exposed evidence supports a model where CTCF or cohesin mutations cause pathogenic effects through a genome-wide disruption of TAD organization and insulation. These effects might compromise the integrity of gene regulatory networks and, more importantly, mediate the formation of aberrant interactions between previously segregated promoters and regulatory elements, causing gene misexpression (Figure 1B).

Disruption of nuclear compartmentalization

Interphase nuclei contain distinct specialized sites that are essential to fulfill the functional requirements of the cell and that, in most cases, interact directly with chromatin to influence its transcriptional status (Figure 1A). These interactions result in the clustering of genomic regions displaying similar epigenetic characteristics, to form distinct nuclear compartments [9]. One of the most well-studied compartments is the nuclear lamina (NL), localized under the inner membrane of the nucleus and composed of many proteins, including the type V intermediate filaments lamins A, B1, B2 and C. The main functions of the NL are to maintain the nuclear shape and structure, and also to mediate transcriptional regulation, heterochromatin organization and nucleopore positioning [55]. Early microscopy studies revealed a physical connection between the NL and certain chromatin regions located at the nuclear periphery, suggesting a potential structural role in nuclear organization [56]. This effect that was further denoted after the development of fluorescence-based approaches [57]. With the recent development of DamID methods, based on the utilization of a methyltransferase fused to a DNA-binding protein, a genome-wide characterization of the genomic sequences associated to the NL was finally possible [58]. This technology revealed the existence of LADs, large genomic regions in the range of 0.1–10 megabases (Mb) that are mainly associated to repressive chromatin marks and low gene expression. Interestingly, LAD and TAD boundaries overlap to a certain extent, suggesting an active involvement of LADs in spatial chromatin organization. Nevertheless, LADs display a more dynamic nature that TADs, either by switching from an epigenetic repressive to an active environment through NL detachment, or vice versa, thereby mediating specific programs of cell specification and differentiation [59]. A recent study also identified the aggregation of non-contiguous TADs into cliques, repressive chromatin complexes that relocate to the NL during differentiation, overlapping with LAD domains and correlated with gene downregulation [14]. Remarkably, the opposite effect is observed during reprogramming, with a reduction on the number of TAD cliques and associated with the opening of repressed chromatin, as cells transition to a pluripotent state. Overall, LADs form a repressive environment that favors the transcriptional silencing of the genes embedded within them [60].

Mutations on individual NL components lead to a broad range of pathological phenotypes, known as laminopathies, which are suggestive of a global dysregulation of gene expression. To date, over 400 different mutations have been linked to different types of laminopathies, the vast majority of them associated to lamin A/C, thus supporting a more critical role of lamin B proteins in development [61]. Lamina-associated mutations induce a wide range of disorders that include premature aging like the Hutchinson–Gilford progeria syndrome (HGPS), as well as myopathies such as the Emery–Dreifuss muscular dystrophy (EDMD) or neuropathies like the Charcot–Marie–Tooth disease. While the molecular mechanisms causing laminopathies are still the subject of intense research, there is compelling indirect evidence that delineates NL alterations as disruptive for 3D chromatin organization. For example, studies on cells lacking the expression of lamin A/C and lamin B receptor (LBR), like nocturnal animal rods with inverted nuclei [62], show a dramatic relocation of heterochromatin from the periphery to the nuclear center [63]. Interestingly, these cells preserve the capacity to form TADs and compartments, although with qualitative differences when compared to other cell types. Another study on mouse embryonic stem cells (mESC) lacking all lamins revealed that, while the overall TAD structure is largely preserved, inter-TAD interactions are dramatically altered with prominent A/B compartment switching that correlates with transcriptional changes [64].

Therefore, the described consequences of altered NL structure might provide important clues to understand the phenotypes arising in certain laminopathies. The most common form of HGPS, caused by a deletion of 50 amino acids in the lamin A protein (LADelta50), results in impaired tethering of heterochromatin to the NL [65], accompanied by a loss of the repressive mark H3K27me3 [66] and disruption of lamin-mediated interactions [67]. 3D models of laminopathies, combining Hi-C and lamin ChIP-Seq, were performed on HeLA cells overexpressing a mutant lamin A protein (R388P) that cause congenital muscle dystrophy and lipodystrophy, as well as on fibroblasts derived from patients affected by familial partial lipodystrophy of Dunnigan type (FPLD2; LMNA p.R482W) [68]. In both cases, the 3D models predict a large reorganization of chromatin, with major intranuclear repositioning of LADs. *In vitro* studies for EDMD carrying distinct mutations for lamin A (p.R453W and p.H222P) also revealed a disruption of lamin A-associated heterochromatinization. Interestingly, these changes were associated to impaired inhibition of the pluripotency factor SOX2, which overexpression was shown to impair myogenic differentiation.

Lamins are also present in another specialized location within the cell nucleus: the nucleolus. While this structure is where rRNA transcription occurs, its periphery is mostly associated to heterochromatinization and gene silencing. Chromatin regions located at the nucleolus periphery are known

as NADs [69, 70], large genomic loci that show strong correlation with B2 sub-compartments, a subtype of B compartments characterized by lack of H3K27me3 and enrichment in LADs [12]. Therefore, NADs represent a subset of LADs that can be relocated from the periphery to this nuclear structure [71]. Interestingly, mutations in *MECP2*, causing the neurodevelopmental disorder Rett syndrome, have been shown to affect heterochromatinization during neuronal maturation, accompanied by a dramatic alteration of nucleolar structures [72, 73]. Remarkably, MECP2 knockdown in mammalian cells compromises nuclear lamin expression, thus providing an interesting mechanistic link between this pathology and 3D nuclear compartmentalization [74].

In summary, the disruption of nuclear compartments, in contrast to CTCF or cohesin mutations, is expected to largely preserve TAD organization but to cause alterations at a compartmental level. In such scenario, altered gene expression patterns would result from the compartmental switching of genomic regions, with consequent activation/repression of genes (Figure 1C). Further investigations will help to clarify the potential involvement of other nuclear locations, which disruption might also induce disease through the described mechanism.

Pathogenic effects of locus-specific genomic misfolding

As previously described, the disruption of factors controlling 3D-chromatin organization can lead to complex syndromes that result from transcriptional alterations occurring in a genome-wide fashion. However, aberrant spatial folding can be also restricted to certain genomic locations, causing more localized phenotypes that are largely dependent on the regulatory configuration of the locus; i.e. identity of genes, specificity of regulatory elements and local spatial organization. In the following section, we describe distinct pathologies affecting 3D chromatin folding at selected loci, mediated by diverse genetic and epigenetic mechanisms, with a highlight on recent findings involving mobile elements.

Structural variation affecting 3D chromatin organization

The partition of vertebrate genomes into functional units, such as TADs, provides a regulatory scaffold that ensures appropriate cross-talk between tissue-specific enhancers and promoters, an aspect that is essential to induce precise spatial-temporal gene expression patterns (Figure 2A). The disruption of TADs, via structural variation, can alter this organization and generate novel functional interactions that result in gene misexpression and disease (Figure 2B) [3]. For example, deletions that affect boundary elements can lead to intermingling of regulatory information through TAD fusion. One prominent example involves the *EPHA4* locus, where large deletions of up to 1.5 Mb that include a CTCF-associated boundary can lead to brachydactyly [75, 76]. The observed phenotype is the result of a complete fusion between adjacent TADs (Figure 2B), thus connecting a cluster of limb enhancers, normally associated to *EPHA4*, with the *PAX3* gene, causing its aberrant expression during embryonic development. Noteworthy, deletions of similar size, but not affecting the boundary region, can prevent these ectopic interactions and the misexpression of *PAX3*, resulting in normal development. Similar perturbations in boundary elements, and consequent TAD disruption, have been also described as a prominent cause of cancer, like on T-cell acute-lymphoblastic leukemia (T-ALL) or medulloblastoma, causing the ectopic activation of proto-oncogenes and abnormal cell proliferation [77–79]. Although these studies highlight the importance of TAD boundaries in the correct segregation of regulatory landscapes, it is important to note that, in some cases, boundary regions might not be the only components required to maintain TAD integrity. This has been exemplified by serial deletions at the *Sox9* locus, where those involving just boundary elements did not cause a complete fusion of TADs, an effect that only took place when other internal CTCF sites were also deleted [80]. Furthermore, individual and double deletions of CTCF-binding sites at the *Shh* TAD can cause a local reorganization of chromatin topology, but domain separation remains largely preserved [81, 82]. Overall, this evidence highlights the capacity of boundary elements to effectively segregate the genome into distinct regulatory units, but also their cooperativity with the internal TAD structure.

The insulating capacity of boundary elements is also illustrated by duplications, which have the potential to form novel isolated structures denominated neo-TADs (Figure 2B). The pathogenic potential of neo-TADs has been also studied at the *SOX9* locus, where duplications involving a boundary element resulted in the formation of a novel TAD that associates *SOX9* enhancers with the *KCNJ2* gene, leading to its misexpression and the appearance of Cooks syndrome [83]. Mouse models revealed that the described neo-TADs create novel regulatory domains that are isolated from other genomic regions. However, smaller duplications that do not include *Kcnj2* can also create a neo-TAD, but without any noticeable changes in gene expression, as the duplicated *Sox9* enhancers cannot find any promoter within the new formed structure. Neo-TAD formation has been also described as a prominent mechanism involved in certain types of cancer, like medulloblastoma [79]. Overall, the pathogenicity of neo-TADs is largely determined by the regulatory elements and genes that are confined within the functional unit, and therefore suitable to interact.

Inversions that include boundary elements, as well as translocations, can lead to gene misexpression through TAD shuffling, by mixing regulatory information from functional units that are otherwise isolated from each other in a healthy condition (Figure 2B). This phenomenon was also exemplified at the *EPHA4* locus [75, 76], where inversions including its centromeric TAD boundary and the cluster of *EPHA4* limb enhancers associate the latest with the *WNT6* gene, causing the fusion of fingers (F-syndrome). Noteworthy, this cluster of limb enhancers is the same described for the *PAX3* brachydactyly-associated deletions, as well as for other variants at the same locus that result in polydactyly through *IHH* misexpression. This is a remarkable finding that highlights the modular nature of vertebrate genomes, with the capability to induce distinct phenotypes depending on the properties of the associated elements, in this case exclusively determined by the identity of the gene (*PAX3*, *WNT6* or *IHH*). Upon inversion, however, not all genes displaying increased interaction frequency with *EPHA4* limb enhancers showed a transcriptional response, an effect also described on a series of additional inversions at the same locus [84]. A recent study on *Drosophila* balancer chromosomes, characterized by their high number of structural rearrangements, revealed selective transcriptional dysregulation, with changes in a subset of the genes that displayed altered interaction profiles [85]. Overall, these findings point to the existence of additional mechanisms contributing to gene activation upon genomic rearrangements, such as enhancer-promoter compatibility [86].

Besides inversions, similar TAD shuffling effects have been also described for translocations. For example, translocations involving the *PITX1* gene can lead to its ectopic activation in developing forelimbs, causing the appearance of the Liebenberg

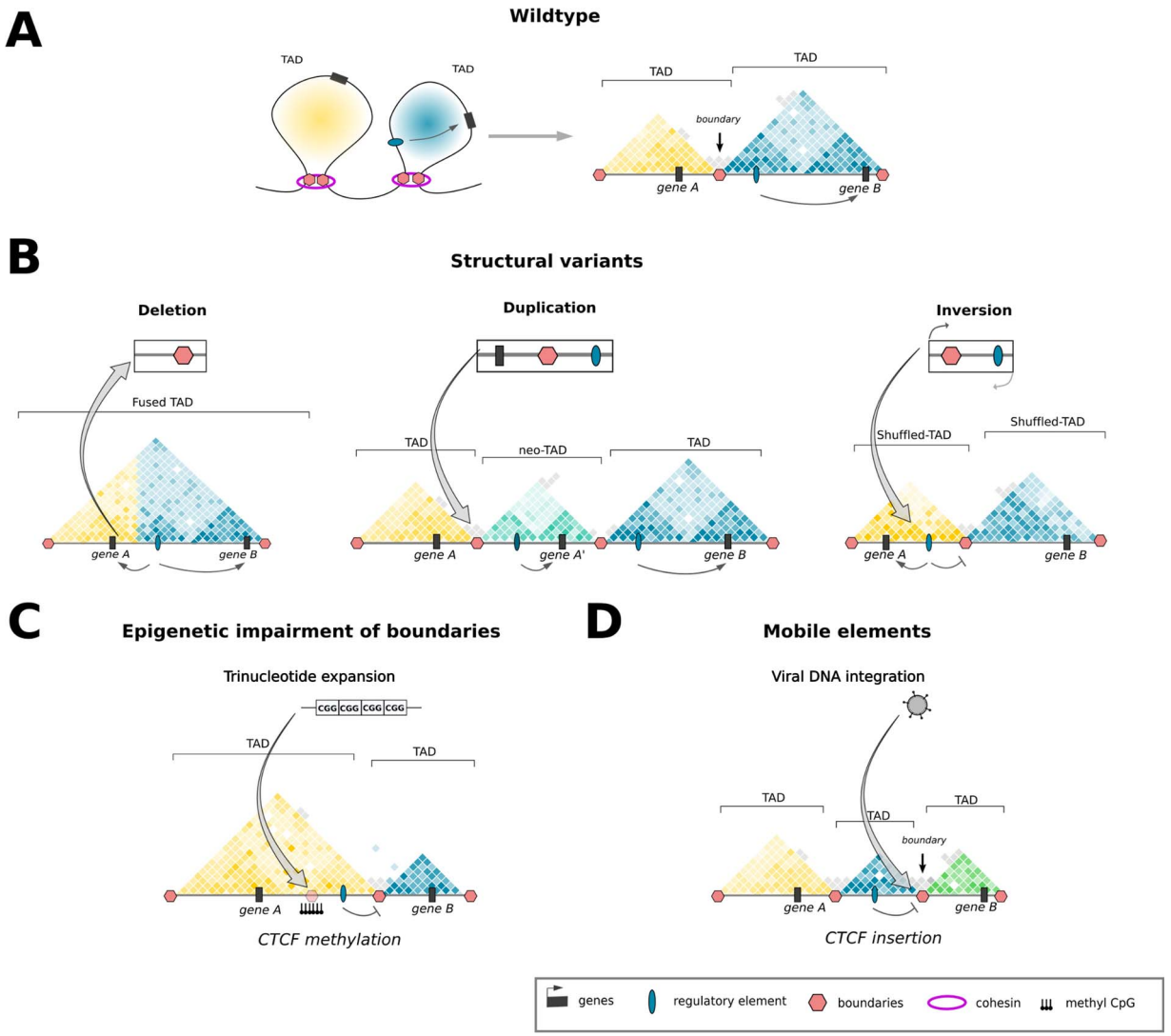

Figure 2. Effects of locus-specific 3D chromatin misfolding. In a wild-type situation, the correct segregation of TADs ensures proper gene activation. Structural variants can affect TAD organization and cause gene miexpression and disease. Left: deletions affecting TAD boundaries can lead to gene misexpression through TAD fusion. Center: duplications including TAD boundaries can form neo-TADs, which pathogenicity is determined by the regulatory elements and genes contained within. Right: inversions involving TAD boundaries can alter the relative position of regulatory elements causing phenotypes by either gene misexpression through novel enhancer-promoter associations (gene A), or by loss-of-function through the disconnection of enhancers from their cognate genes (gene B). TAD boundary function can be affected by epigenetic mechanisms. In the example, a trinucleotide expansion causes hypermethylation of a CTCF element and relocates the boundary to a more telomeric position, originating the disconnection of gene B from its endogenous enhancers and causing loss-of-function. *De novo* insertions of mobile elements can cause disease by altering the 3D chromatin landscape. In the example, a CTCF-associated retrovirus integrates and creates a new boundary element that partitions a TAD and disconnects gene B from its endogenous enhancers, thus causing loss-of-function.

syndrome that is characterized by an arm-to-leg homeotic transformation [87]. Nevertheless, the pathogenic effects of inversions or translocations cannot be only induced by ectopic gene expression through novel regulatory interactions, but also by the disconnection of enhancers from their cognate genes and consequent loss-of-function. One key study involves a large inversion of 89 Mb in patients affected by brachio-oculo-fascila syndrome. Using patient-specific human-induced pluripotent stem cells, the authors demonstrate that, although the genetic content of two distant TADs is intermingled, the phenotypes only arise from the disconnection of the *TFAP2A* gene from its endogenous enhancers [88]. This mechanism has been also reported in patients with balanced translocations that separate *MEF2C* gene, a known driver of 5q14.3 microdeletion syndrome, from its regulatory elements. The functional effects of these translocations were further validated in patient-derived LCLs, which displayed decreased *MEF2C* expression [89].

Finally, it should be noted that the high degree of TAD conservation between cell types facilitates the interpretation of the effects of structural variants even if chromatin interaction data sets are not available for the affected tissue [75]. Yet, such predictions can be challenging in genomic regions that do not display clear TAD structures or that display variable cell-to-cell organization. One example can be found at the *Pitx1* locus, where a tissue-specific loop connects the *Pitx1* gene with a pan-limb enhancer, controlling its exclusive expression in developing hindlimbs. However, large deletions can recapitulate this interaction loop also in forelimbs, causing ectopic enhancer-gene association and Liebenberg syndrome through *PITX1* misexpression [90].

Epigenetic impairment of boundary elements

Factors involved in establishing boundaries between TADs can be also highly susceptible to epigenetic changes that might affect their binding to chromatin and therefore impair chromatin organization. At certain loci, CTCF binding can be particularly sensitive to DNA methylation [91, 92], due to hypermethylation of CpG sites located within its DNA-binding core motif. This mechanism has been proven to be linked to human gliomas caused by mutations on the isocitrate dehydrogenase, originating global hypermethylation of CTCF-binding sites [93]. This effect resulted in a genome-wide decrease of CTCF binding associated with a loss of TAD insulation and ectopic gene activation. One of the observed conformational changes affected a TAD boundary causing a constitutive enhancer to ectopically activate the *platelet-derived growth factor receptor α* (*PDGFRA*) proto-oncogene and increased cell proliferation.

Associations between aberrant DNA-methylation levels and TAD disruption have not only been found in cancer, but also in inherited human disorders. A fascinating study uncovered a novel molecular mechanism behind the *FMR1* gene silencing in Fragile X syndrome (FXS), a disease characterized by an unstable expansion of repetitive DNA sequences termed short tandem repeats (STRs) [94]. The authors make the initial observation that many of these STRs, including those at the *FMR1* locus, frequently localize at TAD boundaries that are highly enriched on CpG islands. Surprisingly, they find that the boundary at the *FMR1* locus is disrupted in FXS, due to an impairment of CTCF occupancy caused by aberrant DNA methylation at the pathological CGG repeat expansion. The disruption of the boundary leads to *FMR1* silencing, likely due to a disconnection from its endogenous regulatory elements that become confined into a different TAD (Figure 2C). Although the exact order of events remains to be elucidated, this study suggest that the differential DNA methylation caused by the repeat expansion might affect the binding of CTCF and therefore rearrange the entire locus configuration, causing pathological gene misexpression.

Mobile elements with an insulating function

Mobile elements represent an important source of genomic variability, due to the innate capacity of these elements to mobilize between or within species. One important class of mobile elements are TEs, estimated to cover up to 44% of the human genome, although just only a small fraction (<0.05%) of these elements remain currently active [95]. Nevertheless, the rate of *de novo* germline transposition in humans for the most abundant TEs is approximately 1 in 21 births for Alu and 1 in 95 births for L1 elements [96–98], thereby representing a wide source of individual variability and associated with a significant number of diseases [99]. Importantly, TEs have been frequently linked to 3D genome organization, due to their tendency to frequently co-localize within the nuclear space [100], or their specific enrichment at TAD boundaries [16, 101]. In a landmark study, Schmidt et al. [102], demonstrated that an activation of retrotransposable elements caused an expansion of CTCF-binding events during the evolution of mammalian lineages. Importantly, many of those novel CTCF sites acted as chromatin insulators, measured by abrupt changes in chromatin states, and influenced 3D genome organization and transcription. A similar mechanism has been described for retroviral mobile elements such as the human T-lymphotropic virus type 1 (HTLV-1), which causes a chronic inflammatory process in the 10% of infected hosts [103]. In this case, the integration of HTLV-1 carried an ectopic CTCF-binding site with the potential of creating novel loops, remodeling local chromatin organization and inducing transcriptional deregulation at the insertion locus (Figure 2D) [104].

However, TEs can also induce conformational changes in a CTCF-independent manner. For example, mammalian-wide interspersed repeats, a class of TEs, were shown to display an insulator function on human T-cells in both *in vitro* and *in vivo* assays, presumably through the recruitment of transcriptional complexes [105]. Similar effects were described for SINE B2 elements at the murine growth hormone locus, which displayed an insulation function associated to Pol II and Pol III-driven transcripts [106]. Recently, two preprints also noted the potential of TEs to shape the 3D genome in early development and during differentiation. Kruse et al. [preprint: 107] studied genome architecture dynamics during the reprogramming of mESCs into 2-cell-like cells (2CLC), characterized by a marked increase on the expression of the murine endogenous retroviral element MERVL. The activation of this mobile element resulted in a *de novo* establishment of boundary regions in 2CLC, an effect also observed *in vivo*, at the early two-cell stage embryo. In another study, the primate-specific endogenous retrotransposons HERV-H were found to be associated with the establishment of TAD boundaries in human embryonic stem cells [preprint: 108]. These boundaries were formed preferentially on HERV copies with high levels of expression, suggesting a transcription-dependent mechanism. Deletions of HERV-H elements resulted in the disappearance of TAD boundaries, while insertions at distinct loci demonstrated their potential to induce the formation of these structures. Overall, the capacity of mobile elements to insert at different genomic locations and to trigger insulation effects in a CTCF-dependent or independent fashion provides a compelling pathomechanism linked to aberrant 3D chromatin organization.

Outlook

The development of novel technologies to map chromatin interactions at high resolution has contributed significantly to our understanding of nuclear organization and the identification of novel disease mechanisms that are the direct consequence of abnormal 3D chromatin folding. The discovery of functional spatial units, such as TADs, provided a useful framework to interpret the pathogenic effects of human mutations and their underlying molecular causes [3, 109]. However, it is becoming increasingly evident that organization into TAD does not apply equally across the entire genome, especially in regions with complex structures or those with marked cell-to-cell variability. Therefore, the identification of the additional organizational rules operating at such regions represents one of the current challenges on the field. The development of single-cell technologies such as scHi-C [110] and genome architectural mapping [111], or the recent advances in super-resolution microscopy [112] are helping to fill this gap, allowing a detailed interrogation of the dynamic principles underlying chromatin interactions. A better elucidation of nuclear-organizational principles, as well as the generation of a wide collection of cell, tissue and temporal-specific interaction data sets, will be essential to better delineate the distinct pathomechanisms that can be the result of disrupted chromatin folding.

The direct application of this growing knowledge is expected to derive in better computation approaches to reconstruct and predict 3D genomic organization. Recent developments have successfully incorporated novel concepts, such as transcriptional-based compartmentalization and loop extrusion, to develop accurate models that recreate this complex biological process to a great extent [68, 113, 114]. Furthermore, these approaches have

demonstrated their intrinsic potential to predict the pathogenic effects of human mutations in chromatin [76].

Nevertheless, an important caveat on the molecular diagnosis of human disease relates to our limited capability to determine the exact breakpoints of structural variants, or the precise genomic location of *de novo* insertions of mobile elements. Such constrains derive from the limitations of short-read sequencing in resolving the repetitive nature of these genomic regions. However, the emergence of third-generation sequencing methods has alleviated this issue by enabling the mapping of low-complexity nucleotide sequences, thus resulting in an increased detection rate of these events. It is the combination of these novel mapping technologies with emerging computational approaches what will result in better predictive strategies to ultimately determine the pathological consequences of abnormal 3D chromatin folding.

Key Points

- Abnormal 3D chromatin organization can lead to disease by rewiring interactions between genes and regulatory elements.
- Pathogenic phenotypes are largely determined by the spatio-temporal activity of regulatory elements and the identity of the newly associated genes, as well as their compatibility.
- Mutations on factors organizing chromatin in space cause widespread effects by global gene misregulation.
- The disruption of 3D genomic architecture at specific loci leads to phenotypes with a more tissue-specific nature.

Acknowledgments

We apologize to colleagues whose work we were unable to discuss or failed to cite owing to space constraints. We thank Francisca Martinez Real and members of the Lupiáñez laboratory for fruitful discussions. We also thank the two reviewers for their valuable comments and suggestions.

Conflicts of interest

The authors declare no conflict of interest.

Funding

C.A. and D.G.L. are supported by a grant from the Deutsche Forschungsgemeinschaft (GA2495/1-1). Work in the D.G.L.'s laboratory is funded by the Deutsche Forschungsgemeinschaft and the Helmholtz Association.

References

1. Dekker J, Mirny L. The 3D genome as moderator of chromosomal communication. *Cell* 2016;**164**:1110–21.
2. Long HK, Prescott SL, Wysocka J. Ever-changing landscapes: transcriptional enhancers in development and evolution. *Cell* 2016;**167**:1170–87.
3. Spielmann M, Lupiáñez DG, Mundlos S. Structural variation in the 3D genome. *Nat Rev Genet* 2018;**19**:453–67.
4. Lupiáñez DG, Spielmann M, Mundlos S. Breaking TADs: how alterations of chromatin domains result in disease. *Trends Genet* 2016;**32**:225–37.
5. Bickmore WA. The spatial organization of the human genome. *Annu Rev Genomics Hum Genet* 2013;**14**:67–84.
6. Dekker J, Rippe K, Dekker M, *et al*. Capturing chromosome conformation. *Science* 2002;**295**:1306–11.
7. Cremer T, Cremer M. Chromosome territories. *Cold Spring Harb Perspect Biol* 2010;**2**(3).
8. Branco MR, Pombo A. Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. *PLoS Biol* 2006;**4**(5).
9. Lieberman-Aiden E, van Berkum NL, Williams L, *et al*. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 2009;**326**:289–93.
10. Bonev B, Mendelson Cohen N, Szabo Q, *et al*. Multiscale 3D genome rewiring during mouse neural development. *Cell* 2017;**171**:557–72.
11. Dixon JR, Jung I, Selvaraj S, *et al*. Chromatin architecture reorganization during stem cell differentiation. *Nature* 2015;**518**:331–6.
12. Rao SSP, Huntley MH, Durand NC, *et al*. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 2014;**159**:1665–80.
13. Fraser J, Ferrai C, Chiariello AM, *et al*. Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation. *Mol Syst Biol* 2015;**11**:852.
14. Paulsen J, Ali TML, Nekrasov M, *et al*. Long-range interactions between topologically associating domains shape the four-dimensional genome during differentiation. *Nat Genet* 2019;**51**:835.
15. Stadhouders R, Vidal E, Serra F, *et al*. Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. *Nat Genet* 2018;**50**:238–49.
16. Dixon JR, Selvaraj S, Yue F, *et al*. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 2012;**485**:376–80.
17. Nora EP, Lajoie BR, Schulz EG, *et al*. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 2012;**485**:381–5.
18. Symmons O, Uslu VV, Tsujimura T, *et al*. Functional and topological characteristics of mammalian regulatory domains. *Genome Res* 2014;**24**:390–400.
19. Phillips-Cremins JE, Sauria MEG, Sanyal A, *et al*. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* 2013;**153**:1281–95.
20. Dowen JM, Fan ZP, Hnisz D, *et al*. Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* 2014;**159**:374–87.
21. Tang Z, Luo OJ, Li X, *et al*. CTCF-mediated human 3D genome architecture reveals chromatin topology for transcription. *Cell* 2015;**163**:1611–27.
22. Ji X, Dadon DB, Powell BE, *et al*. 3D chromosome regulatory landscape of human pluripotent cells. *Cell Stem Cell* 2016;**18**:262–75.
23. Rowley MJ, Nichols MH, Lyu X, *et al*. Evolutionarily conserved principles predict 3D chromatin organization. *Mol Cell* 2017;**67**:837–52.
24. Quinodoz SA, Ollikainen N, Tabak B, *et al*. Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. *Cell* 2018;**174**:744–57.

25. Hug CB, Grimaldi AG, Kruse K, *et al*. Chromatin architecture emerges during zygotic genome activation independent of transcription. *Cell* 2017;**169**:216–28.
26. Li L, Lyu X, Hou C, *et al*. Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing. *Mol Cell* 2015;**58**:216–31.
27. Battulin N, Fishman VS, Mazur AM, *et al*. Comparison of the three-dimensional organization of sperm and fibroblast genomes using the Hi-C approach. *Genome Biol* 2015;**16**:77.
28. Jung YH, Sauria MEG, Lyu X, *et al*. Chromatin states in mouse sperm correlate with embryonic and adult regulatory landscapes. *Cell Rep* 2017;**18**:1366–82.
29. Du Z, Zheng H, Huang B, *et al*. Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature* 2017;**547**:232–5.
30. Sanborn AL, Rao SSP, Huang S-C, *et al*. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc Natl Acad Sci U S A* 2015;**112**:E6456–65.
31. Fudenberg G, Imakaev M, Lu C, *et al*. Formation of chromosomal domains by loop extrusion. *Cell Rep* 2016;**15**:2038–49.
32. Vietri Rudan M, Barrington C, Henderson S, *et al*. Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell Rep* 2015;**10**: 1297–309.
33. Gómez-Marín C, Tena JJ, Acemel RD, *et al*. Evolutionary comparison reveals that diverging CTCF sites are signatures of ancestral topological associating domains borders. *Proc Natl Acad Sci U S A* 2015;**112**:7542–7.
34. Wutz G, Várnai C, Nagasaka K, *et al*. Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL and PDS5 proteins. *EMBO J* 2017;**36**:3573–99.
35. Nora EP, Goloborodko A, Valton A-L, *et al*. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* 2017;**169**:930–44.
36. Rao SSP, Huang S-C, Glenn St Hilaire B, *et al*. Cohesin loss eliminates all loop domains. *Cell* 2017;**171**:305–20.
37. Schwarzer W, Abdennur N, Goloborodko A, *et al*. Two independent modes of chromatin organization revealed by cohesin removal. *Nature* 2017;**551**:51–6.
38. Haarhuis JHI, van der RH, Blomen VA, *et al*. The cohesin release factor WAPL restricts chromatin loop extension. *Cell* 2017;**169**:693–707.
39. Xie L, Dong P, Qi Y, *et al*. Super-resolution imaging reveals 3D structure and organizing mechanism of accessible chromatin. *bioRxiv* 2019; 678649.
40. Cuartero S, Weiss FD, Dharmalingam G, *et al*. Control of inducible gene expression links cohesin to hematopoietic progenitor self-renewal and differentiation. *Nat Immunol* 2018;**19**:932–41.
41. Kim S, Yu N-K, Shim K-W, *et al*. Remote memory and cortical synaptic plasticity require neuronal CCCTC-binding factor (CTCF). *J Neurosci* 2018;**38**:5042–52.
42. Ong C-T, Corces VGCTCF. An architectural protein bridging genome topology and function. *Nat Rev Genet* 2014;**15**: 234–46.
43. Phillips JE, Corces VGCTCF. Master weaver of the genome. *Cell* 2009;**137**:1194–211.
44. Moore JM, Rabaia NA, Smith LE, *et al*. Loss of maternal CTCF is associated with peri-implantation lethality of Ctcf null embryos. *PLoS One* 2012;**7**:e34915.
45. Gregor A, Oti M, Kouwenhoven EN, *et al*. De novo mutations in the genome organizer CTCF cause intellectual disability. *Am J Hum Genet* 2013;**93**:124–31.
46. Ciriello G, Miller ML, Aksoy BA, *et al*. Emerging landscape of oncogenic signatures across human cancers. *Nat Genet* 2013;**45**:1127–33.
47. Filippova GN, Qi C-F, Ulmer JE, *et al*. Tumor-associated zinc finger mutations in the CTCF transcription factor selectively alter tts DNA-binding specificity. *Cancer Res* 2002;**62**:48–52.
48. Kemp CJ, Moore JM, Moser R, *et al*. CTCF haploinsufficiency destabilizes DNA methylation and predisposes to cancer. *Cell Rep* 2014;**7**:1020–9.
49. Katainen R, Dave K, Pitkänen E, *et al*. CTCF/cohesin-binding sites are frequently mutated in cancer. *Nat Genet* 2015;**47**:818–21.
50. Yang M, Vesterlund M, Siavelis I, *et al*. Proteogenomics and Hi-C reveal transcriptional dysregulation in high hyperdiploid childhood acute lymphoblastic leukemia. *Nat Commun* 2019;**10**(1).
51. Uhlmann F. SMC complexes: from DNA to chromosomes. *Nat Rev Mol Cell Biol* 2016;**17**:399–412.
52. Liu J, Krantz ID. Cornelia de Lange syndrome, cohesin and beyond. *Clin Genet* 2009;**76**:303–14.
53. Yuan B, Pehlivan D, Karaca E, *et al*. Global transcriptional disturbances underlie Cornelia de Lange syndrome and related phenotypes. *J Clin Invest* 2015;**125**:636–51.
54. Liu J, Zhang Z, Bando M, *et al*. Transcriptional dysregulation in NIPBL and Cohesin mutant human cells. *PLoS Biol* 2009;**7**(5).
55. Akhtar A, Gasser SM. The nuclear envelope and transcriptional control. *Nat Rev Genet* 2007;**8**:507–17.
56. Fawcett DW. On the occurrence of a fibrous lamina on the inner aspect of the nuclear envelope in certain cells of vertebrates. *Am J Anat* 1966;**119**:129–45.
57. Lanctôt C, Cheutin T, Cremer M, *et al*. Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nat Rev Genet* 2007;**8**: 104–15.
58. Guelen L, Pagie L, Brasset E, *et al*. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* 2008;**453**:948–51.
59. Peric-Hupkes D, Meuleman W, Pagie L, *et al*. Molecular maps of the reorganization of genome–nuclear lamina interactions during differentiation. *Mol Cell* 2010;**38**: 603–13.
60. Leemans C, MCH Z, van der Brueckner L, *et al*. Promoter-intrinsic and local chromatin features determine gene repression in LADs. *Cell* 2019;**177**:852–64.
61. Kang S, Yoon M-H, Laminopathies PB-J. Mutations on single gene and various human genetic diseases. *BMB Rep* 2018;**51**:327–37.
62. Solovei I, Wang AS, Thanisch K, *et al*. LBR and lamin a/C sequentially tether peripheral heterochromatin and inversely regulate differentiation. *Cell* 2013;**152**:584–98.
63. Falk M, Feodorova Y, Naumova N, *et al*. Heterochromatin drives compartmentalization of inverted and conventional nuclei. *Nature* 2019;**570**:395–99.
64. Zheng X, Hu J, Yue S, *et al*. Lamins organize the global three-dimensional genome from the nuclear periphery. *Mol Cell* 2018;**71**:802–15.
65. Goldman RD, Shumaker DK, Erdos MR, *et al*. Accumulation of mutant lamin A causes progressive changes in nuclear

architecture in Hutchinson–Gilford progeria syndrome. *Proc Natl Acad Sci U S A* 2004;**101**:8963–8.
66. Shumaker DK, Dechat T, Kohlmaier A, *et al*. Mutant nuclear lamin A leads to progressive alterations of epigenetic control in premature aging. *Proc Natl Acad Sci U S A* 2006;**103**:8703–8.
67. McCord RP, Nazario-Toole A, Zhang H, *et al*. Correlated alterations in genome organization, histone methylation, and DNA-lamin A/C interactions in Hutchinson-Gilford progeria syndrome. *Genome Res* 2013;**23**:260–9.
68. Paulsen J, Sekelja M, Oldenburg AR, *et al*. Chrom3D: three-dimensional genome modeling from Hi-C and nuclear lamin-genome contacts. *Genome Biol* 2017;**18**:21.
69. Németh A, Conesa A, Santoyo-Lopez J, *et al*. Initial genomics of the human nucleolus. *PLoS Genet* 2010;**6**:e1000889.
70. van Koningsbruggen S, Gierliński M, Schofield P, *et al*. High-resolution whole-genome sequencing reveals that specific chromatin domains from most human chromosomes associate with nucleoli. *Mol Biol Cell* 2010;**21**:3735–48.
71. Kind J, Pagie L, Ortabozkoyun H, *et al*. Single-cell dynamics of genome-nuclear lamina interactions. *Cell* 2013;**153**:178–92.
72. Agarwal N, Becker A, Jost KL, *et al*. MeCP2 Rett mutations affect large scale chromatin organization. *Hum Mol Genet* 2011;**20**:4187–95.
73. Singleton MK, Gonzales ML, Leung KN, *et al*. MeCP2 is required for global heterochromatic and nucleolar changes during activity-dependent neuronal maturation. *Neurobiol Dis* 2011;**43**:190–200.
74. Babbio F, Castiglioni I, Cassina C, *et al*. Knock-down of methyl CpG-binding protein 2 (MeCP2) causes alterations in cell proliferation and nuclear lamins expression in mammalian cells. *BMC Cell Biol* 2012;**13**:19.
75. Lupiáñez DG, Kraft K, Heinrich V, *et al*. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* 2015;**161**:1012–25.
76. Bianco S, Lupiáñez DG, Chiariello AM, *et al*. Polymer physics predicts the effects of structural variants on chromatin architecture. *Nat Genet* 2018.
77. Hnisz D, Weintraub AS, Day DS, *et al*. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* 2016;**351**:1454–8.
78. Northcott PA, Lee C, Zichner T, *et al*. Enhancer hijacking activates GFI1 family oncogenes in medulloblastoma. *Nature* 2014;**511**:428–34.
79. Weischenfeldt J, Dubash T, Drainas AP, *et al*. Pan-cancer analysis of somatic copy-number alterations implicates IRS4 and IGF2 in enhancer hijacking. *Nat Genet* 2017;**49**:65–74.
80. Despang A, Schöpflin R, Franke M, *et al*. Functional dissection of the Sox9–Kcnj2 locus identifies nonessential and instructive roles of TAD architecture. *Nat Genet* 2019;**51**:1263–71.
81. Williamson I, Kane L, Devenney PS, *et al*. Developmentally regulated Shh expression is robust to TAD perturbations. *Development* 2019;**146**:19.
82. Paliou C, Guckelberger P, Schöpflin R, *et al*. Preformed chromatin topology assists transcriptional robustness of Shh during limb development. *Proc Natl Acad Sci U S A* 2019;**116**:12390–9.
83. Franke M, Ibrahim DM, Andrey G, *et al*. Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature* 2016;**538**:265–9.
84. Kraft K, Magg A, Heinrich V, *et al*. Serial genomic inversions induce tissue-specific architectural stripes, gene misexpression and congenital malformations. *Nat Cell Biol* 2019;**21**:305.
85. Ghavi-Helm Y, Klein FA, Pakozdi T, *et al*. Enhancer loops appear stable during development and are associated with paused polymerase. *Nature* 2014;**512**:96–100.
86. Zabidi MA, Arnold CD, Schernhuber K, *et al*. Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation. *Nature* 2015;**518**:556–9.
87. Spielmann M, Brancati F, Krawitz PM, *et al*. Homeotic arm-to-leg transformation associated with genomic rearrangements at the PITX1 locus. *Am J Hum Genet* 2012;**91**:629–35.
88. Laugsch M, Bartusel M, Rehimi R, *et al*. Modeling the pathological long-range regulatory effects of human structural variation with patient-specific hiPSCs. *Cell Stem Cell* 2019;**24**:736–52.
89. Redin C, Brand H, Collins RL, *et al*. The genomic landscape of balanced cytogenetic abnormalities associated with human congenital anomalies. *Nat Genet* 2017;**49**:36–45.
90. Kragesteen BK, Spielmann M, Paliou C, *et al*. Dynamic 3D chromatin architecture contributes to enhancer specificity and limb morphogenesis. *Nat Genet* 2018;**50**:1463–73.
91. Bell AC, Felsenfeld G. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* 2000;**405**:482–5.
92. Hark AT, Schoenherr CJ, Katz DJ, *et al*. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 2000;**405**:486–9.
93. Flavahan WA, Drier Y, Liau BB, *et al*. Insulator dysfunction and oncogene activation in *IDH* mutant gliomas. *Nature* 2016;**529**:110–4.
94. Sun JH, Zhou L, Emerson DJ, *et al*. Disease-associated short tandem repeats co-localize with chromatin domain boundaries. *Cell* 2018;**175**:224–38.
95. Mills RE, Bennett EA, Iskow RC, *et al*. Which transposable elements are active in the human genome? *Trends Genet* 2007;**23**:183–91.
96. Bourque G, Burns KH, Gehring M, *et al*. Ten things you should know about transposable elements. *Genome Biol* 2018;**19**:199.
97. Xing J, Zhang Y, Han K, *et al*. Mobile elements create structural variation: analysis of a complete human genome. *Genome Res* 2009;**19**:1516–26.
98. Ewing AD, Kazazian HH. High-throughput sequencing reveals extensive variation in human-specific L1 content in individual human genomes. *Genome Res* 2010;**20**:1262–70.
99. Hancks DC, Kazazian HH. Roles for retrotransposon insertions in human disease. *Mob DNA* 2016;**7**:9.
100. Cournac A, Koszul R, Mozziconacci J. The 3D folding of metazoan genomes correlates with the association of similar repetitive elements. *Nucleic Acids Res* 2016;**44**:245–55.
101. Harmston N, Ing-Simmons E, Tan G, *et al*. Topologically associating domains are ancient features that coincide with metazoan clusters of extreme noncoding conservation. *Nat Commun* 2017;**8**:441.
102. Schmidt D, Schwalie PC, Wilson MD, *et al*. Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. *Cell* 2012;**148**:335–48.

103. Satou Y, Miyazato P, Ishihara K, *et al*. The retrovirus HTLV-1 inserts an ectopic CTCF-binding site into the human genome. *Proc Natl Acad Sci U S A* 2016;**113**:3054–9.
104. Melamed A, Yaguchi H, Miura M, *et al*. The human leukemia virus HTLV-1 alters the structure and transcription of host chromatin in cis. *Elife* 2018;**7**:e36245.
105. Wang J, Vicente-García C, Seruggia D, *et al*. MIR retrotransposon sequences provide insulators to the human genome. *Proc Natl Acad Sci U S A* 2015;**112**:E4428–37.
106. Lunyak VV, Prefontaine GG, Núñez E, *et al*. Developmentally regulated activation of a SINE B2 repeat as a domain boundary in organogenesis. *Science* 2007;**317**:248–51.
107. Kruse K, Díaz N, Enriquez-Gasca R, *et al*. Transposable elements drive reorganisation of 3D chromatin during early embryogenesis. *bioRxiv* 2019; 523712.
108. Zhang Y, Li T, Preissl S, *et al*. Transcriptionally active HERV-H retrotransposons demarcate topologically associating domains in human pluripotent stem cells. *Nat Genet* 2019;**51**:1380– 88.
109. Dixon JR, Xu J, Dileep V, *et al*. Integrative detection and analysis of structural variation in cancer genomes. *Nat Genet* 2018;**50**:1388–98.
110. Nagano T, Lubling Y, Stevens TJ, *et al*. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* 2013;**502**:59–64.
111. Beagrie RA, Scialdone A, Schueler M, *et al*. Complex multi-enhancer contacts captured by genome architecture mapping. *Nature* 2017;**543**:519–24.
112. Bintu B, Mateo LJ, Su J-H, *et al*. Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science* 2018;**362**(6413).
113. Nuebler J, Fudenberg G, Imakaev M, *et al*. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc Natl Acad Sci U S A* 2018;**115**: E6697–706.
114. Buckle A, Brackley CA, Boyle S, *et al*. Polymer simulations of heteromorphic chromatin predict the 3D folding of complex genomic loci. *Mol Cell* 2018;**72**:786–97.