



## Review

## Structural and functional diversity of Topologically Associating Domains

Job Dekker<sup>a,\*</sup>, Edith Heard<sup>b,\*</sup><sup>a</sup> Howard Hughes Medical Institute, Program in Systems Biology, Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, 368 Plantation Street, Worcester, MA 01605, USA<sup>b</sup> Institut Curie, CNRS UMR3215, INSERM U934, 26 Rue d'Ulm, 75248 Paris Cedex 05, France

## ARTICLE INFO

## Article history:

Received 13 August 2015

Accepted 28 August 2015

Available online 5 September 2015

Edited by Wilhelm Just

## Keywords:

Topologically associating domain

Long-range gene regulation

Chromatin folding

## ABSTRACT

Recent studies have shown that chromosomes in a range of organisms are compartmentalized in different types of chromatin domains. In mammals, chromosomes form compartments that are composed of smaller Topologically Associating Domains (TADs). TADs are thought to represent functional domains of gene regulation but much is still unknown about the mechanisms of their formation and how they exert their regulatory effect on embedded genes. Further, similar domains have been detected in other organisms, including flies, worms, fungi and bacteria. Although in all these cases these domains appear similar as detected by 3C-based methods, their biology appears to be quite distinct with differences in the protein complexes involved in their formation and differences in their internal organization. Here we outline our current understanding of such domains in different organisms and their roles in gene regulation.

© 2015 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

### 1. Introduction

The organization of chromosomes inside the cell nucleus is closely related to regulation of gene expression [1–5]. At the nuclear level this is apparent in the well-known spatial separation of active and inactive chromatin, where heterochromatic loci tend to be near the periphery and actively transcribed genes are localized more internal [1,6]. This separation is also observed within chromosome territories: chromosomes are divided into large multi-Mb compartments that contain either active and open (A-compartments) or inactive and closed chromatin (B-compartments) [7,8]. A compartments cluster with other A compartments, as do B compartments with B compartments. Given that different cell types express different gene sets driven by distinct groups of regulatory elements, the positions of A- and B-compartments change accordingly. Thus, global nuclear organization reflects a high level of compartmentalization that is directly correlated with the cell type-specific gene expression and chromatin status of the genome. However the exact nature of chromosome organization at the sub-megabase scale, which is the level at which most gene regulatory landscapes and long range interactions are thought to occur [9–12] had remained somewhat of a blackbox.

Recently, chromosome conformation capture (3C, [13]) experiments have uncovered the presence of an additional level of compartmentalization at this scale. Throughout the genomes of a wide range of species from bacteria to human, chromosomes are organized as a string of domains. These domains are characterized by preferential chromatin interactions within them, and spatial separation of loci located in different domains. In mammalian genomes these domains are several hundred kb in size, up to 1–2 Mb [14,15], whereas they are smaller in flies (~60 kb) [16,17], and bacteria (~170 kb) [18]. In eukaryotes these domains are referred to as Topological Domains [15] or Topologically Associating Domains ([14]; here referred to as TADs), or as Chromatin Interaction Domains (CIDs) in bacteria [18]. TADs are distinct from A- and B-compartments as they are smaller and largely cell type invariant (see below).

Whether and how any of these chromosomal domains directly contribute to regulation of the genome, e.g. gene expression is less clear. One reason is that control of gene expression is usually thought to occur at a much smaller scale. For instance, genes can be regulated by distal regulatory elements such as enhancers. Enhancers are thought to act over tens of kilobases, up to hundreds of kilobases at most, regulating nearby genes but not necessarily, or exclusively, the closest gene [2,12,19]. Enhancers may regulate target genes by direct looping interactions with their promoters (e.g. [2,12,20,21]). There are now many examples of such interactions, but the molecular mechanisms by which these loops are formed, their dynamics and how these interactions activate

\* Corresponding authors.

E-mail addresses: [Job.Dekker@umassmed.edu](mailto:Job.Dekker@umassmed.edu) (J. Dekker), [edith.heard@curie.fr](mailto:edith.heard@curie.fr) (E. Heard).

expression remains poorly understood. Another unresolved question is what determines specificity of long-range promoter–enhancer interactions. Given that enhancers can apparently loop to reach genes hundreds of kilobases away it is not known how *bona fide* target genes are identified and/or inappropriate interactions are prevented. It has been proposed that TADs play roles in regulating gene expression by either facilitating or preventing looping interactions [3,4,22], which would point to mechanistic links between chromosome compartmentalization, chromatin folding and regulation of gene expression (see below).

Here we will outline our current understanding of chromatin domains, focusing on TADs and CIDs. After describing their structural features and commonalities and differences between species we will present evidence that these domains form key structures involved in gene regulation by defining target regions of regulatory elements. We will then outline outstanding questions and propose future approaches to delineate and dissect the mechanisms of chromatin domain formation, enhancer action and transcription.

## 2. Self-interacting chromosomal domains are present in a wide range of organisms

Here we focus on chromosomal domains that are defined by the increased contact probability of loci located within them, which is readily detected in chromosome conformation capture experiments. In such experiments, e.g. 3C, 5C and Hi-C, comprehensive chromatin interaction datasets are obtained that can be represented as two-dimensional interaction heatmaps, where the genomic coordinates of the interacting pairs of loci are displayed along the two axes (Fig. 1). Chromatin interaction maps typically display a very prominent diagonal that reflects the very frequent contacts between loci located close to each other in the linear genome [23]. Analysis of local and genome-wide chromatin interaction maps for mouse and human genomes, as well as in the fruitfly *Drosophila*, led to the first observation of self interacting chromatin domains, that are apparent in chromatin interaction maps as a series of squares of relatively high interaction frequency along the diagonal [14–16]. These squares represent contiguous regions where loci interact with each other relatively frequently. Self-interacting domains are separated by sharp boundaries that appear to structurally insulate adjacent domains from each other as indicated by the fact that loci located in neighboring domains display a much lower contact frequency [14,15,24].

Self-interacting chromosomal domains have now been detected in bacteria, fungi, flies, nematodes, and in mammals. As detected by chromosome conformation capture experiments these domains all stand out as regions of increased contact frequency, but there are fundamental differences related to their structure and size, the processes and proteins that determine their formation, and possibly the mechanisms by which they affect chromatin state and gene expression. Below we describe the structural features of these domains in different organisms.

## 3. Topologically Associating Domains in mammals

The first evidence for TADs in mammals came from a 5C analysis of the X chromosome inactivation center (Xic) in mouse ESCs and differentiated cells [14] (Fig. 1), and a genome-wide Hi-C study in mouse and human cells [15]. Microscopy studies had previously hinted at globular chromosome structures in the megabase size range, possibly built up of approximately 100 kb domains [25,26]. Super-resolution DNA FISH across the Xic region revealed that TADs might indeed represent such physical entities of preferentially associating chromatin at the single cell level, as FISH probes were found to intermingle more frequently within TADs

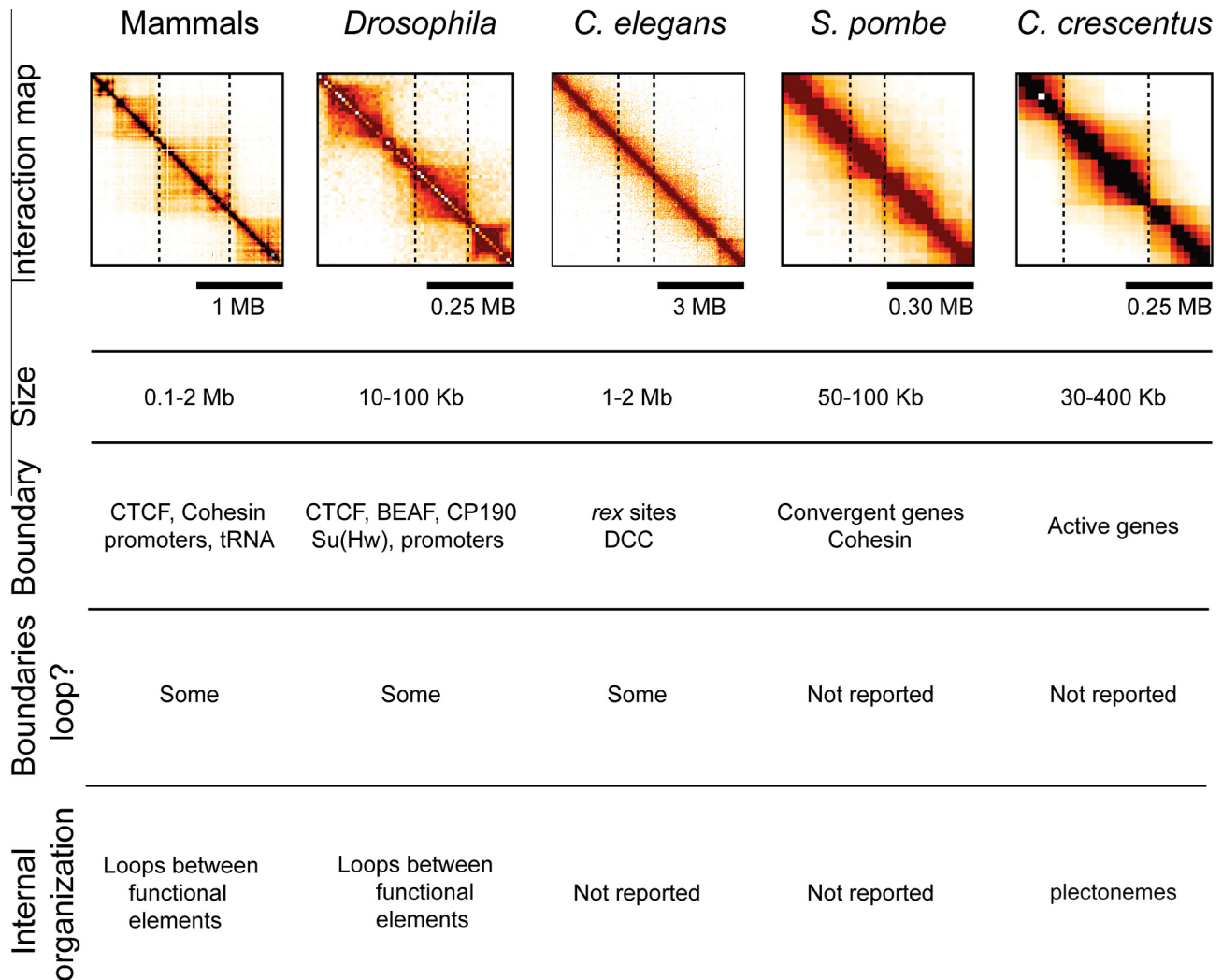
than between them [14,27]. It should be noted that data from FISH and 3C-based techniques are not always concordant, however, and that chromatin interactions, and/or FISH detection at some genomic regions may be subject to specific influences [28].

TADs in mammalian genomes range from tens of kb up to 1 or 2 Mb, with an average of around 800 kb. Two remarkable features of mammalian TADs are their relative invariance during differentiation [14,15] and their general conservation in relative position (though not necessarily in size) between man and mouse [15,29]. Whether this conservation is at the level of boundaries between TADs, or the interacting regions within TADs, remains an open question and indeed the nature of the underlying sequences that are conserved may vary from one genomic region to another, as will be discussed below. Although TADs are generally present and invariant, there are specific cases where TADs are not present. First, the inactive X chromosome appears to be depleted of TADs [14,30,31] and shows rather random interactions along its length [32]. Second, TADs are not observed along mitotic chromosomes [33]. Thus, even in mammalian cells, TADs can be absent in some cases, e.g. during chromosome-wide transcriptional silencing and chromosome condensation, although these processes may be driven by distinct molecular mechanisms.

Although TADs seem to be relatively invariant, the long-range sequence interactions within them can vary significantly between cell types and during differentiation, with specific long-range interactions appearing, while others are lost [14]. Many of these dynamic changes can be linked to the regulatory enhancer–promoter interaction events that orchestrate transcription during development (e.g. [34]), for which many classic examples, such as the  $\beta$ -globin locus [35] have already been described. The emerging picture is that TADs encompass the regulatory landscapes of genes, and that the meeting of enhancers with their target promoters happens usually, if not always, in the context of TADs [36–38].

An important question thus concerns what underlies TAD formation and how sequence interactions are restricted to occur within, but not between domains. An obvious mechanism would be that boundaries between TADs have specific insulating properties. An alternative but not mutually exclusive mechanism could be that sequence interactions within TADs are sufficient to ensure spatial segregation, although asymmetry in interactions (i.e. preferential interactions within the domain as compared to interactions between domains) must be provided somehow. Deletion of a boundary at the Xic locus resulted in aberrant interactions between previously separate TADs and misregulation of genes, presumably due to *de novo* enhancer–promoter interactions. Importantly, this deletion resulted in only partial fusion of the two adjacent TADs, with the appearance of a new boundary within one TAD implying that in the context of *de novo* interactions, a novel boundary can actually form [14]. More recent studies of structural variations that disrupt TAD boundaries in the context of malformation syndromes reveal that ectopic interactions leading to aberrant gene expression can be caused by disruption of boundary elements rather than merely by distance effects, highlighting the importance of TADs and their boundaries in genomic compartmentalization and normal gene regulation [37] at least in some regions.

Mammalian TAD boundaries are reported to be enriched in active transcription, housekeeping genes, tRNA genes and short interspersed nuclear elements (SINES), as well as binding sites for the architectural proteins CTCF and cohesin [15]. However such binding sites also exist within TADs, and CTCF and cohesin depletion reduce the intensity of intra-TAD interactions without affecting overall TAD location or organization [39–41]. This is consistent with their putative role in mediating enhancer–promoter contacts within TADs but leaves open the question of their role at boundaries between TADs.



**Fig. 1.** Overview of the features of TADs in different species. Heatmaps represent Hi-C interaction maps. Mammals: Mouse Xic region in ES cells [14]; *Drosophila*: chrX:4000001–4550001 (S2 cells, unpublished data Dekker lab); *C. elegans* embryos chrX:5760001–12780001 [24]. *S. pombe*: chr3:300001–610001 [55]; *Caulobacter*: 1040001–1620001 [18].

How CTCF and cohesin organize chromatin in such a way as to prevent interactions between particular TADs and isolate gene expression states from one another still remains unclear. Such insulation can occur through a local activity of protein complexes bound to an individual TAD boundary, or through formation of boundary-boundary interactions (e.g. through CTCF and/or cohesin bound to each [24,42,43] leading to a “looped configuration”. Chromatin looping leads to physical insulation of loci located within the loop from loci outside the loop [44]. Consistent with this model, new insights into the finer details of TAD organization, revealed intriguing orientation-specific looping interactions between CTCF sites at domain boundaries [42]. This study identified smaller contact domains within TADs, in the order of 100–200 kb, containing multiple specific loops that occur between CTCF sites in a predominantly (>90%) convergent orientation, with asymmetric motifs “facing” one another. Thus (0.2–2 Mb) TADs appear to represent just one level of folding in a more intricate hierarchy – something already hinted at in previous Hi-C and 5C maps where smaller domains in the order of tens of kilobases, as well as much larger domains spanning a few megabases are visible [14,15]. The precise relationship between TAD boundaries and these CTCF-anchored loops and contact domains, is not known in detail yet.

Clearly facultative enhancer–promoter interactions cannot underlie the apparent stability of TADs during development.

Indeed, some long-range interactions within mammalian TADs appear to be invariant [14], raising the interesting possibility that architectural elements, distinct from regulatory elements such as enhancers and promoters, may exist. Consistent with this, approximately one third of the long-range interactions mediated by cohesin and CTCF do not involve enhancer and promoter sequences in mammalian cells [43]. Recent physical modeling of 5C data at the Xic points to the existence of such structural elements required for TAD formation [27]. A recent study applying Hi-C to four different mammals revealed that the modular organization of chromosomes is robustly conserved in syntenic regions and that this is compatible with conservation of the CTCF binding landscape [29]. The most highly conserved CTCF sites were found to co-localize with cohesin and to be enriched at strong TAD boundaries. Furthermore, CTCF DNA motif orientations defined the directionality of the long-range interactions [42]. On the other hand divergent CTCF binding between species correlated with divergence of internal domain structure. Furthermore, the authors found that TADs are reorganised as intact modules during evolution, providing further support that TADs represent functional domains of long-range gene regulation.

An understanding of the stability and dynamics of the long range interactions within TADs will be critical to assess how this level of chromatin folding impacts on TAD structure and gene

expression. Indeed, a predictive physical model of the chromatin fiber suggests that some TADs represent domains of probabilistic interactions between the sequences lying within them, rather than to stable looping structures [27].

Hi-C data previously demonstrated that the genome is partitioned into distinct compartments [8]. The relationships between compartments and TADs are still being explored. Compartments represent large (up to several Mb) chromosomal domains defined by their preferential interactions with other compartments, whereas TADs are defined by the preferential interactions within them. Compartments tend to interact with other compartments that share their chromatin and/or transcriptional state: chromosomal regions enriched in active (A compartment) or inactive (B compartment) chromatin preferentially interact. Recent high-resolution Hi-C maps suggest that A and B-compartment can be further split into several sub-types [42]. Compartments can encompass several directly adjacent TADs that share chromatin state and that display similar genome-wide interactions with other sets of TADs. This has led to a model where interphase chromosome organization is a hierarchy of chromatin domains with TADs as the universal building blocks [4]; tissue invariant TADs come together in 3D space to form larger cell type-specific compartments. Compartment differences between cell types are due to relocation of entire TADs from one compartment type to another. This model accommodates the observation that TAD boundary positions are mostly invariant across tissues, while the chromatin states within TADs can change dramatically in different cell types and conditions, reflecting changes in gene activity and leading to altered compartment associations. A further prediction of this model is that the principles underlying compartment formation may be rather different to those underlying TADs. Indeed, the former are likely to depend on chromatin associated factors such as trithorax or polycomb; whereas TADs and their boundaries are not dependent on such factors, and instead rely on architectural proteins such as CTCF.

#### 4. Topologically Associating Domains in flies

The partitioning of the *Drosophila* genome into approximately 1000 physical domains each in the range of tens to 100 kb (average 60 kb; Fig. 1), that may be equivalent to TADs, was first described using a genome-wide 3C analysis (3C-seq) on early embryos by the Cavalli lab [16]. Similar domains were identified by the Corces lab by Hi-C analysis of kc167 cells [17]. These domains were found to correlate strongly with epigenomic features, including histone modifications, active gene density, association with the nuclear lamina, replication timing, nucleotide and repetitive element composition. Many of the physical domains identified by Hi-C could thus be classified into previous, statistically defined epigenomic groups [45] e.g. active domains (domains showing active transcription), repressive domains (at the nuclear periphery), Polycomb and HP1 domains bound by Polycomb group complexes and HP1 respectively and null domains, lacking specific epigenetic marks. Although the precise association between physical domains and their epigenomic status still remains unclear, recent studies focusing on Polycomb-repressed domains suggest that they correspond to cooperative interactions among low-affinity sequences, DNA-binding factors such as PHO, and the Polycomb machinery, with PHO recruitment to sites within Polycomb domains being stabilized by PRC1. On the other hand, chromosomal domains categorised as active show rather distinct folding patterns, with more rapid decay in contact frequency as a function of genomic distance than other domains. The local structure of active domains may thus be rather different to repressive domains [46]. However, whether physical domains or TADs provide a basic chromosome

architecture onto which epigenomic domains are laid down, or whether epigenomic demarcation is involved in the formation or maintenance of a TAD will require genetic disruption of the enzymes involved and Hi-C assessment. In support of the former model we note that in mammals, it would appear that disruption of large domains of H3K27me3 or H3K9me2 at the Xic locus did not impact on TAD segmentation [14] although an impact on local compaction could not be ruled out entirely.

As in mammals, the specification of TAD boundaries in *Drosophila* probably relies at least in part, on architectural proteins [16,17]. Unlike mammals however, numerous DNA binding architectural proteins, including CTCF, have been identified in *Drosophila*, each recognizing a unique DNA motif [47]. There are also multiple accessory proteins, in addition to Rad21 (cohesin) that can associate with these DNA binding proteins. The specific combinations of architectural and accessory proteins at different genomic regions, as well as the number and orientation of their binding sites can easily be imagined to produce a diversity of 3D organization states, that can vary in a cell type specific fashion. Very little is currently known about the differences in chromosome folding states between tissues or developmental stages. One study found extensive looping between functional elements that was stable across development [48]. On the other hand, a recent study investigated the changes induced during heat shock [49]. Temperature stress induced a dramatic rearrangement in 3D chromosome organization, with the relocalization of architectural proteins from TAD boundaries to sites within TADs, leading to an increase in long-distance inter-TAD interactions, with increased contacts among enhancers and promoters of silenced genes. These results reinforce the notion that architectural protein complexes play critical role in TAD boundary formation.

#### 5. Topologically Associating Domains in *Caenorhabditis elegans*

Recently the first genome-wide chromatin interaction map for *C. elegans* embryos was obtained by combining conventional 3C with deep sequencing [24] (Fig. 1). This map revealed known features of *C. elegans* nuclear organization, e.g. the tethering of large multi-Mb domains near the ends of the chromosomes to the nuclear lamina [50]. As a result of the peripheral localization of these domains, they interact with each other as well, both in cis and in trans, leading to the formation of higher order nuclear “compartments”, comparable to those observed in mammalian cells.

Perhaps surprisingly, no strong TADs were observed along the five autosomes, although some weak TAD boundaries could be detected. However, TADs are present along the two X-chromosomes of *C. elegans* in hermaphrodites. This is in contrast to the genome of *Drosophila* that has a genome of comparable size and complexity (e.g. gene number), and where chromosomal domains are prominently present along all chromosomes [16,17]. Thus, TADs are clearly not a universal feature of metazoan chromosomes. Indeed, no TADs have been observed in *Arabidopsis* [51–53]. The global lack of TADs in *C. elegans* and *Arabidopsis* may be related to the fact that long-range enhancers do not appear to be required for developmental gene regulation in these two organisms. Another difference between *Drosophila* and *C. elegans* is that TADs in *C. elegans* are considerably larger: 1–2 Mb compared to ~60 kb.

The fact that TADs are present along the X-chromosomes is interesting because in hermaphrodites gene expression along the X chromosomes is repressed by a factor of two to make the expression similar to that in males that carry only a single X chromosome. This chromosome-wide process of dosage compensation is specific to the X chromosome and is mediated by the condensin-like Dosage Compensation Complex (DCC). In mutants that cannot



recruit the DCC to the X-chromosomes, leading to loss of dosage compensation, most of the TAD boundaries were no longer detected, or strongly reduced in strength. This observation points to a direct role of the DCC and the process of dosage compensation in formation of many, but not all, TADs along X in *C. elegans*.

The DCC is recruited to the X chromosomes through binding to *rex* (recruitment on X) sites [54]. Interestingly, the strongest TAD boundaries on X contain strong *rex* sites. These are also the TAD boundaries that are most affected in DCC mutants. Further, deletion of *rex* sites from a TAD boundary is sufficient to eliminate the boundary. Thus, DCC binding to *rex* sites is critical for TAD boundary formation.

The molecular mechanism by which the DCC induces TAD formation is not known. One intriguing finding is that *rex* sites at TAD boundaries engage in DCC-dependent long-range looping interactions, especially with the adjacent *rex*-containing neighboring TAD boundary over 1 Mb away [24]. These results show that the DCC induces and reinforces TADs through binding high-affinity *rex* sites and mediating long-range looping interactions between them. This is reminiscent of TADs in humans and flies, where at least a subset of them display chromatin loops between CTCF-bound sites located within their boundaries [17,42]. The role of TAD formation in dosage compensation along the X chromosome is still an open question.

## 6. Chromatin globules in *Schizosaccharomyces pombe*

Self-interacting chromatin domains have also been detected in *S. pombe*, referred to as “globules” [55] (Fig. 1). These globules are 50–100 kb in size and are found all along the genome. In this case globule boundaries are enriched for 3' ends of convergent genes. Such convergent sites are bound by the cohesin complex. Interestingly, in a partial loss-of-function cohesin mutant globule boundaries are lost pointing to important roles of the cohesin complex in domain formation in this organism. *S. pombe* cultures are mainly composed of G2 cells, and thus globules could be related to the prominent sister chromatid cohesion in G2. Importantly however, globules were also observed in G1 cells, and these were cohesin dependent, indicating that globule formation depends on an activity of the cohesin complex that is separate from its role in sister chromatid cohesion.

The functional consequences of globule formation are not known in detail. In partial loss-of-function cohesin mutants, when globule formation is affected, widespread aberrant transcriptional read-through is observed. However, whether this is due to loss of globules *per se* or due to other functions of the cohesin complex is not known.

## 7. Chromatin Interaction Domains in bacteria

Hi-C and 5C analyses of the spatial organization of the circular *Caulobacter crescentus* genome showed that the chromosome adopts an elongated structure where the origin of replication is anchored at one pole of the cell with the two chromosome arms running in parallel along the length of the cell [18,56]. The higher resolution Hi-C data further revealed the presence of self-interacting chromatin domains, referred to Chromatin Interaction Domains (CIDs) that are on average 170 kb (30–420 kb) in size [18] (Fig. 1). Several lines of evidence indicate that active transcription plays a major role in CID formation. First, almost all boundaries contain highly transcribed genes. Second, genetically relocating an active gene also relocates the associated CID boundary. Third, blocking transcription by addition of rifampicin disrupted CIDs and reduced their boundaries. Thus, in *C. crescentus* transcription, esp. at CID boundaries is a major driver of domain

formation, possibly by forming locally unwound and plectoneme-free stretches of DNA at boundaries. This is in contrast to TADs in metazoans that are independent of transcription.

A fundamental difference between prokaryotic chromatin and eukaryotic chromatin is the fact that in bacteria DNA is nucleosome-free and supercoiled leading to the formation of plectonemes. Modeling of the *C. crescentus* genome as a circular chromosome composed of a series of plectonemes, on average 13 kb in size, with plectoneme-free areas at CID boundaries produced predicted Hi-C maps that are very similar to the experimentally observed data. This suggests that each CID contains around 10–15 such plectonemes that can migrate throughout the CID but cannot pass the plectoneme-free regions at the actively transcribed genes at their boundaries. Consistent with supercoiling playing a major role, treatment of cells with novobiocin that inhibits gyrase and negative supercoiling, reduced the sharpness of CID boundaries as well as their locations.

The data from *C. crescentus* provides a striking example of the fact that self-interacting chromatin domains can be observed in many genomes, but that the molecular mechanisms of their formation, and the folding of the DNA within them can be very different in different species.

## 8. TADs are functional domains

There is now growing and strong evidence that TADs are critical chromosome structural units of long-range gene regulation. The first data relating TADs to gene expression came from analysis of expression patterns of genes located within the same TAD across ES cell differentiation [14]. It was found that genes embedded in the same TAD show similar dynamics of expression during differentiation, whereas genes located in different TADs were less correlated. Further, some TADs correspond to Lamin Associated Domains (LADs), or domains covered by certain histone modifications such as H3K9Me2 and H3K27Me3, which all mark repressed chromatin states. This data indicates that at least some TADs are units of chromatin state and histone modification that correlates with regions of gene repression. Importantly, deletion of genes encoding enzymes that deposit such histone modifications, results in loss of the modifications, while TADs are maintained [14]. This shows that TADs are not the result of formation of domains of histone modifications, but that instead histone-modifying complexes act on pre-existing TADs to regulate chromatin state, and possibly gene expression, at the level of the entire domain. Further evidence that TADs can be regulated as units is provided by experiments where gene expression in T47D cells was induced by addition of nuclear hormones such as progesterone [46]. It was found that up to 20% of the TADs behaved as discrete regulatory units where the majority of the genes embedded within them are either activated or repressed.

As mentioned above mammalian TAD positions are to a significant extent conserved between different cell types, and even between mouse and human. Despite this universal and cell-type invariant architecture, TADs are believed to be involved in highly tissue-specific gene regulation. First, using correlation analysis across large panels of cell types, target genes of cell-type specific enhancer-like elements have been predicted [11,57]. These predictions identified enhancer–promoter pairs that are significantly enriched for pairs located within the same TAD [11]. Further, we have found that long-range looping interactions between promoters and distal regulatory elements detected by 5C are highly cell-type specific but occur within generally invariant TADs (Smith, Lajoie, Jain and Dekker, unpublished results). These observations may explain another feature of TADs that is readily observed in 5C and high-resolution Hi-C datasets: whereas the boundaries of

TADs are highly conserved between cell types, the internal folding and interaction patterns of TADs are highly cell type-specific [14,15,34,42] and may represent intra-TAD loops between genes and regulatory elements.

Perhaps the strongest evidence that TADs correspond to functional domains is provided by a completely independent approach. Simmons and co-workers employed an enhancer trap-like strategy and generated mice with a reporter sensor construct inserted at different positions along the chromosome [38]. Analysis of expression patterns of this panel of reporters identified functional chromosomal domains: wherever the reporter is inserted within such domain, the expression pattern is the same. These results suggest that enhancers exert their activities throughout such regulatory domains to control cell type-specific expression of any receptive promoter within the region. Importantly, these regulatory domains identified based solely on gene expression patterns show a remarkable correlation with TADs. Therefore, TADs are structural as well as functional units of gene regulation.

Recently two studies have shown that genetic rearrangements that affect TAD organization alter gene expression by changing patterns of long-range enhancer–promoter interactions. An inversion on chromosome 3 is implicated in AML. It was found that this inversion disrupts two TADs at the breakpoints and results in formation of new hybrid TADs containing parts of each flanking genomic region [58]. As a result, an enhancer that normally regulates the *GATA2* gene is repositioned and now located within the TAD that contains the *EVI* oncogene. This enhancer activates the oncogene and contributes to tumor formation. At the same time the *GATA2* gene is no longer located within the TAD containing the enhancer and this leads to *GATA2* haploinsufficiency which is implicated in sporadic familial AML/MDS and MonoMac/Emberger syndromes.

A second example is provided by naturally occurring genomic rearrangements involved in human limb malformations. Reconstruction of such alterations in mice showed that these affect TAD boundaries and as a result lead to ectopic enhancer–promoter connections that normally do not happen as these elements are located in different TADs [37]. Specifically, different rearrangements place a set of *Eph4* enhancers within the same TAD as *WNT*, *IHH*, or *PAX3*, depending on the precise nature of the TAD reorganization, leading to inappropriate interactions between the enhancers and the promoters of these genes.

Finally, RNAi-mediated knock down, or cell-type specific knock-out of cohesin subunits and CTCF has confirmed that these protein complexes play roles in TAD formation [40,41,59]. These studies found that removal of these complexes leads to weakening of TAD boundaries with concomitant gain of ectopic interactions between genes and regulatory elements located in different TADs. This also results in some altered gene expression. It should be noted though, that the effects on chromatin organization and transcription are rather modest, which could indicate that other factors are also involved, or that the knock downs are not complete and remaining levels of proteins are sufficient for maintaining significant chromatin structure. Still, overall these examples of functional effects of genetic perturbations of TAD organization point to deep mechanistic relationships between chromatin domain formation and gene regulation.

We note that information on TAD organization might also prove powerful in interpreting genome-wide association studies. Such studies typically identify non-coding regions linked to disease that likely contain gene regulatory elements. Current data outlined above suggest that TAD organization will help predict target genes for these regulatory elements: target genes should be located within the same TAD. Thus, insights into the domainal organization of chromosomes and its relation to long-range gene regulation can

contribute to uncovering the molecular mechanisms of the genetic basis of disease.

## 9. Other roles of TADs, globules and CIDs

Besides roles in gene regulation, TADs have also been linked to patterns of DNA replication [60]. Replication timing fluctuates along chromosomes in units of several hundred kilobases. Intriguingly, almost all TAD borders were found to be located at borders of replication domains in at least some cell types. In a given cell type, series of adjacent TADs can all replicate early, but the transition to a late replication domains occurs at TAD boundaries. At which TAD boundary this transition in replication timing happens can depend on the cell type. Consistently, TADs typically replicated as a whole either early or late, and switched replication timing as units during differentiation in accordance with changes in their transcriptional activity and chromatin state. The mechanisms by which replication timing is regulated at the TAD level are not known in detail.

Most functional studies have been focused on TADs in mammalian cells. Whether insights obtained from these studies can be extrapolated to chromatin domains in other organisms remains an open question. For instance, it is not known whether globules in *S. pombe* play similar roles in transcriptional control and DNA replication as TADs in mammals. Fungi such as *S. pombe* are not thought to regulate genes through long-range interactions between promoters and distal enhancers. Thus, it is not clear whether globules play similar roles as TADs in constraining such looping interactions. As mentioned above, cohesin mutants disrupt globule formation and also affect 3' ends processing of transcript pointing to different roles of these structures in gene regulation.

CIDs appear to be a fundamentally different type of chromatin domain than TADs and globules, despite a similar appearance in Hi-C interaction maps. As mentioned above CIDs have been proposed to be composed of a series of migrating plectonemes that are blocked at CID boundaries. Supercoiling of DNA is important for gene expression in bacteria, and CIDs may be important for regulation of gene expression by constraining supercoiling and preventing local dissipation of plectonemes. On the other hand, CIDs may form simply as a result of local DNA unwinding at boundaries, and may not play an active role in transcriptional control. Much more work is required to elucidate the mechanistic relationships between transcription, supercoiling and CID formation.

Finally, TADs in *C. elegans* present another example where they may play roles other than constraining looping interactions between genes and enhancers. In this case, the formation of TADs along the X chromosome somehow impacts gene expression uniformly all along the chromosome. How this is accomplished is not known. It is interesting to note that these TADs differ from mammalian TADs in that they depend on a condensin-related complex. On the other hand, *C. elegans* TADs also share features with mammalian TADs, including the presence of looping interactions between TAD boundaries, and their overall large (Mb) size.

Clearly, while TAD-like chromatin domains are observed across organisms, the mechanisms of their formation and the protein complexes involved, their internal organization and their functional roles in regulating genomic processes such as transcription and replication differ greatly.

## 10. Outstanding questions and future studies

The presence of TAD-like chromatin domains in a range of organisms is now well established, but molecular insights into the mechanisms of their formation, and their roles in regulating a range of genomic activities are still largely lacking. Major questions include: (1) When and how (during development, during

the cell cycle) are TADs established and how are they maintained? (2) How can adjacent TAD-like domains be prevented from mixing? Is chromatin looping between TAD boundaries sufficient for such spatial separation? (3) What is the internal organization of TADs? Is supercoiling involved in eukaryotic TADs, as it is in CID formation in bacteria? Does looping within TADs play a role in TAD formation and/or stabilization? (4) What is the internal dynamics of chromatin folding within TADs in real time in single cells? Are enhancer–promoter looping interactions within TADs stable, or dynamic? All of these questions need to be addressed in the various model organisms described above, as it is likely that different mechanisms are at work.

Some of the outstanding questions can now be experimentally addressed by genetic perturbation approaches using genome editing tools such as those based the CRISPR/Cas9 system. There are already examples where targeted deletion of TAD boundaries [37] or binding sites of candidate protein complexes [24] were introduced followed by analysis of the conformation of the chromatin by 3C-based methods, and by analysis of effects on local gene expression.

To gain insights into the dynamics of TAD-like domains, live cell imaging will be essential. Again CRISPR/Cas9-based tools could be employed to visualize targeted elements within TADs and correlate their interactions over time, and possible relate such interactions with local transcription.

Further, the molecular machines that fold chromatin at the level of TADs need to be identified and their mechanism of action elucidated. Whether the proteins that drive chromatin folding are different from those that maintain it (through the cell cycle, during DNA repair etc.) must also be explored. Several protein complexes are already known, including cohesin, condensin, and CTCF. These, and their associated molecules, provide fruitful starting points but there are likely other complexes involved as well.

Given the important roles of TADs in gene control and other processes, deeper insights into their biology promises to lead to a better understanding of how cells regulate their genome and how genetic variants can lead to inappropriate gene expression over hundreds of kilobases leading to disease.

## Acknowledgements

We would like to thank members of the Heard and Dekker Labs for discussions and ideas. Supported by Grants from the National Human Genome Research Institute (R01 HG003143) to J.D., an ERC Advanced Investigator award, EU FP7 Grants SYBOSS and MODHEP, La Ligue, Labex DEEP (ANR-11-LBX-0044) part of the IDEX Idex PSL (ANR-10-IDEX-0001-02 PSL) to E.H.

## References

- [1] Bickmore, W.A. (2013) The spatial organization of the human genome. *Annu. Rev. Genomics Hum. Genet.* 14.
- [2] de Laat, W. and Duboule, D. (2013) Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* 502, 499–506.
- [3] Nora, E.P., Dekker, J. and Heard, E. (2013) Segmental folding of chromosomes: a basis for structural and regulatory chromosomal neighborhoods? *BioEssays* 35, 818–828.
- [4] Gibcus, J.H. and Dekker, J. (2013) The hierarchy of the 3D genome. *Mol. Cell* 49, 773–782.
- [5] Sexton, T., Schober, H., Fraser, P. and Gasser, S.M. (2007) Gene regulation through nuclear organization. *Nat. Struct. Mol. Biol.* 14, 1049–1055.
- [6] Cremer, T. and Cremer, C. (2001) Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat. Rev. Genet.* 2, 292–301.
- [7] Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B. and de Laat, W. (2006) Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat. Genet.* 38, 1348–1354.
- [8] Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., Sandstrom, R., Bernstein, B., Bender, M.A., Grroudine, M., Gnirke, A., Stamatoyannopoulos, J.A., Mirny, L., Lander, E.S. and Dekker, J. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289–293.
- [9] Kleinjan, D.A. and van Heyningen, V. (2005) Long-range control of gene expression: emerging mechanisms and disruption in disease. *Am. J. Hum. Genet.* 76, 8–32.
- [10] Kleinjan, D.J. and Coutinho, P. (2009) Cis-rupture mechanisms: disruption of cis-regulatory control as a cause of human genetic disease. *Brief Funct. Genomic Proteomic* 8, 317–332.
- [11] Jin, F., Li, Y., Dixon, J.R., Selvaraj, S., Ye, Z., Lee, A.Y., Yen, C.A., Schmitt, A.D., Espinoza, C.A. and Ren, B. (2013) A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* 503, 290–294.
- [12] Sanyal, A., Lajoie, B.R., Jain, G. and Dekker, J. (2012) The long-range interaction landscape of gene promoters. *Nature* 489, 109–113.
- [13] Dekker, J., Rippe, K., Dekker, M. and Kleckner, N. (2002) Capturing chromosome conformation. *Science* 295, 1306–1311.
- [14] Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., Gribnau, J., Barillot, E., Blüthgen, N., Dekker, J. and Heard, E. (2012) Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 485, 381–385.
- [15] Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S. and Ren, B. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376–380.
- [16] Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A. and Cavalli, G. (2012) Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell* 148, 458–472.
- [17] Hou, C., Li, L., Qin, Z.S. and Corces, V.G. (2012) Gene density, transcription, and insulators contribute to the partition of the drosophila genome into physical domains. *Mol. Cell* 48, 471–484.
- [18] Lee, T.B., Imakaev, M.V., Mirny, L.A. and Laub, M.T. (2013) High-resolution mapping of the spatial organization of a bacterial chromosome. *Science* 342, 731–734.
- [19] Schoenfelder, S., Furlan-Magaril, M., Mifsud, B., Tavares-Cadete, F., Sugar, R., Javierre, B.M., Nagano, T., Katsman, Y., Sakthidevi, M., Wingett, S.W., Dimitrova, E., Dimond, A., Edelman, L.B., Elderkin, S., Tabbada, K., Darbo, E., Andrews, S., Herman, B., Higgs, A., LeProust, E., Osborne, C.S., Mitchell, J.A., Luscombe, N.M. and Fraser, P. (2015) The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements. *Genome Res.* 25, 582–597.
- [20] Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F. and de Laat, W. (2002) Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol. Cell* 10, 1453–1465.
- [21] Vernimmen, D., De Gobbi, M., Sloane-Stanley, J.A., Wood, W.G. and Higgs, D.R. (2007) Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. *EMBO J.* 26, 2041–2051.
- [22] Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenko, V.V. and Ren, B. (2012) A map of the cis-regulatory sequences in the mouse genome. *Nature* 488, 116–120.
- [23] Dekker, J., Marti-Renom, M.A. and Mirny, L.A. (2013) Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat. Rev. Genet.* 14, 390–403.
- [24] Crane, E., Bian, Q., McCord, R.P., Lajoie, B.R., Wheeler, B.S., Ralston, E.J., Uzawa, S., Dekker, J. and Meyer, B.J. (2015) Condensin-driven remodeling of X-chromosome topology during dosage compensation. *Nature* 523, 240–244.
- [25] Cremer, T. and Cremer, M. (2010) Chromosome territories. *Cold Spring Harb. Perspect. Biol.* 2, a003889.
- [26] Markaki, Y., Gunkel, M., Schermelleh, L., Beichmanis, S., Neumann, J., Heidemann, N., Leonhardt, H., Eick, D., Cremer, C. and Cremer, T. (2010) Functional nuclear organization of transcription and DNA replication: a topographical marriage between chromatin domains and the interchromatin compartment. *Cold Spring Harb. Symp. Quant. Biol.* 75, 475–492.
- [27] Giorgetti, L., Galupa, R., Nora, E.P., Piolot, T., Lam, F., Dekker, J., Tiana, G. and Heard, E. (2013) Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. *Cell* 157, 950–963.
- [28] Williamson, I., Berlivet, S., Eskeland, R., Boyle, S., Llingworth, R.S., Paquette, D., Dostie, J. and Bickmore, W.A. (2014) Spatial genome organization: contrasting views from chromosome conformation capture and fluorescence in situ hybridization. *Genes Dev.* 28, 2778–2791.
- [29] Vietri Rudan, M., Barrington, C., Henderson, S., Ernst, C., Odom, D.T., Tanay, A. and Hadjur, S. (2015) Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell Rep.* 10, 1297–1309.
- [30] Minajigi, A., Froberg, J.E., Wei, C., Sunwoo, H., Kesner, B., Colognori, D., Lessing, D., Payer, B., Boukhali, M., Haas, W. and Lee, J.T. (2015) A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* 349, <http://dx.doi.org/10.1126/science.aab2276>. pii: aab2276.
- [31] Deng, X., Ma, W., Ramani, V., Hill, A., Yang, F., Ay, F., Berletch, J.B., Blau, C.A., Shendure, J., Duan, Z., Noble, W.S. and Distche, C.M. (2015) Bipartite structure of the inactive mouse X chromosome. *Genome Biol.* 16, 152.
- [32] Splinter, E., de Wit, E., Nora, E.P., Klous, P., van de Werken, H.J., Zhu, Y., Kaaij, L. J., van Ijcken, W., Gribnau, J., Heard, E. and de Laat, W. (2011) The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. *Genes Dev.* 25, 1371–1383.



- [33] Naumova, N., Imakaev, M., Fudenberg, G., Zhan, Y., Lajoie, B.R., Mirny, L.A. and Dekker, J. (2013) Organization of the mitotic chromosome. *Science* 342, 948–953.
- [34] Phillips-Cremins, J.E., Sauria, M.E., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S., Ong, C.T., Hookway, T.A., Guo, C., Sun, Y., Bland, M.J., Wagstaff, W., Dalton, S., McDevitt, T.C., Sen, R., Dekker, J., Taylor, J. and Corces, V.G. (2013) Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* 153, 1281–1295.
- [35] Palstra, R.J., Tolhuis, B., Splinter, E., Nijmeijer, R., Grosveld, F. and de Laat, W. (2003) The beta-globin nuclear compartment in development and erythroid differentiation. *Nat. Genet.* 35, 190–194.
- [36] Anderson, E., Devenney, P.S., Hill, R.E. and Lettice, L.A. (2014) Mapping the Shh long-range regulatory domain. *Development* 141, 3934–3943.
- [37] Lupiáñez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R., Santos-Simarro, F., Gilbert-Dussardier, B., Wittler, L., Borschiwer, M., Haas, S.A., Osterwalder, M., Franke, M., Timmermann, B., Hecht, J., Spielmann, M., Visel, A. and Mundlos, S. (2015) Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* 161, 1012–1025.
- [38] Symmons, O., Uslu, V.V., Tsujimura, T., Ruf, S., Nassari, S., Schwarzer, W., Ettwiller, L. and Spitz, F. (2014) Functional and topological characteristics of mammalian regulatory domains. *Genome Res.* 24, 390–400.
- [39] Sofueva, S., Yaffe, E., Chan, W.C., Georgopoulou, D., Vietri Rudan, M., Mira-Bontenbal, H., Pollard, S.M., Schroth, G.P., Tanay, A. and Hadjir, S. (2013) Cohesin-mediated interactions organize chromosomal domain architecture. *EMBO J.* 32, 3119–3129.
- [40] Seitan, V., Faure, A., Zhan, Y., McCord, R., Lajoie, B., Ing-Simmons, E., Lenhard, B., Giorgetti, L., Heard, E., Fisher, A., Flicek, P., Dekker, J. and Merkenschlager, M. (2013) Cohesin-based chromatin interactions enable regulated gene expression within pre-existing architectural compartments. *Genome Res.* 23, 2066–2077.
- [41] Zuin, J., Dixon, J.R., van der Reijden, M.I., Ye, Z., Kolovos, P., Brouwer, R.W., van de Corput, M.P., van de Werken, H.J., Knoch, T.A., van Ijcken, W.F., Grosveld, F. G., Ren, B. and Wendt, K.S. (2014) Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc. Natl. Acad. Sci. U.S.A.* 111, 996–1001.
- [42] Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S. and Lieberman Aiden, E. (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665–1680.
- [43] Downen, J.M., Fan, Z.P., Hnisz, D., Ren, G., Abraham, B.J., Zhang, L.N., Weintraub, A.S., Schuijers, J., Lee, T.I., Zhao, K. and Young, R.A. (2014) Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* 159, 374–387.
- [44] Doyle, B., Fudenberg, G., Imakaev, M. and Mirny, L.A. (2014) Chromatin loops as allosteric modulators of enhancer–promoter interactions. *PLoS Comput. Biol.* 10, e1003867.
- [45] Fillion, G.J., van Bommel, J.G., Braunschweig, U., Talhout, W., Kind, J., Ward, L.D., Brugman, W., de Castro, I.J., Kerkhoven, R.M., Bussemaker, H.J. and van Steensel, B. (2010) Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell* 143, 212–224.
- [46] Le Dily, F., Baù, D., Pohl, A., Vicent, G.P., Serra, F., Soronellas, D., Castellano, G., Wright, R.H., Ballare, C., Filion, G., Marti-Renom, M.A. and Beato, M. (2014) Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes Dev.* 28, 2151–2162.
- [47] Ceballos-Potts, C. and Corces, V.G. (2015) Architectural proteins, transcription, and the three-dimensional organization of the genome. *FEBS Lett.* pii: S0014-5793(15)00401-9.
- [48] Ghavi-Helm, Y., Klein, F.A., Pakozdi, T., Ciglar, L., Noordermeer, D., Huber, W. and Furlong, E.E. (2014) Enhancer loops appear stable during development and are associated with paused polymerase. *Nature* 512, 96–100.
- [49] Li, L., Lyu, X., Hou, C., Takenaka, N., Nguyen, H.Q., Ong, C.T., Ceballos-Potts, C., Hu, M., Lei, E.P., Bosco, G., Qin, Z.S. and Corces, V.G. (2015) Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing. *Mol. Cell* 58, 216–231.
- [50] Ikegami, K., Egelhofer, T.A., Strome, S. and Lieb, J.D. (2010) *Caenorhabditis elegans* chromosome arms are anchored to the nuclear membrane via discontinuous association with LEM-2. *Genome Biol.* 11, R120.
- [51] Moissiard, G., Cokus, S.J., Cary, J., Feng, S., Billi, A.C., Stroud, H., Husmann, D., Zhan, Y., Lajoie, B.R., McCord, R.P., Hale, C.J., Feng, W., Michaels, S.D., Frand, A. R., Pellegrini, M., Dekker, J., Kim, J.K. and Jacobsen, S. (2012) MORC family ATPases required for heterochromatin condensation and gene silencing. *Science* 336, 1448–1451.
- [52] Feng, S., Cokus, S.J., Schubert, V., Zhai, J., Pellegrini, M. and Jacobsen, S.E. (2014) Genome-wide Hi-C analyses in wild-type and mutants reveal high-resolution chromatin interactions in *Arabidopsis*. *Mol. Cell* 55, 694–707.
- [53] Grob, S., Schmid, M.W. and Grossniklaus, U. (2014) Hi-C analysis in *Arabidopsis* identifies the KNOT, a structure with similarities to the flamenco locus of *Drosophila*. *Mol. Cell* 55, 678–693.
- [54] Jans, J., Gladden, J.M., Ralston, E.J., Pickle, C.S., Michel, A.H., Pferdehirt, R.R., Eisen, M.B. and Meyer, B.J. (2009) A condensin-like dosage compensation complex acts at a distance to control expression throughout the genome. *Genes Dev.* 23, 602–618.
- [55] Mizuguchi, T., Fudenberg, G., Mehta, S., Belton, J.M., Taneja, N., Folco, H.D., FitzGerald, P., Dekker, J., Mirny, L., Barrowman, J. and Grewal, S.I. (2014) Cohesin-dependent globules and heterochromatin shape 3D genome architecture in *S. pombe*. *Nature* 516, 432–435.
- [56] Umbarger, M.A., Toro, E., Wright, M.A., Porreca, G.J., Baù, D., Hong, S., Fero, M.J., Zhu, L.J., Marti-Renom, M.A., McAdams, H.H., Shapiro, L., Dekker, J. and Church, G.M. (2011) The three-dimensional architecture of a bacterial genome and its alteration by genetic perturbation. *Mol. Cell* 44, 252–264.
- [57] Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., Garg, K., John, S., Sandstrom, R., Bates, D., Boatman, L., Canfield, T.K., Diegel, M., Dunn, D., Ebersol, A.K., Frum, T., Giste, E., Johnson, A.K., Johnson, E.M., Kutayavin, T., Lajoie, B.R., Lee, B.-K., Lee, K., London, D., Lotakis, D., Neph, S., Neri, F., Nguyen, E.D., Qu, H., Reynolds, A.P., Roach, V., Safi, A., Sanchez, M.E., Sanyal, A., Shafer, A., Simon, J.M., Song, L., Vong, S., Weaver, M., Yan, Y., Zhang, Z., Zhang, Z., Lenhard, B., Tewari, M., Dorschner, M.O., Hansen, R.S., Navas, P.A., Stamatoyannopoulos, G., Iyer, V.R., Lieb, J.D., Sunyaev, S.R., Akey, J.M., Sabo, P.J., Kaul, R., Furey, T.S., Dekker, J., Crawford, G.E. and Stamatoyannopoulos, J.A. (2012) The accessible chromatin landscape of the human genome. *Nature* 489, 75–82.
- [58] Gröschel, S., Sanders, M.A., Hoogenboezem, R., de Wit, E., Bouwman, B.A., Erpelinck, C., van der Velden, V.H., Havermans, M., Avellino, R., van Lom, K., Rombouts, E.J., van Duin, M., Döhner, K., Beverloo, H.B., Bradner, J.E., Döhner, H., Löwenberg, B., Valk, P.J., Bindels, E.M., de Laat, W. and Delwel, R. (2014) A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia. *Cell* 157, 369–381.
- [59] Hadjir, S., Williams, L.M., Ryan, N.K., Cobb, B.S., Sexton, T., Fraser, P., Fisher, A. G. and Merkenschlager, M. (2009) Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature* 460, 410–413.
- [60] Pope, B.D., Ryba, T., Dileep, V., Yue, F., Wu, W., Denas, O., Vera, D.L., Wang, Y., Hansen, R.S., Canfield, T.K., Thurman, R.E., Cheng, Y., Gülsoy, G., Dennis, J.H., Snyder, M.P., Stamatoyannopoulos, J.A., Taylor, J., Hardison, R.C., Kahveci, T., Ren, B. and Gilbert, D.M. (2014) Topologically associating domains are stable units of replication-timing regulation. *Nature* 515, 402–405.