

TADs and Their Borders: Free Movement or Building a Wall?

Li-Hsin Chang, Sourav Ghosh and Daan Noordermeer

Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, University Paris-sud, University Paris-Saclay, Gif-sur-Yvette, France

Correspondence to Daan Noordermeer: daan.noordermeer@i2bc.paris-saclay.fr

<https://doi.org/10.1016/j.jmb.2019.11.025>

Edited by Romain Koszul

Abstract

The tridimensional (3D) organization of mammalian genomes combines structures from different length scales. Within this organization, Topologically Associating Domains (TADs) are visible in Hi-C heat maps at the sub-megabase scale. The integrity of TADs is important for correct gene expression, but in a context-dependent and variable manner. The correct structure and function of TADs require the binding of the CTCF protein at both borders, which appears to block an active and dynamic mechanism of “Cohesin-mediated loop extrusion.” As a result, mammalian TADs appear as so-called “loop domains” in Hi-C data, which are the focus of this review. Here, we present a reanalysis of TADs from three “golden-standard” mammalian Hi-C data sets. Despite the prominent presence of TADs in Hi-C heat maps from all studies, we find consistently that regions within these domains are only moderately insulated from their surroundings. Moreover, single-cell Hi-C and superresolution microscopy have revealed that the structure of TADs and the position of their borders can vary from cell to cell. The function of TADs as units of gene regulation may thus require additional aspects, potentially incorporating the mechanism of loop extrusion as well. Recent developments in single-cell and multi-contact genomics and superresolution microscopy assays will be instrumental to link TAD formation and structure to their function in transcriptional regulation.

© 2019 Elsevier Ltd. All rights reserved.

TADs as Functional Units of Gene Regulation

A major biophysical challenge for cells is to compact their large genome—up to 2 m for a mammalian cell—into a nucleus that is only a few microns in diameter, all the while ensuring both proper genome function and the programmable flexibility required for cellular differentiation [1]. The first order of DNA compaction, the wrapping of 147 bp of DNA around a histone octamer core to form the nucleosome (“10 nm fiber”), is well agreed upon [2–4]. Subsequent higher-order chromatin organization is formed by more diverse combinations of DNA structures, ranging from chromatin loops and contact domains at the smallest scales to chromosome territories that occupy large volumes within interphase nuclei [5,6].

The development of Chromosome Conformation Capture (3C) and genomics-based derivatives has greatly advanced our capacity to explore structural

aspects of 3D genome architecture [7,8]. Particularly Hi-C has proven to be instrumental in our understanding of 3D genome organization, by providing information on all possible genomewide interactions within a single experiment [9]. When analyzed at sufficient resolution (typically below 100 kb), Hi-C experiments reveal the existence of distinct structural domains [10]. These domains, referred to as Topologically Associating Domains (TADs), are characterized by increased intradomain interactions and insulation from their surroundings and thus appear as “pyramids” in Hi-C heat maps [10,11] (Fig. 1A). Whereas TADs in mammalian cells are generally identified with median sizes just below one megabase, in other organisms they may manifest themselves at considerably different sizes [10–16]. Moreover, studies with increasingly higher resolution have suggested the presence of additional finer-scale levels of subdomains (“sub-TADs”), nested within mammalian TADs [17–20].

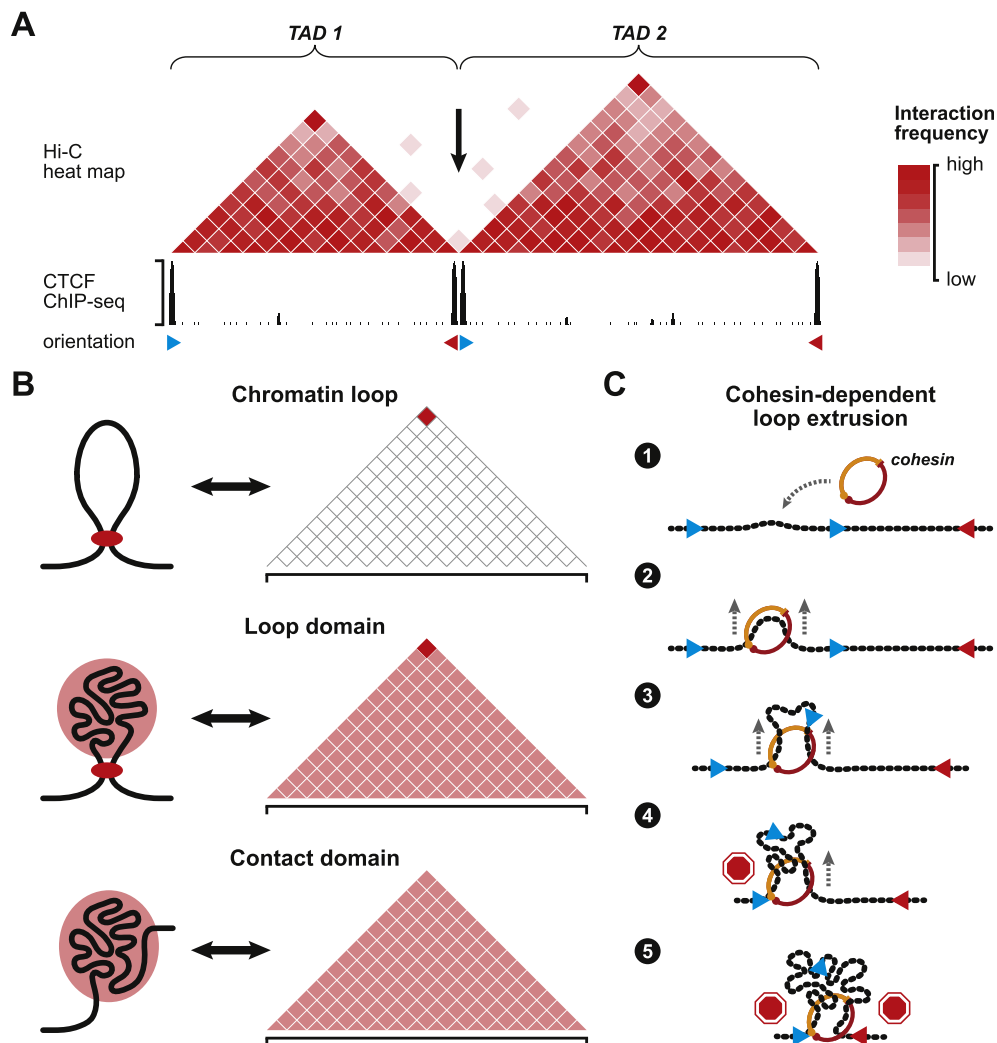


Fig. 1. TADs and loop extrusion. **A.** Schematic of two TADs in a Hi-C heat map, with the arrow indicating the border between the TADs. Simulated CTCF binding is depicted below, with convergently oriented CTCF peaks at the borders of both TADs (orientation of CTCF sites shown as blue and red triangles). **B.** Appearance of different types of chromatin structures in a Hi-C heat map. Left: cartoon of 3D chromatin structure; right: corresponding simplified Hi-C heat maps. **C.** Model for Cohesin-dependent loop extrusion. After random association of the Cohesin complex with the chromosome (1, 2), a loop will start to bidirectionally extrude (2, 3). Extrusion on either side will continue until a correctly oriented CTCF-bound site is encountered (left blue triangle; 4). A complete block is only achieved when both sides have encountered convergently orientated CTCFs (blue and red triangles; 5).

Initial correlative studies revealed enrichment within TADs of coregulated gene clusters, enhancer–promoter pairs, and replication domains [10,11,21,22]. Moreover, a large fraction of the borders that surround mammalian TADs bind the CTCF protein (CCCTC-binding factor), the only somatic insulator protein in somatic cells [10,18] (Fig. 1A). Until now, CTCF binding at large subsets of TAD borders has only been reported in vertebrates (e.g., Refs. [10,15]). Other TAD borders can be linked to (at least) two other biological phenomena. First, the presence of highly expressed genes can introduce a strong insulation between surround-

ing regions in the Hi-C heat map of a diverse range of organisms [10,23–25]. Second, a sharp separation may be observed between regions bound by histone modifications associated with opposing transcriptional and regulatory output [12,18,26]. This latter separation, in mammals usually at the smaller subdomain levels, appears not to directly involve the borders though, but rather to be caused by the self-aggregation of regions that carry the same epigenetic marks due to phase separation, thereby forming so-called “contact domains” or “compartmental domains” ([5,18,27–29], see also [30] in this issue of *Journal of Molecular Biology* for further

discussion). Unless mentioned otherwise, in the remainder of this review we focus on TADs where CTCF is bound at the borders.

Why CTCF is present at a large fraction of mammalian TAD borders, and if it is linked to transcriptional regulation, has been addressed in various pathologies and developmental contexts (see Table 1 for summary). The CTCF insulator protein was previously described as an “enhancer blocker” that precludes transcriptional activation when bound in between an enhancer–promoter pair [31]. Lupiáñez and colleagues showed that a group of congenital hand malformations could be traced to deletions, inversions, and duplications around the *EPHA4* gene that had in common that they created chromosomal configurations where CTCF-bound sites between the enhancers of the *EPHA4* gene and neighboring inactive genes were removed [32]. In turn, ectopic enhancer–promoter interactions could be formed that activated these neighboring genes, which in turn caused the various hand malformations [32]. Similarly, in Cooks syndrome, a duplication that spanned parts of two TADs

and CTCF-bound border in between resulted in the formation of a new TAD (“neo-TAD”) where the genes encountered a reorganized regulatory landscape that imposed aberrant activation [33]. In further analogy, mutations or epimutations of CTCF sites at TAD borders resulted in ectopic enhancer–promoter interactions and inappropriate gene activation in cancer cells as well [34,35]. Combined, these studies suggest that the enhancer blocking function of CTCF is achieved by positioning genes and nonrelevant enhancers in separate TADs.

More recent studies have provided nuance to the impact of perturbed CTCF binding at TAD borders. At the *Shh* locus, the limb-specific ZRS enhancer is located near the TAD border, where it is surrounded by multiple CTCF sites. Deletion of CTCF sites, either one or multiple, reduced *Shh* promoter–ZRS enhancer contacts but only had moderate influence on *Shh* gene activity. Removal of multiple CTCF sites increased interactions with neighboring sequences, which coincided with a mildly increased activity of a gene that is contained within this region [36,37]. Similarly, at the *Sox9* locus, a fusion with the

Table 1. Key studies showing the influence of TAD borders on gene expression

Key findings	Gene locus	Context	Ref.
Early and late-phase <i>HoxD</i> gene regulations in limb buds depend on enhancers located in two gene deserts located in TADs on either side of the <i>HoxD</i> cluster. The transition from the early to the late phase involves a functional and conformational switch, whereby a relocation of the TAD border within the <i>HoxD</i> clusters allows <i>Hoxd11</i> and <i>Hoxd9</i> genes to switch contacts from one TAD to the other.	<i>HoxD</i>	Limb development	[86]
Disruptions of a TAD border due to deletions, inversions, or duplications near the <i>EPHA4</i> gene result in the formation of aberrant interactions between the <i>EPHA4</i> enhancers and the neighboring <i>WNT6</i> , <i>IHH</i> , or <i>PAX3</i> genes. Their ectopic activation in turn results in hand malformations in human.	<i>WNT6-EPHA4</i>	Congenital hand malformations	[32]
A duplication of the <i>SOX9</i> and <i>KCNJ2</i> TADs and the border in between create a new chromatin domain (“neo-TAD”). In this structure, the <i>SOX9</i> enhancers ectopically activate the <i>KCNJ2</i> gene, resulting in limb malformations in human.	<i>SOX9-KCNJ2</i>	Limb malformations (Cooks syndrome)	[33]
Complete fusion of the <i>Sox9</i> and <i>Kcnj2</i> TADs requires the deletion of eight CTCF sites, located both at the border and inside the TADs. Upon fusion, <i>Sox9</i> expression is moderately reduced, whereas the neighboring <i>Kcnj2</i> gene is activated twofold.	<i>Sox9-Kcnj2</i>	Limb development	[38]
Deletion of all CTCF sites near the limb-specific ZRS enhancer at the <i>Shh</i> locus, which is located near a TAD border, results in the loss of the <i>Shh</i> –ZRS preformed interaction and a 50% decrease in <i>Shh</i> expression, but without a significant phenotype.	<i>Shh</i>	Limb development	[36]
Reorganization of <i>Shh</i> TAD structure and the deletion of one or several CTCF sites at the TAD border result in no detectable effect on <i>Shh</i> expression or phenotype.	<i>Shh</i>	Limb development	[37]
Expansion of the short tandem repeats around the <i>FMR1</i> gene in fragile X syndrome patients reduces CTCF binding at the nearby TAD border. As a result, the <i>FMR1</i> gene localizes to a different TAD, which is accompanied by the silencing of the gene.	<i>FMR1</i>	Fragile X syndrome	[87]
Human <i>IDH</i> mutant gliomas exhibit DNA hypermethylation. This hypermethylation reduces CTCF binding at the <i>PDGFRA</i> TAD border. This loss of insulation results in a constitutive enhancer to ectopically activate the <i>PDGFRA</i> oncogene.	<i>PDGFRA</i>	Glioma	[34]
Recurrent CTCF site microdeletions at TAD borders are common in T-cell acute lymphoblastic leukemia (T-ALL). Upon deletions, border insulation is eliminated, which can result in the activation of proto-oncogenes such as <i>TAL1</i> and <i>LMO2</i> .	<i>TAL1, LMO2</i>	T-cell acute lymphoblastic leukemia	[35]

neighboring TAD required the deletion of eight CTCF sites at and near the border. Yet, upon this fusion, the effect on gene activity was relatively mild, with *Sox9* activity moderately reduced and a twofold upregulation of the lowly expressed neighboring *Kcnj2* gene [38]. Combined, these studies show that TAD borders can be composed of more complex ensembles of CTCF-bound sites. Moreover, the effect of TAD-border perturbations on gene expression may be relatively moderate. Intriguingly, acute depletion of CTCF in embryonic stem cells resulted in similar moderate changes to transcriptional output. After 2 days of CTCF removal, insulation between TADs was strongly reduced, yet the transcriptional deregulation of only 900 genes could be assigned to TAD deregulation [39]. In a specific cell type, the contribution of individual TADs on gene expression may thus be relatively small. This effect may be related to the cell-type specificity of enhancers, which reduces the chance of inappropriate enhancer–promoter contacts within a given cell type [40,41].

In conclusion, genes and their associated enhancers are clustered in TADs, which act as functional but variable and context-dependent units of gene regulation. Demarcation and insulation of TADs are primarily dependent on the correct location of their borders, which are strongly enriched for the binding of the CTCF insulator protein (see also Table 1).

Cohesin-dependent Loop Extrusion and TAD Formation

TADs can be identified from a Hi-C heat map, where they appear as pyramid-shaped domains with increased intradomain interactions (Fig. 1A). Initial TAD-calling algorithms used a “Directionality Index” to identify bin-by-bin transitions in the asymmetry of contacts, thus indicating the switch from one insulated TAD to the next [10]. More recently, algorithms to determine “Insulation Scores” have been developed, which determine localized minima in interaction frequency by looking at a more regional scale, thus being more robust to individual outlying bins ([42,43]; see Ref. [30] for further considerations).

How punctuated CTCF binding at TAD borders can result in increased intradomain contacts remains not fully established. An important mechanistic suggestion comes from the Hi-C data itself, where many TADs with CTCF at their borders display strongly enriched signal on the top of the pyramid [18]. This punctuated signal indicates that the borders of these domains loop together, which is particularly frequently observed when CTCF is bound at both sides. These TADs, with both increased intradomain interactions and looping of the borders, are therefore also referred to as “loop

domains” [18,44] (Fig. 1B). This feature distinguishes these domains from conventional chromatin loops, which are often associated with enhancer–promoter contacts, and “contact/compartamental domains,” which are associated with the aggregation of domains carrying identical histone modifications [5,12,18,26,45] (Fig. 1B).

CTCF binds nonsymmetric DNA motifs in the genome, where it often colocalizes with the ring-shaped Cohesin complex [46–48]. Intriguingly, up to 90% of CTCF motifs that are at the base of TAD-spanning loops are in a convergent orientation [18,49,50]. The importance of the correct orientation of CTCF motifs has been experimentally confirmed by artificial inversions, which result in reduced TAD insulation [49,51].

How CTCF motif orientation guides TAD formation may be explained through the colocalization with the Cohesin complex. Based on this link, a “Cohesin-dependent loop extrusion” model has been developed, which reproduces Hi-C data *in silico* with high degrees of similarity to experimentally obtained data [52–55]. In this model, the Cohesin complex randomly associates with the chromatin fiber, where it starts the progressive formation of a DNA loop through bidirectional extrusion (Fig. 1C). Loop extrusion is blocked, on one side only, when a CTCF molecule is encountered that is bound in the correct orientation. When Cohesin has encountered CTCF on both borders of the TAD, the loop is stabilized, as reflected by a visible loop in the Hi-C map (Fig. 1C). The visible enrichment of intradomain interactions indicates that this process is simultaneously accompanied by more transient DNA contacts within the extruded loop, suggesting that it adopts a more crumpled configuration. Together, these two patterns of 3D interactions are reflected as a loop domain in the Hi-C heat map (Fig. 1B).

Importantly, to reconstruct TADs *in silico*, the loop extrusion in the model needs to be a dynamic process. This implies that the Cohesin complex does not only continuously associate with the DNA at random positions in the genome, but also dissociates from the DNA after a certain amount of time [54]. In practice, this dissociation will mostly occur after Cohesin has co-occupied a CTCF-bound site for a certain period of time. Various experimental studies have dissected how the combined action of the Cohesin complex and CTCF influences TAD structure, thereby confirming key aspects of the Cohesin-dependent loop extrusion model. Rapid elimination of the Cohesin component Rad21/Scc1, or of its loading factor Nipbl, results in elimination of all TADs that appear as loop domains. In contrast, contact domains appear mostly untouched or even increased in structure [44,56–58]. These studies therefore uncouple the process of loop domain formation from that of contact domains. Loss of TADs upon mutations in the ATPase domains of

Smc3, another Cohesin complex member, further confirms that loop extrusion is energy-consuming and thus an active mechanism [59]. In the same study, a subset of TADs was identified where one or both borders showed contiguous enriched contacts within the entire domain, visible as “architectural stripes.” These stripes are mostly found at TAD borders that combine strong insulation and multiple instances of CTCF binding with a nearby Cohesin loading area [59]. At such borders, the Cohesin complex will be blocked instantly on one side. According to loop extrusion model, such unidirectional loop formation is indeed predicted to manifest itself as horizontal and vertical stripes in the Hi-C matrix. Similar to Cohesin removal, the rapid removal of CTCF from TAD borders reduced the insulated nature of loop domains, whereas contact domains remained intact as well [39]. Interestingly though, the elimination of Cohesin and CTCF differentially affected genome compaction: whereas overall compaction was not affected by CTCF elimination, the loss of Cohesin association caused a global decompaction of the genome [39,57]. This supports the model that Cohesin can create loops in the absence of CTCF but that it needs to be blocked at specific sites in the genome to create noticeable loop domains.

An intriguing addition to the model of loop formation has been identified upon the deletion of Wapl, the protein that is required for the release of Cohesin from the DNA [58,60]. As expected from reinforced residence of the Cohesin complex at TAD borders, intra-TAD interactions become stronger. Unexpectedly though, TADs increased in size as well, which was visible as multiple TADs fusing into a single loop domain [60]. Prolonged blocking of Cohesin at TAD borders may therefore result in a “read through” of loop extrusion toward the next site occupied by CTCF in the correct orientation. This phenomenon may be due to a “traffic jam” effect, whereby multiple TAD borders cluster together [61]. In turn, this may be caused by the dynamics of CTCF binding to the DNA, whose dissociation is considerably more dynamic than that of the Cohesin complex (~1–2 min vs. ~20 min residence time) [62]. Cohesin-dependent loop extrusion may therefore not only depend on dynamic association and dissociation of the Cohesin complex but also involve a dynamic CTCF-binding component.

Combined, the results from *in silico* modeling and experimental studies provide accumulating evidence that TADs, as visible as loop domains in Hi-C heat maps, are shaped by a dynamic and active process that consists of a constant and energy-consuming bidirectional extrusion of DNA loops by the Cohesin complex. Within this process, CTCF functions to temporarily fix extruded chromatin loops at defined sites in the genome, which in turn creates specific insulated TADs.

TADs as Structural Units of Genome Organization: Consistency or Heterogeneity?

How TADs are structured and how they are involved in context-dependent transcriptional regulation is increasingly understood. In contrast, important questions remain on how stable TADs are over time, if they are consistent between different cells, and what is the nature of their internal structure. Most insights on TAD organization have been obtained from Hi-C studies on large cell populations, whose averaged interaction data preclude the interpretation of truly dynamic aspects.

Nonetheless, a more detailed appraisal of population-based Hi-C heat maps does provide valuable insights on how the chromatin is organized within and between TADs. Whereas pairing of CTCF-bound borders in a loop domain is represented by the strong signal at the top of the triangle in the Hi-C heat map, the internal organization of a TAD is described by the signal within the triangle itself. Moreover, the degree of insulation between neighboring TADs is measured by the depletion of signal between two TADs (Fig. 1A; absence of signal in the “inverted” pyramid between TADs). Similarly, the strength of borders between TADs influences insulation as well. Strong borders, visible as sharp valleys in the Hi-C heat map, are indicative of a separation between highly discrete TADs. In contrast, the presence of a more gradual valley between two TADs indicates a border that is more transitional and dynamic in nature (Fig. 2A). Inspection of high-resolution Hi-C data reveals that many TADs are separated by zones of transition that can span a considerable distance (Fig. 2B, see further for additional characterization). TAD-calling algorithms will call discrete borders within such zones of transition, yet the actual insulation at these zones is more gradual (Fig. 2B; dashed versus solid line in Hi-C heat map). In turn, according to the population-based Hi-C data, sequences near the borders do not strictly localize within one TAD (Fig. 2B; virtual 4C tracks). Although the presence of zones of transition may be due to artifacts in the Hi-C assay, similar interaction patterns are observed in the crosslink- and ligation-free DamC assay as well [63]. A more biological explanation for the appearance of these zones of transition may be that TAD borders can be dynamic, linked to both the discontinuous nature of CTCF binding and the observed clustering of CTCF binding sites at the majority of TAD borders [62,64,65]. In line with the model of Cohesin-dependent loop extrusion, the actual position of a TAD border may therefore not be fixed but vary within these defined zones of transition over time and between cells.

More recently developed single-cell Hi-C methods have started to shed light into the structural organization within and between TADs [56,66–68].

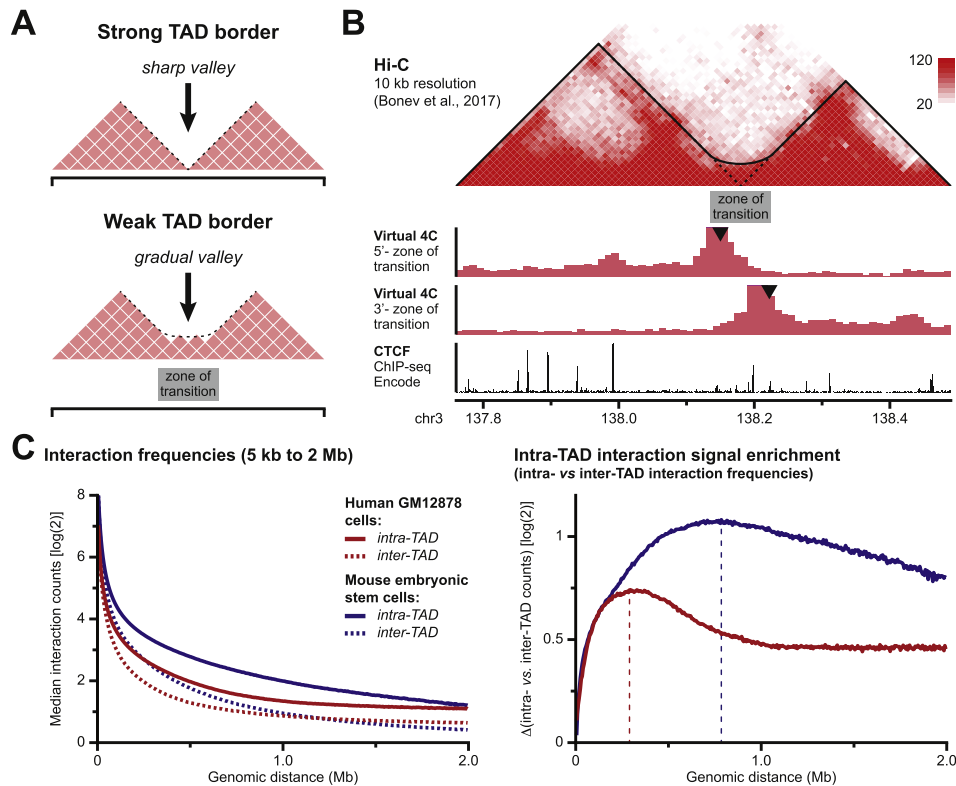


Fig. 2. The insulated nature of TADs and their borders. **A.** Schematic of strong and weak TAD borders. Weak TAD borders are separated by a “zone of transition” that appears as a gradual valley between two TADs. **B.** Reanalyzed Hi-C data showing a nearly 100 kb “zone of transition” between two TADs in mouse embryonic stem cells [13]. Top: Hi-C heat map with the algorithmically called TAD border indicated with dotted lines, and the more gradual transition between the TADs with a solid line. Middle: virtual 4C signal for two bins at either side of the zone of transition [12]. Bottom: published CTCF ChIP-seq signal in mouse embryonic stem cells [22]. **C.** Genomewide assessment of TAD insulation in human GM12878 B-lymphoblastoid cells [18] (red) and mouse embryonic stem cells [13] (blue) using reanalyzed Hi-C data at 5 kb resolution. Left: median genomewide interaction counts for intra-TAD (solid lines) and inter-TAD (dashed lines) read pairs at distances from 5 kb to 2 Mb. Right: enrichment of intra-TAD signal at different distance scales. Vertical dotted lines indicate the distance between read pairs with maximum signal enrichment (human B-lymphoblastoid cells: 290 kb – mouse embryonic stem cells: 785 kb).

Compared to population-based Hi-C, chromatin interactions within individual cells tend to be more stochastic and often span multiple TADs. Yet, combination of single-cell Hi-C data from multiple individual cells tends to faithfully reproduce TAD structure as observed in population-based Hi-C data [67,68]. TADs in population-based Hi-C data may thus represent the statistically most prominent average 3D configuration, which is built up from a variety of structures present within the cell population.

To get a more detailed and quantitative insight into the structural organization of TADs, we reanalyzed currently highest-resolution mouse and human Hi-C data sets, both of which are commonly used as reference contact maps [13,18]. Whereas TADs appear as (relatively) discrete pyramids in the Hi-C heat map of both data sets, a comparison of interaction frequencies between equally distant loci

that are located either within or between TADs reveals only a maximum twofold degree of insulation (Fig. 2C). A similar moderate difference was previously reported for lower-resolution mouse Hi-C data as well [10,69]. A comparison of TAD structure and borders between the two mouse embryonic stem cell data sets confirms their similarities (Fig. S1A-D). The calling of these structural aspects of TAD organization is therefore highly robust, despite considerable experimental and analytical differences. Interestingly, the maximum difference between intra- and inter-TAD interaction frequencies is only observed at a considerable distance from the borders between the TADs (Fig. 2C; dashed lines at 290 and 785 kb). This distance for maximum insulation exceeds the span of many TADs and appears not reduced at smaller TADs (Fig. 2C and Fig S1E). The zones of transition that separate TADs therefore encompass both the

borders and (at least partially) the domains themselves. In turn, these results suggest that the insulating capacity of TADs is relatively weak, yet this capacity scales with the size of the TAD.

Importantly though, a low level of insulation is not incompatible with the model of loop extrusion, as the extrusion process does not prevent the resulting loops in different domains from crossing borders and intermingling with each other. In contrast, it may suggest that the creation of regulatory units within TADs is not determined by the insulated nature of these domains alone. Rather, it may be that the ongoing loop extrusion that structures TADs is involved in guiding enhancer–promoter contacts as well. Indirect support for this hypothesis may be found at architectural stripes, where active (super-) enhancer elements that are in close proximity to stripe anchors have considerably elevated contacts with other regulatory elements [59]. Prevalent promoter–enhancer contacts between chromosomes in dipteran insects (including transvection) and an artificial mammalian system using an ectopic superenhancer and its endogenous target gene have ruled out though that the formation of regulatory contacts depends on loop extrusion alone [70,71].

Recent developments in superresolution microscopy, particularly the combination of STORM (STochastic Optical Reconstruction Microscopy) with Oligopaint labeling [72], constitute an alternative approach to characterize TAD structure and insulation in individual cells. Multiple studies in *Drosophila* cells were able to confirm the presence of TADs, albeit with a certain degree of heterogeneity and intermingling [73–75]. The small size of these domains and their overlap with specific histone modifications suggest that they are more associated with contact domains, rather than the process of Cohesin-dependent loop extrusion. In mammals, larger spatially segregated nanocompartments that correspond to TADs could also be visualized in individual cells [76]. Similar to Hi-C data, an enrichment of intra-TAD interactions could be detected, yet with relatively little specificity. Moreover, frequent contacts between TADs were detected as well, but with a reduced consistency between cells. Similar patterns of moderately enriched, two- to threefold, intra-TAD interactions were reported in a study that mapped large number of pairwise interactions at a genomewide scale [77].

Interestingly, the depletion of Cohesin in these single-cell microscopy studies did not result in a loss of TAD-like structures in individual cells, which contradicts published population-based Hi-C data [44,57,76]. In contrast, the preferential positioning of boundaries at CTCF-bound sites was lost, which is more in line with single-cell Hi-C studies [67,68]. The overall 3D organization of TADs may therefore be shaped by additional nested structures, including

smaller loop domains that are linked to epigenetic state.

Conclusions and Perspectives

In this review, we have discussed the current evidence for the existence of TADs, how they are structured, and how they are involved in transcriptional regulation. Whereas their presence can be unambiguously detected in population-based Hi-C heat maps, both quantitative interpretation of Hi-C data and the comparison of single-cell Hi-C and microscopy data suggest a certain degree of heterogeneity in TAD structure. As such, these data imply that TADs are dynamic structures, regarding both their internal structure and the position of their borders, with a relatively low degree of insulation from their surroundings. How this lack of structural consistency can be reconciled with their function as regulatory units in the genome remains to be determined. Here, we hypothesize that the setting of transcriptional programs may incorporate the process of TAD formation through loop extrusion as well.

To unambiguously link TAD formation and structure to transcriptional regulation, we will need to comprehensively address these aspects at highest precision in individual cells. Particularly instrumental will be the recent developments in superresolution microscopy, including the combined visualization of DNA conformation and RNA [78,79], and the development of genomics-based approaches that can identify and quantify ensembles of chromatin contacts. Recently, several technologies have emerged that allow either genomewide nontargeted characterization of chromatin hubs (C-walks [80], GAM (Genome Architecture Mapping) [81], SPRITE [82], and ChIA-Drop [83]) or the quantitative measurement of three-way interactions at defined sites (Multi-Contact 4C [61] and Tri-C [84]). The use of these technologies, and potential further developments, together with advanced modeling approaches and cellular models with modified aspects of TAD structure, will provide the tools to comprehensively unravel the dynamic nature of TADs.

Methods

Reanalysis of published Hi-C data

Unprocessed sequencing data from published Hi-C studies was obtained from the GEO data repository: human GM12878 B-lymphoblastoid cells: GSE63525 [18] and mouse embryonic stem cells: GSE96107 [13]. Data was processed at 5 kb resolution using the HiC-Pro tool, using standard settings and ICE-normalization of the

resulting matrices [85]. TADs and their borders were called using the TADtool from the 5 kb Hi-C matrices based on insulation score [43]. Lower-resolution Hi-C data [10], including normalized matrices and positions of TADs, were downloaded from <http://chromosome.sdsc.edu/mouse/hi-c/download.html>. For the comparison of TAD borders in mouse embryonic stem cells (Fig. S1A and C), data from Ref. [10] was converted to mm10 using the LiftOver tool (<http://genome.ucsc.edu>).

Autosomal read pairs spanning 5 kb–2 Mb were filtered for the presence within TADs, followed by sorting for being within the same TAD or in different TADs. Median interaction signal for all read pairs was calculated based on category and distance.

Acknowledgments

We thank the members of the Noordermeer lab for useful discussion. L.C. and S.G. are supported by funding from the Agence Nationale de Recherche to D.N. (grants ANR-14-ACHN-0009-01, ANR-16-TERC-0027-01 and ANR-17-CE12-0001-02).

Declarations of interest

None.

Author contributions

S.G. reanalyzed published Hi-C data. L.-H.C. and D.N. wrote the manuscript with input from S.G.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jmb.2019.11.025>.

Received 19 June 2019;

Received in revised form 1 November 2019;

Accepted 26 November 2019

Available online 27 December 2019

Keywords:

TADs;
topologically associating domains;
CTCF;
loop extrusion;
Hi-C

References

- [1] J. Mozziconacci, M. Merle, A. Lesne, The 3D Genome Shapes the Regulatory Code of Developmental Genes, *J. Mol. Biol.* 432 (2020) 712–723, <https://doi.org/10.1016/j.jmb.2019.10.017>.
- [2] S. Khorasanizadeh, The nucleosome: from genomic organization to genomic regulation, *Cell* 116 (2004) 259–272.
- [3] T. Schalch, S. Duda, D.F. Sargent, T.J. Richmond, X-ray structure of a tetranucleosome and its implications for the chromatin fibre, *Nature* 436 (2005) 138–141.
- [4] E. Fussner, M. Strauss, U. Djuric, R. Li, K. Ahmed, M. Hart, et al., Open and closed domains in the mouse genome are configured as 10-nm chromatin fibres, *EMBO Rep.* 13 (2012) 992–996.
- [5] M.J. Rowley, V.G. Corces, Organizational principles of 3D genome architecture, *Nat. Rev. Genet.* 19 (2018) 789–800.
- [6] R. Stadhouders, G.J. Filion, T. Graf, Transcription factors and 3D genome conformation in cell-fate decisions, *Nature* 569 (2019) 345–354.
- [7] J. Dekker, K. Rippe, M. Dekker, N. Kleckner, Capturing chromosome conformation, *Science* 295 (2002) 1306–1311.
- [8] J.O. Davies, A.M. Oudelaar, D.R. Higgs, J.R. Hughes, How best to identify chromosomal interactions: a comparison of approaches, *Nat. Methods* 14 (2017) 125–134.
- [9] E. Lieberman-Aiden, N.L. van Berkum, L. Williams, M. Imakaev, T. Ragozy, A. Telling, et al., Comprehensive mapping of long-range interactions reveals folding principles of the human genome, *Science* 326 (2009) 289–293.
- [10] J.R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, et al., Topological domains in mammalian genomes identified by analysis of chromatin interactions, *Nature* 485 (2012) 376–380.
- [11] E.P. Nora, B.R. Lajoie, E.G. Schulz, L. Giorgetti, I. Okamoto, N. Servant, et al., Spatial partitioning of the regulatory landscape of the X-inactivation centre, *Nature* 485 (2012) 381–385.
- [12] T. Sexton, E. Yaffe, E. Kenigsberg, F. Bantignies, B. Leblanc, M. Hoichman, et al., Three-dimensional folding and functional organization principles of the *Drosophila* genome, *Cell* 148 (2012) 458–472.
- [13] B. Bonev, N. Mendelson Cohen, Q. Szabo, L. Fritsch, G.L. Papadopoulos, Y. Lubling, et al., Multiscale 3D genome rewiring during mouse neural development, *Cell* 171 (2017) 557–572 e524.
- [14] C. Liu, Y.J. Cheng, J.W. Wang, D. Weigel, Prominent topologically associated domains differentiate global chromatin packing in rice from *Arabidopsis*, *Nat. Plants* 3 (2017) 742–748.
- [15] L.J.T. Kaaij, R.H. van der Weide, R.F. Ketting, E. de Wit, Systemic loss and gain of chromatin architecture throughout zebrafish development, *Cell Rep.* 24 (2018) 1–10 e14.
- [16] V. Fishman, N. Battulin, M. Nuriddinov, A. Maslova, A. Zlotina, A. Strunov, et al., 3D organization of chicken genome demonstrates evolutionary conservation of topologically associated domains and highlights unique architecture of erythrocytes' chromatin, *Nucleic Acids Res.* 47 (2019) 648–665.
- [17] J.E. Phillips-Cremins, M.E. Sauria, A. Sanyal, T.I. Gerasimova, B.R. Lajoie, J.S. Bell, et al., Architectural protein subclasses shape 3D organization of genomes during lineage commitment, *Cell* 153 (2013) 1281–1295.
- [18] S.S.P. Rao, M.H. Huntley, N.C. Durand, E.K. Stamenova, I.D. Bochkov, J.T. Robinson, et al., A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping, *Cell* 159 (2014) 1665–1680.
- [19] T.-H.S. Hsieh, E. Slobodyanyuk, A.S. Hansen, C. Cattoglio, O.J. Rando, R. Tjian, et al., Resolving the 3D landscape of transcription-linked mammalian chromatin folding, *bioRxiv* (2019) 638775.

- [20] N. Krietenstein, S. Abraham, S.V. Venev, N. Abdennur, J.H. Gibcus, T.-H.S. Hsieh, et al., Ultrastructural details of mammalian chromosome architecture, *bioRxiv* (2019) 639922.
- [21] A. Sanyal, B.R. Lajoie, G. Jain, J. Dekker, The long-range interaction landscape of gene promoters, *Nature* 489 (2012) 109–113.
- [22] Y. Shen, F. Yue, D.F. McCleary, Z. Ye, L. Edsall, S. Kuan, et al., A map of the cis-regulatory sequences in the mouse genome, *Nature* 488 (2012) 116–120.
- [23] C. Hou, L. Li, Z.S. Qin, V.G. Corces, Gene density, transcription, and insulators contribute to the partition of the *Drosophila* genome into physical domains, *Mol. Cell* 48 (2012) 471–484.
- [24] T.B. Le, M.V. Imakaev, L.A. Mirny, M.T. Laub, High-resolution mapping of the spatial organization of a bacterial chromosome, *Science* 342 (2013) 731–734.
- [25] U. Eser, D. Chandler-Brown, F. Ay, A.F. Straight, Z. Duan, W.S. Noble, et al., Form and function of topologically associating genomic domains in budding yeast, *Proc. Natl. Acad. Sci. U. S. A.* 114 (2017) E3061–E3070.
- [26] D. Noordermeer, M. Leleu, E. Splinter, J. Rougemont, W. De Laat, D. Duboule, The dynamic architecture of Hox gene clusters, *Science* 334 (2011) 222–225.
- [27] A.G. Larson, D. Elnatan, M.M. Keenen, M.J. Trnka, J.B. Johnston, A.L. Burlingame, et al., Liquid droplet formation by HP1 α suggests a role for phase separation in heterochromatin, *Nature* 547 (2017) 236–240.
- [28] A.R. Strom, A.V. Emelyanov, M. Mir, D.V. Fyodorov, X. Darzacq, G.H. Karpen, Phase separation drives heterochromatin domain formation, *Nature* 547 (2017) 241–245.
- [29] R. Tatavosian, S. Kent, K. Brown, T. Yao, H.N. Duc, T.N. Huynh, et al., Nuclear condensates of the Polycomb protein chromobox 2 (CBX2) assemble through phase separation, *J. Biol. Chem.* 294 (2019) 1451–1463.
- [30] E. de Wit, TADs as the Caller Calls Them, *J. Mol. Biol.* 432 (2020) 638–642, <https://doi.org/10.1016/j.jmb.2019.09.026>.
- [31] M. Merkenschlager, E.P. Nora, CTCF and cohesin in genome folding and transcriptional gene regulation, *Annu. Rev. Genom. Hum. Genet.* 17 (2016) 17–43.
- [32] D.G. Lupianez, K. Kraft, V. Heinrich, P. Krawitz, F. Brancati, E. Klopocki, et al., Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions, *Cell* 161 (2015) 1012–1025.
- [33] M. Franke, D.M. Ibrahim, G. Andrey, W. Schwarzer, V. Heinrich, R. Schopflin, et al., Formation of new chromatin domains determines pathogenicity of genomic duplications, *Nature* 538 (2016) 265–269.
- [34] W.A. Flavahan, Y. Drier, B.B. Liau, S.M. Gillespie, A.S. Venteicher, A.O. Stemmer-Rachamimov, et al., Insulator dysfunction and oncogene activation in IDH mutant gliomas, *Nature* 529 (2016) 110–114.
- [35] D. Hnisz, A.S. Weintraub, D.S. Day, A.L. Valton, R.O. Bak, C.H. Li, et al., Activation of proto-oncogenes by disruption of chromosome neighborhoods, *Science* 351 (2016) 1454–1458.
- [36] C. Paliou, P. Guckelberger, R. Schopflin, V. Heinrich, A. Esposito, A.M. Chiariello, et al., Preformed chromatin topology assists transcriptional robustness of Shh during limb development, *Proc. Natl. Acad. Sci. U. S. A.* 116 (2019) 12390–12399.
- [37] I. Williamson, L. Kane, P.S. Devenney, I.M. Flyamer, E. Anderson, F. Kilanowski, et al., Developmentally regulated Shh expression is robust to TAD perturbations, *Development* 146 (2019) dev179523.
- [38] A. Despang, R. Schopflin, M. Franke, S. Ali, I. Jerkovic, C. Paliou, et al., Functional dissection of the Sox9-Kcnj2 locus identifies nonessential and instructive roles of TAD architecture, *Nat. Genet.* 51 (2019) 1263–1271.
- [39] E.P. Nora, A. Goloborodko, A.L. Valton, J.H. Gibcus, A. Uebersohn, N. Abdennur, et al., Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization, *Cell* 169 (2017) 930–944 e922.
- [40] N.D. Heintzman, G.C. Hon, R.D. Hawkins, P. Kheradpour, A. Stark, L.F. Harp, et al., Histone modifications at human enhancers reflect global cell-type-specific gene expression, *Nature* 459 (2009) 108–112.
- [41] A. Visel, M.J. Blow, Z. Li, T. Zhang, J.A. Akiyama, A. Holt, et al., ChIP-seq accurately predicts tissue-specific activity of enhancers, *Nature* 457 (2009) 854–858.
- [42] E. Crane, Q. Bian, R.P. McCord, B.R. Lajoie, B.S. Wheeler, E.J. Ralston, et al., Condensin-driven remodelling of X chromosome topology during dosage compensation, *Nature* 523 (2015) 240–244.
- [43] K. Kruse, C.B. Hug, B. Hernandez-Rodriguez, J.M. Vaquerizas, TADtool: visual parameter identification for TAD-calling algorithms, *Bioinformatics* 32 (2016) 3190–3192.
- [44] S.S.P. Rao, S.C. Huang, B. Glenn St Hilaire, J.M. Engreitz, E.M. Perez, K.R. Kieffer-Kwon, et al., Cohesin loss eliminates all loop domains, *Cell* 171 (2017) 305–320 e324.
- [45] B. Tolhuis, R.J. Palstra, E. Splinter, F. Grosveld, W. de Laat, Looping and interaction between hypersensitive sites in the active beta-globin locus, *Mol. Cell* 10 (2002) 1453–1465.
- [46] T.H. Kim, Z.K. Abdullaev, A.D. Smith, K.A. Ching, D.I. Loukinov, R.D. Green, et al., Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome, *Cell* 128 (2007) 1231–1245.
- [47] V. Parelho, S. Hadjur, M. Spivakov, M. Leleu, S. Sauer, H.C. Gregson, et al., Cohesins functionally associate with CTCF on mammalian chromosome arms, *Cell* 132 (2008) 422–433.
- [48] K.S. Wendt, K. Yoshida, T. Itoh, M. Bando, B. Koch, E. Schirghuber, et al., Cohesin mediates transcriptional insulation by CCCTC-binding factor, *Nature* 451 (2008) 796–801.
- [49] E. de Wit, E.S. Vos, S.J. Holwerda, C. Valdes-Quezada, M.J. Verstegen, H. Teunissen, et al., CTCF binding polarity determines chromatin looping, *Mol. Cell* 60 (2015) 676–684.
- [50] Z. Tang, O.J. Luo, X. Li, M. Zheng, J.J. Zhu, P. Szalaj, et al., CTCF-mediated human 3D genome architecture reveals chromatin topology for transcription, *Cell* 163 (2015) 1611–1627.
- [51] Y. Guo, Q. Xu, D. Canzio, J. Shou, J. Li, D.U. Gorkin, et al., CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function, *Cell* 162 (2015) 900–910.
- [52] E. Alipour, J.F. Marko, Self-organization of domain structures by DNA-loop-extruding enzymes, *Nucleic Acids Res.* 40 (2012) 11202–11212.
- [53] A.L. Sanborn, S.S. Rao, S.C. Huang, N.C. Durand, M.H. Huntley, A.I. Jewett, 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.* 112 (2015) E6456–E6465.
- [54] G. Fudenberg, M. Imakaev, C. Lu, A. Goloborodko, N. Abdennur, L.A. Mirny, formation of chromosomal domains by loop extrusion, *Cell Rep.* 15 (2016) 2038–2049.

- [55] J. Nuebler, G. Fudenberg, M. Imakaev, N. Abdennur, L.A. Mirny, Chromatin organization by an interplay of loop extrusion and compartmental segregation, *Proc. Natl. Acad. Sci. U. S. A.* 115 (2018) E6697–E6706.
- [56] J. Gassler, H.B. Brandao, M. Imakaev, I.M. Flyamer, S. Ladstätter, W.A. Bickmore, et al., A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture, *EMBO J.* 36 (2017) 3600–3618.
- [57] W. Schwarzer, N. Abdennur, A. Goloborodko, A. Pekowska, G. Fudenberg, Y. Loe-Mie, et al., Two independent modes of chromatin organization revealed by cohesin removal, *Nature* 551 (2017) 51–56.
- [58] G. Wutz, G. Varnai, K. Nagasaka, D.A. Cisneros, R.R. Stocsits, W. Tang, et al., Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins, *EMBO J.* 36 (2017) 3573–3599.
- [59] L. Vian, A. Pekowska, S.S.P. Rao, K.R. Kieffer-Kwon, S. Jung, L. Baranello, et al., The energetics and physiological impact of cohesin extrusion, *Cell* 173 (2018) 1165–1178 e1120.
- [60] J.H.I. Haarhuis, R.H. van der Weide, V.A. Blomen, J.O. Yanez-Cuna, M. Amendola, M.S. van Ruiten, et al., The cohesin release factor WAPL restricts chromatin loop extension, *Cell* 169 (2017) 693–707 e614.
- [61] A. Allahyar, C. Vermeulen, B.A.M. Bouwman, P.H.L. Krijger, M. Versteegen, G. Geeven, et al., Enhancer hubs and loop collisions identified from single-allele topologies, *Nat. Genet.* 50 (2018) 1151–1160.
- [62] A.S. Hansen, I. Pustova, C. Cattoglio, R. Tjian, X. Darzacq, CTCF and cohesin regulate chromatin loop stability with distinct dynamics, *Elife* 6 (2017).
- [63] J. Redolfi, Y. Zhan, C. Valdes-Quezada, M. Kryzhanovska, I. Guerreiro, V. Iesmantavicius, et al., DamC reveals principles of chromatin folding in vivo without crosslinking and ligation, *Nat. Struct. Mol. Biol.* 26 (2019) 471–480.
- [64] S.A.M. Tonekaboni, P. Mazrooei, V. Kofia, B. Haibe-Kains, M. Lupien, Identifying clusters of cis-regulatory elements underpinning TAD structures and lineage-specific regulatory networks, *Genome Res.* 29 (2019) 1733–1743.
- [65] E. Kentepozidou, S.J. Aitken, C. Feig, K. Stefflova, X. Ibarra-Soria, D.T. Odom, et al., Clustered CTCF binding is an evolutionary mechanism to maintain topologically associating domains, *bioRxiv* (2019) 668855.
- [66] T. Nagano, Y. Lubling, T.J. Stevens, S. Schoenfelder, E. Yaffe, W. Dean, et al., Single-cell Hi-C reveals cell-to-cell variability in chromosome structure, *Nature* 502 (2013) 59–64.
- [67] I.M. Flyamer, J. Gassler, M. Imakaev, H.B. Brandao, S.V. Uljanov, N. Abdennur, et al., Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition, *Nature* 544 (2017) 110–114.
- [68] T.J. Stevens, D. Lando, S. Basu, L.P. Atkinson, Y. Cao, S.F. Lee, et al., 3D structures of individual mammalian genomes studied by single-cell Hi-C, *Nature* 544 (2017) 59–64.
- [69] J. Dekker, L. Mirny, The 3D genome as moderator of chromosomal communication, *Cell* 164 (2016) 1110–1121.
- [70] V. Pirrotta, Transvection and chromosomal trans-interaction effects, *Biochim. Biophys. Acta* 1424 (1999) M1–M8.
- [71] D. Noordermeer, E. de Wit, P. Klous, H. van de Werken, M. Simonis, M. Lopez-Jones, et al., Variegated gene expression caused by cell-specific long-range DNA interactions, *Nat. Cell Biol.* 13 (2011) 944–951.
- [72] B.J. Beliveau, A.N. Boettiger, M.S. Avendano, R. Jungmann, R.B. McCole, E.F. Joyce, et al., Single-molecule super-resolution imaging of chromosomes and in situ haplotype visualization using Oligopaint FISH probes, *Nat. Commun.* 6 (2015) 7147.
- [73] A.N. Boettiger, B. Bintu, J.R. Moffitt, S. Wang, B.J. Beliveau, G. Fudenberg, et al., Super-resolution imaging reveals distinct chromatin folding for different epigenetic states, *Nature* 529 (2016) 418–422.
- [74] D.I. Cattoni, A.M. Cardozo Gizzi, M. Georgieva, M. Di Stefano, A. Valeri, D. Chamoussat, et al., Single-cell absolute contact probability detection reveals chromosomes are organized by multiple low-frequency yet specific interactions, *Nat. Commun.* 8 (2017) 1753.
- [75] Q. Szabo, D. Jost, J.M. Chang, D.I. Cattoni, G.L. Papadopoulos, B. Bonev, et al., TADs are 3D structural units of higher-order chromosome organization in *Drosophila*, *Sci. Adv.* 4 (2018), eaar8082.
- [76] B. Bintu, L.J. Mateo, J.H. Su, N.A. Sinnott-Armstrong, M. Parker, S. Kinrot, et al., Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells, *Science* 362 (2018).
- [77] E.H. Finn, G. Pegoraro, H.B. Brandao, A.L. Valton, M.E. Oomen, J. Dekker, et al., Extensive heterogeneity and intrinsic variation in spatial genome organization, *Cell* 176 (2019) 1502–1515 e1510.
- [78] A.M. Cardozo Gizzi, D.I. Cattoni, J.-B. Fiche, S.M. Espinola, J. Gurgo, O. Messina, et al., Microscopy-based chromosome conformation capture enables simultaneous visualization of genome organization and transcription in intact organisms, *Mol. Cell* 74 (2019) 212–222, e215.
- [79] L.J. Mateo, S.E. Murphy, A. Hafner, I.S. Cinquini, C.A. Walker, A.N. Boettiger, Visualizing DNA folding and RNA in embryos at single-cell resolution, *Nature* 568 (2019) 49–54.
- [80] P. Olivares-Chauvet, Z. Mukamel, A. Lifshitz, O. Schwartzman, N.O. Elkayam, Y. Lubling, et al., Capturing pairwise and multi-way chromosomal conformations using chromosomal walks, *Nature* 540 (2016) 296–300.
- [81] R.A. Beagrie, A. Scialdone, M. Schueler, D.C. Kraemer, M. Chotalia, S.Q. Xie, et al., Complex multi-enhancer contacts captured by genome architecture mapping, *Nature* 543 (2017) 519–524.
- [82] S.A. Quinodoz, N. Ollikainen, B. Tabak, A. Palla, J.M. Schmidt, E. Detmar, et al., Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus, *Cell* 174 (2018) 744–757 e724.
- [83] M. Zheng, S.Z. Tian, D. Capurso, M. Kim, R. Maurya, B. Lee, et al., Multiplex chromatin interactions with single-molecule precision, *Nature* 566 (2019) 558–562.
- [84] A.M. Oudelaar, J.O.J. Davies, L.L.P. Hanssen, J.M. Telenius, R. Schwesinger, Y. Liu, et al., Single-allele chromatin interactions identify regulatory hubs in dynamic compartmentalized domains, *Nat. Genet.* 50 (2018) 1744–1751.
- [85] N. Servant, N. Varoquaux, B.R. Lajoie, E. Viara, C.J. Chen, J.P. Vert, et al., HiC-Pro: an optimized and flexible pipeline for Hi-C data processing, *Genome Biol.* 16 (2015) 259.
- [86] G. Andrey, T. Montavon, B. Mascres, F. Gonzalez, D. Noordermeer, M. Leleu, et al., A switch between topological domains underlies HoxD genes collinearity in mouse limbs, *Science* 340 (2013) 1195.
- [87] J.H. Sun, L. Zhou, D.J. Emerson, S.A. Phyo, K.R. Titus, W. Gong, et al., Disease-associated short tandem repeats Co-localize with chromatin domain boundaries, *Cell* 175 (2018) 224–238 e215.