genes

Review

Contribution of Topological Domains and Loop Formation to 3D Chromatin Organization

Vuthy Ea [1], Marie-Odile Baudement [1], Annick Lesne [1,2,*] and Thierry Forné [1,*]

[1] Institut de Génétique Moléculaire de Montpellier, UMR5535, CNRS, Université de Montpellier, 1919 Route de Mende, 34293 Montpellier cedex 5, France; E-Mails: vuthy.ea@igmm.cnrs.fr (V.E.); mbaudement@igmm.cnrs.fr (M.-O.B.)

[2] Laboratoire de Physique de la Matière Condensée, CNRS UMR 7600, UPMC, Sorbonne Universités, 75252 Paris, France

* Authors to whom correspondence should be addressed; E-Mails: thierry.forne@igmm.cnrs.fr (T.F.); annick.lesne@igmm.cnrs.fr or annick.lesne@lptmc.jussieu.fr (A.L.); Tel.: +33-467-34-35-96-82 (T.F.); Fax: +33-467-34-35-96-34 (T.F.).

Abstract: Recent investigations on 3D chromatin folding revealed that the eukaryote genomes are both highly compartmentalized and extremely dynamic. This review presents the most recent advances in topological domains' organization of the eukaryote genomes and discusses the relationship to chromatin loop formation. CTCF protein appears as a central factor of these two organization levels having either a strong insulating role at TAD borders, or a weaker architectural role in chromatin loop formation. TAD borders directly impact on chromatin dynamics by restricting contacts within specific genomic portions thus confining chromatin loop formation within TADs. We discuss how sub-TAD chromatin dynamics, constrained into a recently described statistical helix conformation, can produce functional interactions by contact stabilization.

Keywords: chromatin dynamics; topological domains; TAD borders; chromatin loops; CTCF; statistical helix

1. Introduction

In eukaryotic species, the assembly of DNA and histones into chromatin is essential to genome compaction and functions. Observations with optical and electron microscopy led to a classical bipartite description of eukaryotic chromatin. The gene-rich euchromatin is permissive to transcription and positioned at the center of the nucleus while heterochromatin, having a darkish shade, corresponds to a more compact/gene-poor form of chromatin and is generally positioned at the nuclear periphery. One can further distinguish two types of heterochromatin: constitutive heterochromatin is conserved from one cell type to another while facultative heterochromatin is not.

These cytological considerations have been recently renewed by the experimental access to the 3D organization of the genome in the nuclear space [1,2]. This exploration revealed that distinct organization levels exist at the supranucleosomal scale, between the nucleosome scale (nucleofilament) and the nuclear scale (chromosome territories) (Figure 1).
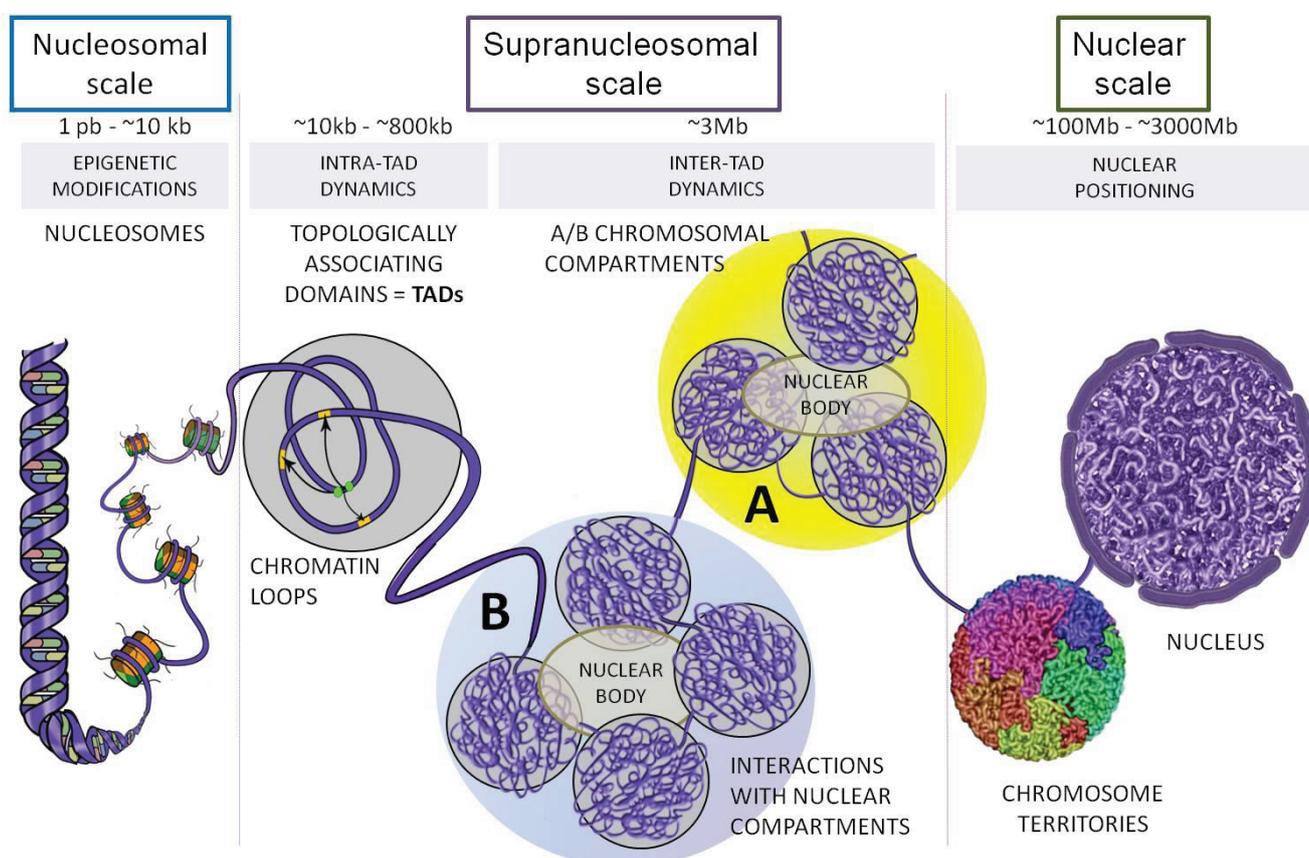

Figure 1. Schematic representation of genome organization in mammals. Between the nucleosomal scale (nucleofilament) and the nuclear scale (chromosome territories), 3D organization of the genome at the supranucleosomal scale has been recently explored thanks to 3C-derived methods (see Figure 2). Beyond the transcriptionally active (**A**) or inactive (**B**) chromosomal compartments, Topologically Associating Domains (TADs) and chromatin loops are two essential determinants of eukaryotic genome organization.

This breakthrough was made possible by the development of the Chromosome Conformation Capture (3C) assay and its combination with the Next-Generation Sequencing (NGS) technology allowing the

development of the so-called 4C, 5C and Hi-C methods (Figure 2). The advantages and limits of these methods are summarized in Table 1. They essentially differ by the dynamic ranges achieved for quantifying contact frequencies between distant chromatin sites and by the throughput of data generated.

Table 1. Advantages and limits of 3C-derived methods.

Method	Genomic Scale Investigated	Advantages	Limits
3C-qPCR	~250 kilobases	Very high dynamic range (highly quantitative), easy data analysis	Very low throughput: limited to few viewpoints in a selected region
4C	Complete genome	Good sensitivity at large separation distances	Genome-wide contact map limited to a unique viewpoint (few viewpoints if multiplex sequencing is used)
5C	Few megabases	Good dynamic range, complete contact map (all possible viewpoints) of a specific locus	The contact map obtained is limited to a selected region
Hi-C	Complete genome	Very high throughput (complete contact map)	Poor dynamic range, complex data processing

The first 4C (Chromosome Conformation Capture on ChIP) experiments, that evaluated for a given region all contacts occurring between this region and the whole genome, showed that chromosomal loci are not only able to establish close contacts in *cis* all along their chromosome, but also in *trans* with other chromosomes [3]. These experiments revealed that such contacts occur in active or inactive "chromatin hubs". This dichotomy between active and inactive chromatin was then confirmed by Hi-C experiments [4], leading to split the human and mouse genomes into two categories named the A and B compartments (also called "megadomains", [5]) (Figure 1). After developing a new Hi-C approach named TCC (Tethered Conformation Capture) that considerably reduces the background of inter-chromosomal contacts, Kalhor *et al.* [6] reached similar conclusions. Globally, along the chromosomes, one can distinguish transcriptionally active regions (A-compartments) from inactive regions (B-compartments) (Figure 1). Long-range homologous contacts (A–A or B–B) are largely favored over heterologous contacts (A–B). This organization principle seems to be universal, even if, from one cell type to another, a given chromosomal region can change from a class A compartment to a class B compartment [4].

This dichotomy appears very relevant both structurally and functionally. Indeed, A/B compartments also allow the integration of diverse experimental data [7], such as those issued from transcriptomics or histone modification and protein binding-site maps [8,9]. A-compartments are thus constituted of transcriptionally active gene-rich regions displaying high nuclease sensitivity, while B-compartments are gene-poor, transcriptionally silent and largely insensitive to nucleases [4]. However, this bimodal distribution does not totally reflect the *in vivo* situation since one can observe a continuum between A and B compartments that probably favors compartment switches during cell differentiation processes [10]. Finally, this genomic organization has been proposed to be coordinated through contacts with nuclear compartments or nuclear bodies [11] (Figure 1).

In this review, we present the finer 3D organization of chromatin compartments and discuss its relationships with genome functions and gene regulation.
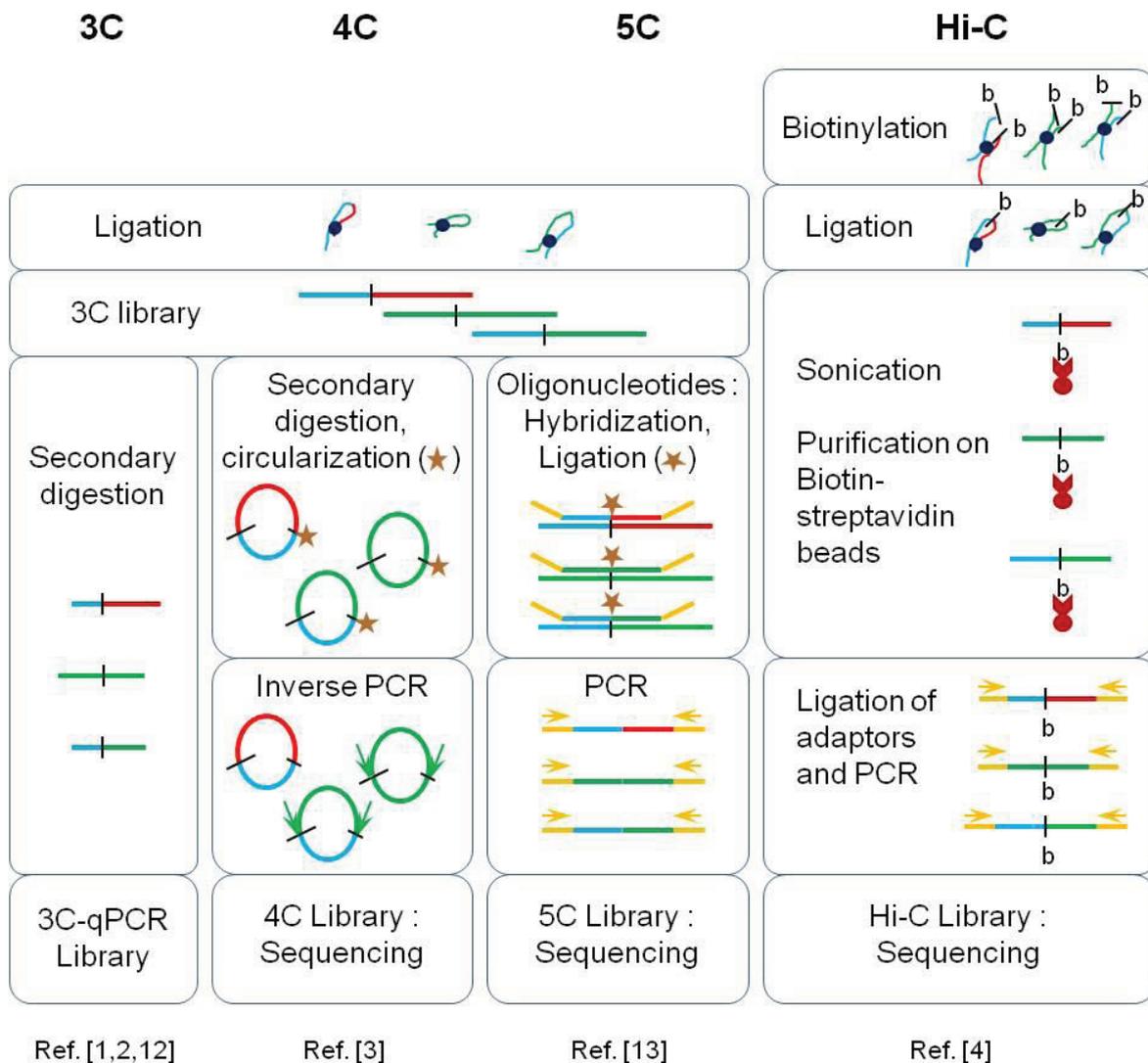
Figure 2. Principles of 3C-derived methods. All the methods derived from the Chromosome Conformation Capture (3C) protocol involve a formaldehyde cross-linking step followed by an enzymatic digestion and a ligation step. Each method uses different approaches to generate genomic libraries: secondary digestion for 3C-qPCR, circularization and inverse PCR for the 4C, "carbon copy" amplification for the 5C and biotinylation and purification on streptavidine beads for Hi-C. Ligation products are quantified by real-time qPCR in the first method or by high-throughput sequencing in the others (adapted from [14]).

2. Topologically Associating Domains (TADs)

2.1. Identifications of TADs and Their Conservation across Species

The organization of chromatin compartments was further refined in *Drosophila*, in mouse and in human, thanks to 5C and Hi-C experiments [15–18]. These experiments have shown that chromosomal compartments can in fact be subdivided into domains that display high internal contact frequencies, *i.e.*, contact frequencies of chromosomal regions located within these domains are much higher than across adjacent domains. Because of such preferential contacts, these domains have been named Topologically Associating Domains (TADs) (Figure 1).

The human and mouse genomes are composed of more than 2000 TADs having a median size lower than 1 Mb and covering, altogether, more than 90% of the entire genome (Table 2). TADs have also been evidenced from single-cell Hi-C experiments, indicating that they represent a genuine stable organization principle of mammalian genomes [19].

In the fly *Drosophila melanogaster*, for which genome size is about 30 times smaller than mammalian genomes (~130 Mb *vs.* 3400 Mb in the human), two independent studies have described more than 1100 TADs having a median size of about 60 kb (Table 2): the positions of 42% of TAD borders defined in these two studies coincide at ±4 kb [16,18].

The yeast genomes seem to be devoid of such TAD organization, as shown by Hi-C experiments performed in *Saccharomyces cerevisiae* [20] or in *Schizosaccharomyces pombe* [21] while, in plants, the genome of *Arabidopsis thaliana* appears organized through contacts with pericentromeric regions [22]. However, in these latter two organisms, the existence of domain-like organizations has been suggested [23,24]. Despite these differences, TADs appear as the basic functional units of a chromosome organization that seems widespread at least in the animal kingdom.

TAD organization is also confirmed by cell imaging approaches demonstrating that two loci located in the same TAD are on average closer in 3D space than two loci each located in different TADs [15,25]. More precisely, high-resolution FISH experiments show that two groups of probes tend to superimpose upon one another when they belong to the same TAD, but not when they belong to two different TADs [17].

Table 2. Characteristics of TADs identified in Hi-C experiments. Bin size determines the resolution achieved in the experiments. In addition to the references mentioned in this table, other works have been also performed in many other cell types (including neural precursors and astrocytes in the mouse; Hela S3 and HEK293T cells in the human) but without specifying the number and sizes of TADs [26–29].

Organism	Sample or Cell Line	Number of TADs	Median Size (kb)	Mean Size (kb)	Bin Size (kb)	Reference
M. musculus	E14 ESC	2200	880	1093		
	Cortex	1518	800	1063	20–40	[15]
H. sapiens	ESC	3127	680	852		
	IMR90	2348	840	1123		
D. melanogaster	Embryo	1169	62	100	~10	[18]
	Kc167	1110	61	107	4	[16]

2.2. Characteristics of TADs

Remarkably, TAD positions are similar from one cell type to another and they are also highly conserved between mouse and human. TAD stability appears as a powerful means to coordinate chromatin dynamics during cell differentiation. Indeed, in the mouse, cell-specific long-range interactions occurring for separation distances over 20 kb are almost systematically found in the same TAD (less than 4% take place between different TADs) [15]. Furthermore, in *Drosophila*, TADs can be classified according to their epigenetic landscape. At least four classes of chromatin, labeled as "colors", can be defined: transcriptionally active domains ("red domains", active domains) are associated to the trimethylation of the lysines 4 and 36 of histone H3 (H3K4me3 and H3K36me3); repressed chromatin

includes two kinds of domains, one linked to Polycomb group (PcG) proteins and to trimethylation of the lysine 27 of histone H3 (H3K27me3) ("blue domains", PcG domains) and the other associated to the HP1 and Su(var)3-9 heterochromatin proteins as well as to dimethylation of the lysine 9 of histone H3 (H3K9me2) ("green domains", HP1/centromeric domains); finally, the fourth class is not specifically associated to any chromatin mark investigated ("black domains", null domains) [18]. The euchromatin corresponds to active domains, the constitutive heterochromatin to HP1 domains and the facultative heterochromatin to PcG domains.

2.3. TAD Borders

A characteristic feature of TAD borders is that two loci each located on a side of a border establish much less contacts than loci located within the TAD [15,25]. 54% of TAD borders described in mouse Embryonic Stem Cells (ESC) are also found at identical positions in the mouse cortex and 43% of them are conserved in syntenic regions of the human genome. A similar observation was made in two studies performed in *Drosophila*: the first one investigating whole embryos [18] and the second one studying a homogenous cell line (Kc167) [16]. This indicates that borders that demarcate TADs are quite stable structures that possess specific topological characteristics.

However, the mechanisms involved in TAD establishment are not yet fully understood. In mammals, borders are enriched in Transcriptional Start Sites (TSS), more particularly TSSs of housekeeping genes. Accordingly, transcriptional activity in general appears to be high at TAD borders. TAD borders are also enriched in CTCF (CCCTC-binding Factor) insulator protein and Smc1 or Smc3, belonging to the cohesin complex, as well as in active histone marks (H3K4me3 and H3K36me3). Finally, TAD borders are enriched in Med1 and Med12 [15] that belong to the *Mediator* multiprotein complex, which is involved in transcription activation and is known to be important for maintaining cell pluripotency [30]. Similar enrichments are also observed in *Drosophila* where the CP190 insulator protein [31] as well as the *Chromator* factor and DNase hypersensitive sites accumulate at TAD borders. Enrichment in H3K4me3 mark is also found at TAD borders, but in this peculiar case one can observe a 500 bp shift of the enrichment peak toward the active domain [18].

In the mouse, the CTCF protein is found at more than 75% of TAD borders, but only 15% of CTCF sites correspond with TAD borders [15]. Indeed, this zinc-finger protein is highly conserved among mammals where it plays diverse important roles (for reviews see [32,33]). For example, CTCF controls gene expression during the X-inactivation process [34] or the V(D)J recombination in lymphocytes (see [35] for a review). CTCF is also involved at the parentally imprinted *Igf2/H19* locus, leading to distinct 3D organization of the parental alleles [36–38].

CTCF depletion affects essentially intra-TAD contacts at separation distances below 100 kb [28,29]. However, TAD borders become much more permissive and inter-TAD contacts are also significantly increased. Therefore, CTCF seems to be involved not only in strengthening intra-TAD interactions but also in preventing long-range contacts between adjacent TADs. These results [29] indicate that some CTCF sites play important roles for maintaining the integrity of TAD borders while so far, the importance of the cohesin complex is less clearly established.

Smc1 and Smc3 are two sub-units of the cohesin complex that also contains the Rad21 and SA1/2 proteins. This complex plays an essential role in maintaining the cohesion between sister chromatids

during DNA replication and cell division. CTCF and the cohesin complex map to numerous identical sites on the mammalian genome [30,39,40]. It is therefore not surprising that they are together enriched at TAD borders.

Globally, the insulator proteins identified at TAD borders appear to play a direct structural role in TAD organization [31,41], rather than to block the interactions between functional regulatory elements. Therefore, in that case, the term of "architectural proteins" may be more appropriate [42]. However, the position of a TAD probably also depends on its wider genomic context as shown by the deletion of a 58 kb region removing a TAD border at the X-Inactivation Center in the mouse (ΔXTX deletion): this removal perturbs TAD organization but only leads to an incomplete fusion of the adjacent TADs. In this case, other elements, located farther within the domains, contribute to limit inter-TADs contacts [17].

Finally, disruption of TAD borders (e.g., by impairment of binding of associated factors) [28,29] seems to have different architectural and functional effects. This hints at the possible existence of several types of TAD borders.

3. Sub-TAD Organization and Contact Domains

Given that TAD borders are highly stable across cell-types, the dynamic regulations of long-range contacts that are required for cell differentiation are expected to take place within TADs. Such contacts, possibly stabilized into cell type-specific interactions, would determine chromatin loops that can contribute to a sub-TAD organization regulated during developmental processes and maintained henceforth. In the mouse, using 5C experiments combined with Next Generation Sequencing (NGS), Phillips-Cremins *et al.* [43] have determined contact frequencies in 6 regions that are 1 to 2 Mb in size around genes important during early development (*Oct4*, *Nanog*, *Nestin*, *Sox2*, *Klf4* and *Olig1-Olig2*). These experiments, which achieved the highest genomic resolution available at that time, allowed identification of around 60 sub-TADs nested within the previously described TADs [43]. Some of them change during the differentiation of ESCs into Neural Progenitor Cells (NPCs) [15]. Furthermore, it was shown that CTCF and cohesin mediate developmentally stable interactions, while CTCF-free contacts often change [43].

Similarly, using 5C experiments, additional sub-TAD structures have been described at the mouse *HoxA* locus that appear to be specific of the developmental stages [44]. Globally, these results strengthen the idea that the genome may be hierarchically organized with a sub-TAD organization in chromatin loops linked to some cell type specificities whereas the TADs would represent stable units maintained throughout cell differentiation and development.

This sub-TAD organization was confirmed in human by high resolution *in situ* Hi-C experiments that identified about 10,000 loops (mostly linked to enhancer/promoter interactions) and defined contact domains having a median size of 185 kb [5]. Interestingly, similar to *Drosophila* TADs, such contact domains can be segregated into at least six classes according to their associated epigenetic landscape. A1 and A2 contact domains are gene dense and linked to active chromatin marks (such as H3K36me3, H3K27ac and H3K4me1), but A2 contact domains have a slightly lower G/C content and are more strongly associated to H3K9me3 than A1 contact domains. B1, B2 and B3 sub-TAD compartments are linked to repressive chromatin. B1 contact domains are linked to facultative heterochromatin, while the B2 compartments mostly include pericentromeric heterochromatin and are

strongly associated to Lamina-Associated Domains (LADs) [45,46] and Nucleolus-Associated Domains (NADs) [47]. Finally, the sub-compartment B3 is enriched in LADs but strongly depleted in NADs, while the sub-compartment B4 is restricted to a small 11 Mb portion of chromosome 19 containing both active chromatin (*i.e.*, H3K36me3) and heterochromatin-associated (*i.e.*, H3K9me3) marks [5].

About 65% of such chromatin domains correspond to chromatin loops, the vast majority of which involves a pair of convergent CTCF/RAD21/SMC3 binding sites. Contact domains also tend to be conserved across cell types and between human and mouse [5]. Interestingly, comparative Hi-C in four mammals (mouse/dog/macaque/rabbit) showed that a modular organization of chromosomes is conserved in syntenic regions in relation to the conservation of some CTCF binding sites [48]. Evolutionarily conserved CTCF sites enriched for RAD21 are co-occurring with TAD borders while divergent CTCF binding sites drive loop organization within TADs. Finally, high-resolution 4C-seq showed that conserved CTCF sites display strong directional interactions with neighboring conserved CTCF sites (*i.e.*, evolutionarily conserved orientations of CTCF binding sites) [48].

In conclusion, CTCF appears to have distinct functions in 3D organization of the chromatin: having either a strong insulating role at TAD borders [49,50], or a weaker architectural role in securing chromatin loops [31,51].

4. Defining TAD Borders *vs.* Contact Domain Borders and Loop Closures

Globally, the wealth of studies performed in mammals point to a modular model whereby stable TADs are themselves subdivided into multiple sub-TAD structures (chromatin loops) that undergo more dynamic cell-type specific rearrangements. However, our ability to discriminate these two organization levels relies on our ability to define "contact borders": how can TAD borders be distinguished from a simpler insulating activity mediated by chromatin loops? In Hi-C experiments, this is of course a matter of resolution since chromatin loops, having a relatively small size, can be better evidenced only in the higher-resolution experiments [5]. Moreover, accurate definition of contact borders also depends on our ability to accurately quantify local contact frequencies. It is thus tightly linked to the dynamic range achieved in each experiment, which depends not only on the deepness of sequencing, but also on the complexity of the 3C library. Indeed, while lowering resolution, using a 6-cutter restriction enzyme and binning the data through several tens of kilobases increases the dynamic range of the experiments thus allowing to better evidence TAD borders [15].

Finally, the best methods to distinguish TAD borders from chromatin loops are those having the best dynamic range. In that respect, the more classical quantitative 3C approaches [12] remain the best methods to characterize not only local chromatin loops, as largely exemplified in the literature for a vast number of loci (see for example [52]), but also TAD borders. Indeed, performing quantitative 3C experiments through TAD borders identified by Hi-C experiments can greatly help to assess whether such borders are more or less permissive to long-range *cis* inter-TAD contacts. To illustrate the relevance of quantitative 3C approaches, we present experimental results obtained in mouse embryonic Stem Cells (ESC) and Neural Precursor Cells (NPC) at two loci: the *HoxD* gene locus (Figure 3) and the *Ttll7* gene locus (Figure 4). The *HoxD* gene cluster has been described as precisely located at a TAD border [15] and higher resolution experiments have shown that a switch of enhancers between adjacent TADs underlies the collinearity of *HoxD* gene expression observed in mouse limbs [53]. The border

located at the *Ttll7* gene locus was identified in previous Hi-C experiments [15] but was not further characterized. For each locus, contact frequencies were determined from two viewpoints (anchors in the figures) located on the telomeric or centromeric sides of the borders (see Figures 3 and 4, a and b panels, respectively). At the *HoxD* gene locus, contact frequencies fall dramatically 61.6 kb upstream the telomeric viewpoint (red vertical dashed line in Figure 3a) and 127 kb downstream the centromeric viewpoint (red vertical dashed line in Figure 3b). This allows mapping the border within a gene-rich 38 kb large region delimited by two CTCF sites. In contrast, at the *Ttll7* gene locus, no drastic decrease could be evidenced, neither from the telomeric (Figure 4a) nor from the centromeric (Figure 4b) viewpoints. Therefore, while a large typical TAD border could be easily evidenced at the *HoxD* gene locus, no strong border could be defined at the *Ttll7* locus. However, a unique CTCF site located at the expected position may display the property of a weak border (red vertical dashed line in Figure 4b).
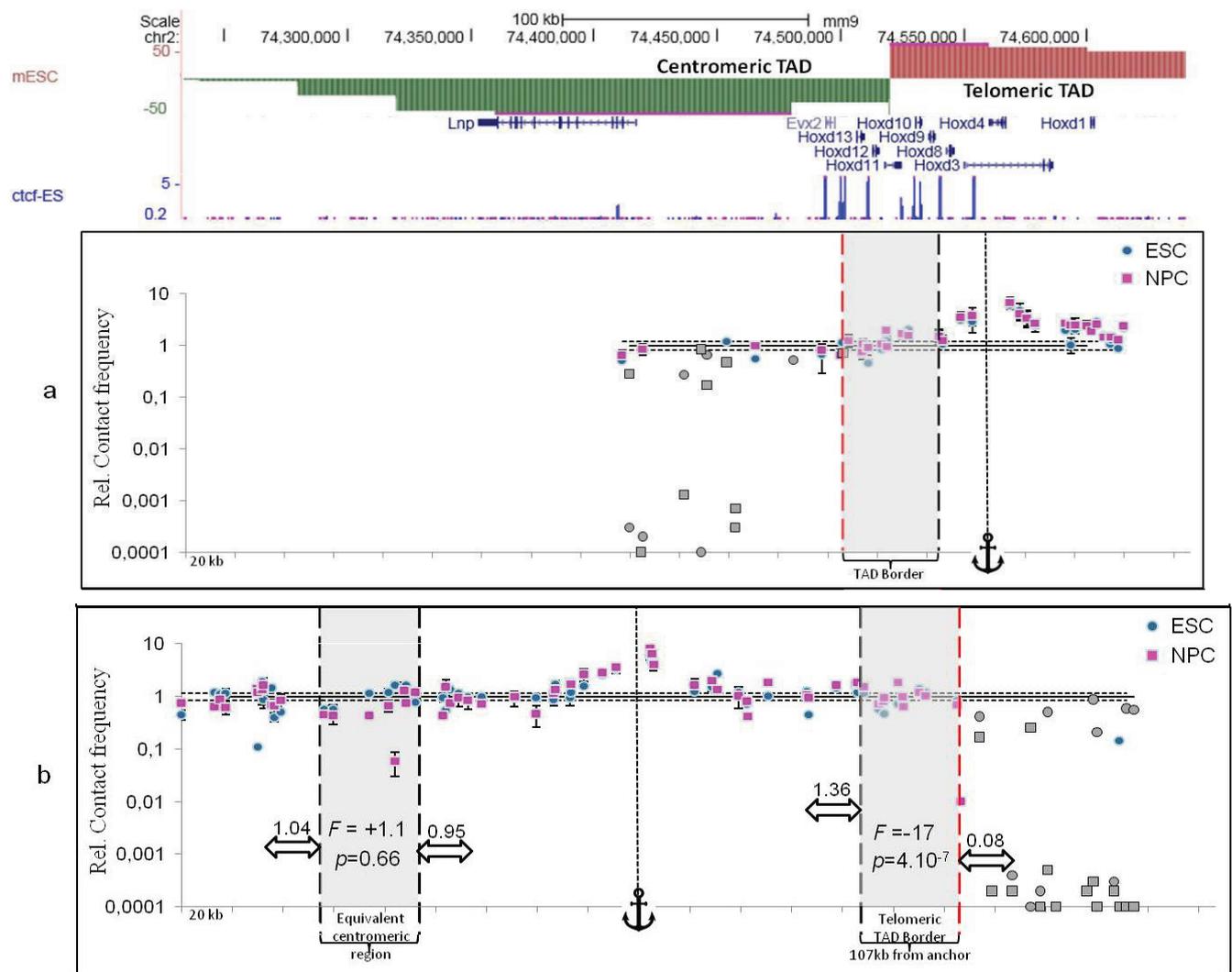

Figure 3. Contact frequencies at the *HoxD* locus in ESC and NPC. Contact frequencies have been determined at the *HoxD* locus (mouse chromosome 2) in mouse ESC (blue circles) and neural progenitor cells (NPC, pink rectangles) using 3C-qPCR as described in [54]. The data have been aligned with the directivity index and the CTCF sites previously described for mouse ESC [15]. The viewpoint used in each 3C experiment is indicated by an anchor. Contact frequencies have been measured from two different viewpoints located

in (**a**) the *Hoxd4* gene in the most telomeric TAD (colored in red) or (**b**) the *Lnp* gene in the most centromeric TAD (colored in green). The horizontal lines represent the mean basal contact level (continuous lines) and the associated background noise (dashed lines) [55]. The position of the TAD border is delimited by thick dashed vertical lines (note that this TAD border is delimited by two conserved CTCF binding sites). Error bars correspond to standard error of the mean of five biological replicates each quantified at least in triplicate. Experimental points for which the extremely low contact frequency could not be quantified in all of the five replicates are depicted in grey. In these cases, the points represented in the figure involve an upper bound of the contact frequency corresponding to the detection limits of our qPCR assays: for these points, a Ct of 45 cycles was assigned to replicates that could not be detected by quantitative PCR because of extremely low contact frequencies. These data points were then analyzed according to the same procedure as all the others experimental points as previously described [55]. Mean contact frequencies (values above white horizontal arrows) were calculated over 20 kb on each side of the TAD border (grey rectangle on the right) or on each side of an equivalent region located on the centromeric side of the same viewpoint (grey rectangle on the left). The fold change (ratio) *F* between these mean contact frequencies is indicated together with the *p*-value of their difference (Student *t*-test).

As our experiments are highly quantitative, we could take advantage of that feature to better characterize these two putative types of border. To this aim, we calculated the mean contact frequencies 20 kb upstream to the borders and we compared them with the mean contact frequencies observed 20 kb downstream. The strong TAD border observed at the *HoxD* gene locus, located 107 kb from the viewpoint on the telomeric side, is associated with a very strong decrease in contact frequencies: the mean contact frequency observed downstream to this border is 17 fold lower than that observed upstream to the border (Figure 3b, grey rectangle on the right). We also checked the significance of the difference between these means, as given by its *p*-value ($p = 4.10^{-7}$, Student *t*-test). Importantly, an equivalent region, located 107 kb on the centromeric side of the same viewpoint and displaying no evidence for any border, displays no significant change (fold change of 1.1 and difference with *p*-value of $p = 0.66$) (Figure 3b, grey rectangle on the left). These results demonstrate that the decrease observed on the telomeric side of this viewpoint corresponds to a specific feature linked to the presence of a strong border preventing long-range contacts between adjacent TADs.

In contrast, mean contact frequencies across the putative border identified at the *Ttll7* locus (CTCF site) display a fold change of only 1.7 (Figure 4b). While the difference between these mean contact frequencies is significant ($p = 2.10^{-3}$, Student *t*-test), their fold change is at least 10 times weaker than that observed for the TAD border of the *HoxD* locus (Figure 3b). This result confirms the presence of a weak border at this locus that may be explained, for instance, by the presence of a chromatin loop.

Globally, these results are also fully consistent with our previous observation [54] indicating that, in gene rich regions as those investigated here, contact frequencies are modulated with a period of 100 kb and that, consequently, contact frequencies of sites located around 100 kb are stable (Figure 3b, grey rectangle on the left) except, of course, if a strong TAD border (Figure 3b, grey rectangle on the right) or a weak border (Figure 4b) interferes with this underlying chromatin dynamics.

Interestingly, as shown above, the *HoxD* cluster thus appears as a highly structured locus, even in cells where it is repressed. Therefore, similar to sub-TAD chromatin loops, TAD borders can sometimes appear as specialized 3D structures extending over several tens of kilobases. However, TAD borders are stable structures and no or very low contacts can be detected through them (Figure 5).
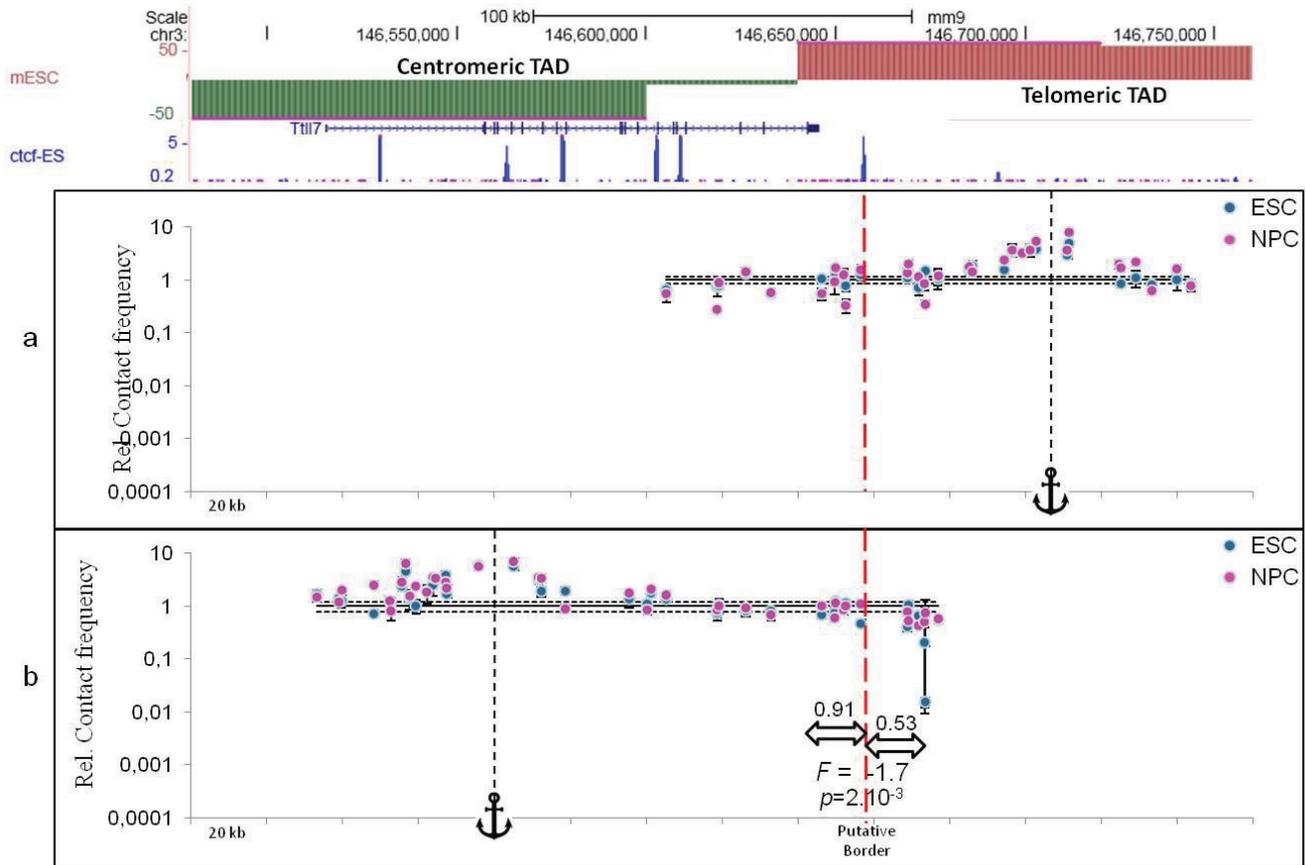

Figure 4. Contact frequencies at the *Ttll7* locus in ESC and NPC. Following the conventions described in Figure 3, contact frequencies have been determined at the *Ttll7* gene locus (mouse chromosome 3) in mouse ESC (blue circles) and neural progenitor cells (NPC, pink circles) from (**a**) an intergenic viewpoint located in the most telomeric TAD (colored in red) or (**b**) from an intragenic (*Ttll7* gene) viewpoint in the most centromeric TAD (colored in green). Error bars correspond to standard error of the mean of five biological replicates each quantified at least in triplicate. The most telomeric CTCF site appears as an interesting element delineating a putative border (dashed red vertical line). The fold change (*F*) between the mean contact frequencies on each side of this putative border and the *p*-value of their difference were determined as explained in the legend of Figure 3.

In contrast, sub-TAD 3D organization is characterized by its highly dynamic nature. Chromatin loops, mediated by stable locus-specific interactions, occur in that fluctuating context [56] (Figure 5). At this scale, conformational fluctuations of the chromatin fiber are primarily constrained by its own stiffness, which favors or restricts the formation of long-range contacts, as argued in Kleckner [57] in the context of meiosis. In interphase, with no anchoring of the chromatin loops onto a scaffold, the array of long-range contacts when observed in a large population of cells, as in 3C experiments, tends to fold

statistically the chromatin into a helical shape [54] (Figure 6, left part). This 3D organization is fully compatible with rosette-like structures (Figure 6, middle part) and is naturally prone to form chromatin loops provided that their closure may be achieved by stabilizing factors (Figure 5). However, for the sake of simplicity, such rosette-like structures are usually represented as a 2D projection taking the form of the chromatin loops classically represented in most models [5,58] (Figure 6 right part).
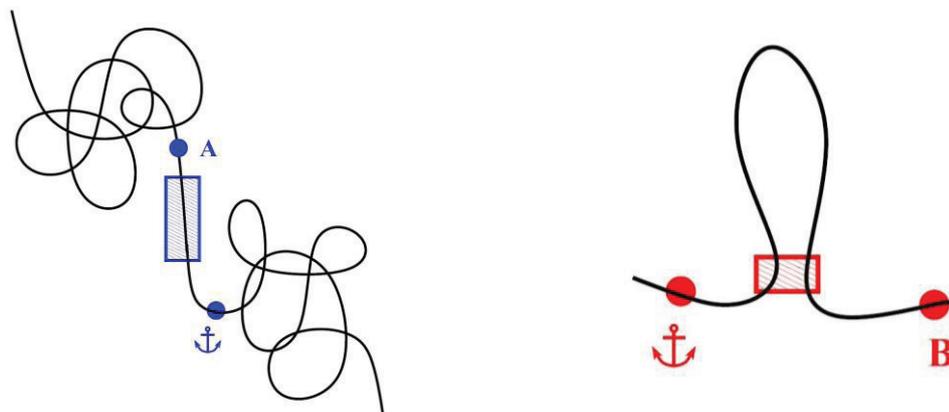

Figure 5. Two different kinds of borders: TAD borders and loop closures. The figure depicts a border separating two adjacent TADs along the genome, sketched as a blue box (**Left**), and a chromatin loop and its closure by stabilizing factors, sketched as a red box (**Right**). When performing a quantitative 3C experiment with the viewpoint represented by the anchor, TAD border plays a strong insulating role by preventing contacts between the viewpoint (anchor) and point A, as observed in Figure 3, thus demarcating two adjacent domains. In contrast, the formation of a chromatin loop, while locally increasing the contact frequency between the two stretches of DNA in the stem region (red rectangle), weakly lowers (or does not affect) the contact frequency between the anchor and point B. This weak-border signature may correspond to that seen in Figure 4.
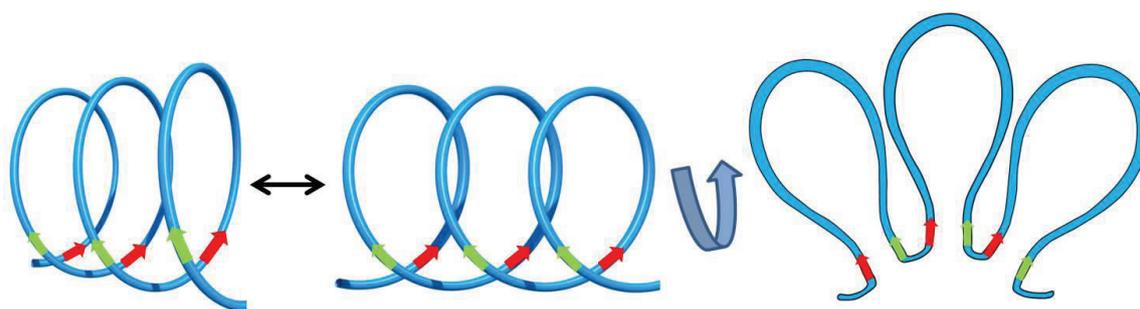

Figure 6. From statistical helix to 2D representation of chromatin loops. As recently shown [36,54], spontaneous conformational fluctuations of the chromatin fiber are constrained at supranucleosomal scale, which produces a helical statistical shape. Specific factors stabilize statistical contacts (*i.e.*, random collisions), turning them into stable interactions and creating 3D topological loops of chromatin [59]. 2D projection of such 3D rosette-like structures (as sketched by the curved arrow) leads to the current plane representation of chromatin loops [5,58]. Green and red arrows depict putative CTCF sites involved in chromatin loop formation and represent the directionality of these sites [5,48].

5. Conclusions

As the resolution, the dynamic range and the panel of methods available for studying chromatin dynamics increase, the organization principles of the eukaryote genomes become more and more clear. Our current view on eukaryote genomes points to a compartmentalized multi-scale organization that controls chromatin dynamics to restrict or favor intra- or inter-chromosomal contacts [11]. While TADs now appear as an essential unit of genome organization, exploring chromatin dynamics at the sub-TAD/chromatin-loop level is now essential to unravel the tight relationships that should exist between the functions and the fluctuating 3D organization of the eukaryote genomes.

Author Contributions

All authors contributed to conception and writing of this manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Dekker, J.; Rippe, K.; Dekker, M.; Kleckner, N. Capturing chromosome conformation. *Science* **2002**, *295*, 1306–1311.
2. Tolhuis, B.; Palstra, R.J.; Splinter, E.; Grosveld, F.; de Laat, W. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol. Cell* **2002**, *10*, 1453–1465.
3. Simonis, M.; Klous, P.; Splinter, E.; Moshkin, Y.; Willemsen, R.; de Wit, E.; van Steensel, B.; de Laat, W. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4c). *Nat. Genet.* **2006**, *38*, 1348–1354.
4. 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.; *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **2009**, *326*, 289–293.
5. Rao, S.S.; 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.; *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **2014**, *159*, 1665–1680.
6. Kalhor, R.; Tjong, H.; Jayathilaka, N.; Alber, F.; Chen, L. Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. *Nat. Biotechnol.* **2012**, *30*, 90–98.

7. Lesne, A.; Riposo, J.; Roger, P.; Cournac, A.; Mozziconacci, J. 3D genome reconstruction from chromosomal contacts. *Nat. Methods* **2014**, *11*, 1141–1143.

8. Kundaje, A.; Meuleman, W.; Ernst, J.; Bilenky, M.; Yen, A.; Heravi-Moussavi, A.; Kheradpour, P.; Zhang, Z.; Wang, J.; Ziller, M.J.; *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **2015**, *518*, 317–330.

9. Shen, Y.; Yue, F.; McCleary, D.F.; Ye, Z.; Edsall, L.; Kuan, S.; Wagner, U.; Dixon, J.; Lee, L.; Lobanenkov, V.V.; *et al.* A map of the cis-regulatory sequences in the mouse genome. *Nature* **2012**, *488*, 116–120.

10. Imakaev, M.; Fudenberg, G.; McCord, R.P.; Naumova, N.; Goloborodko, A.; Lajoie, B.R.; Dekker, J.; Mirny, L.A. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat. Methods* **2012**, *9*, 999–1003.

11. Gibcus, J.H.; Dekker, J. The hierarchy of the 3D genome. *Mol. Cell* **2013**, *49*, 773–782.

12. Hagège, H.; Klous, P.; Braem, C.; Splinter, E.; Dekker, J.; Cathala, G.; de Laat, W.; Forné, T. Quantitative analysis of chromosome conformation capture assays (3C-QPCR). *Nat. Protoc.* **2007**, *2*, 1722–1733.

13. Dostie, J.; Richmond, T.A.; Arnaout, R.A.; Selzer, R.R.; Lee, W.L.; Honan, T.A.; Rubio, E.D.; Krumm, A.; Lamb, J.; Nusbaum, C.; *et al.* Chromosome Conformation Capture Carbon Copy (5C): A massively parallel solution for mapping interactions between genomic elements. *Genome Res.* **2006**, *16*, 1299–1309.

14. De Wit, E.; de Laat, W. A decade of 3C technologies: Insights into nuclear organization. *Genes Dev.* **2012**, *26*, 11–24.

15. Dixon, J.R.; Selvaraj, S.; Yue, F.; Kim, A.; Li, Y.; Shen, Y.; Hu, M.; Liu, J.S.; Ren, B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **2012**, *485*, 376–380.

16. Hou, C.; Li, L.; Qin, Z.S.; Corces, V.G. Gene density, transcription, and insulators contribute to the partition of the *Drosophila* genome into physical domains. *Mol. Cell* **2012**, *48*, 471–484.

17. 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.; *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **2012**, *485*, 381–385.

18. Sexton, T.; Yaffe, E.; Kenigsberg, E.; Bantignies, F.; Leblanc, B.; Hoichman, M.; Parrinello, H.; Tanay, A.; Cavalli, G. Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **2012**, *148*, 458–472.

19. Nagano, T.; Lubling, Y.; Stevens, T.J.; Schoenfelder, S.; Yaffe, E.; Dean, W.; Laue, E.D.; Tanay, A.; Fraser, P. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **2013**, *502*, 59–64.

20. Duan, Z.; Andronescu, M.; Schutz, K.; McIlwain, S.; Kim, Y.J.; Lee, C.; Shendure, J.; Fields, S.; Blau, C.A.; Noble, W.S. A three-dimensional model of the yeast genome. *Nature* **2010**, *465*, 363–367.

21. Tanizawa, H.; Iwasaki, O.; Tanaka, A.; Capizzi, J.R.; Wickramasinghe, P.; Lee, M.; Fu, Z.; Noma, K. Mapping of long-range associations throughout the fission yeast genome reveals global genome organization linked to transcriptional regulation. *Nucleic Acids Res.* **2010**, *38*, 8164–8177.

22. Feng, S.; Cokus, S.J.; Schubert, V.; Zhai, J.; Pellegrini, M.; Jacobsen, S.E. Genome-wide Hi-C analyses in wild-type and mutants reveal high-resolution chromatin interactions in arabidopsis. *Mol. Cell* **2014**, *55*, 694–707.

23. Grob, S.; Schmid, M.W.; Grossniklaus, U. Hi-C analysis in arabidopsis identifies the knot, a structure with similarities to the flamenco locus of *Drosophila*. *Mol. Cell* **2014**, *55*, 678–693.

24. Mizuguchi, T.; Fudenberg, G.; Mehta, S.; Belton, J.M.; Taneja, N.; Folco, H.D.; FitzGerald, P.; Dekker, J.; Mirny, L.; Barrowman, J.; *et al.* Cohesin-dependent globules and heterochromatin shape 3D genome architecture in S. Pombe. *Nature* **2014**, *516*, 432–435.

25. Eskeland, R.; Leeb, M.; Grimes, G.R.; Kress, C.; Boyle, S.; Sproul, D.; Gilbert, N.; Fan, Y.; Skoultchi, A.I.; Wutz, A.; *et al.* Ring1b compacts chromatin structure and represses gene expression independent of histone ubiquitination. *Mol. Cell* **2010**, *38*, 452–464.

26. Jin, F.; Li, Y.; Dixon, J.R.; Selvaraj, S.; Ye, Z.; Lee, A.Y.; Yen, C.A.; Schmitt, A.D.; Espinoza, C.A.; Ren, B. A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* **2013**, *503*, 290–294.

27. Naumova, N.; Imakaev, M.; Fudenberg, G.; Zhan, Y.; Lajoie, B.R.; Mirny, L.A.; Dekker, J. Organization of the mitotic chromosome. *Science* **2013**, *342*, 948–953.

28. Sofueva, S.; Yaffe, E.; Chan, W.C.; Georgopoulou, D.; Vietri Rudan, M.; Mira-Bontenbal, H.; Pollard, S.M.; Schroth, G.P.; Tanay, A.; Hadjur, S. Cohesin-mediated interactions organize chromosomal domain architecture. *EMBO J.* **2013**, *32*, 3119–3129.

29. 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, I.W.F.; *et al.* Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 996–1001.

30. Kagey, M.H.; Newman, J.J.; Bilodeau, S.; Zhan, Y.; Orlando, D.A.; van Berkum, N.L.; Ebmeier, C.C.; Goossens, J.; Rahl, P.B.; Levine, S.S.; *et al.* Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **2010**, *467*, 430–435.

31. Liang, J.; Lacroix, L.; Gamot, A.; Cuddapah, S.; Queille, S.; Lhoumaud, P.; Lepetit, P.; Martin, P.G.; Vogelmann, J.; Court, F.; *et al.* Chromatin immunoprecipitation indirect peaks highlight long-range interactions of insulator proteins and pol ii pausing. *Mol. Cell* **2014**, *53*, 672–681.

32. Merkenschlager, M.; Odom, D.T. CTCF and cohesin: Linking gene regulatory elements with their targets. *Cell* **2013**, *152*, 1285–1297.

33. Ong, C.T.; Corces, V.G. CTCF: An architectural protein bridging genome topology and function. *Nat. Rev. Genet.* **2014**, *15*, 234–246.

34. Spencer, R.J.; del Rosario, B.C.; Pinter, S.F.; Lessing, D.; Sadreyev, R.I.; Lee, J.T. A boundary element between tsix and xist binds the chromatin insulator CTCF and contributes to initiation of X-chromosome inactivation. *Genetics* **2011**, *189*, 441–454.

35. Shih, H.Y.; Krangel, M.S. Chromatin architecture, ccctc-binding factor, and V(D)J recombination: Managing long-distance relationships at antigen receptor loci. *J. Immunol.* **2013**, *190*, 4915–4921.

36. Caré, B.R.; Carrivain, P.; Forné, T.; Victor, J.-M.; Lesne, A. Finite-size conformational transitions: A unifying concept underlying chromosome dynamics. *Commun. Theor. Phys.* **2014**, *62*, 607–616.

37. Court, F.; Baniol, M.; Hagège, H.; Petit, J.S.; Lelay-Taha, M.N.; Carbonell, F.; Weber, M.; Cathala, G.; Forné, T. Long-range chromatin interactions at the mouse *Igf2/H19* locus reveal a novel paternally expressed long non-coding RNA. *Nucleic Acids Res.* **2011**, *39*, 5893–5906.

38. Kurukuti, S.; Tiwari, V.K.; Tavoosidana, G.; Pugacheva, E.; Murrell, A.; Zhao, Z.; Lobanenkov, V.; Reik, W.; Ohlsson, R. CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 10684–10689.

39. DeMare, L.E.; Leng, J.; Cotney, J.; Reilly, S.K.; Yin, J.; Sarro, R.; Noonan, J.P. The genomic landscape of cohesin-associated chromatin interactions. *Genome Res.* **2015**, *23*, 1224–1234.

40. Li, Y.; Huang, W.; Niu, L.; Umbach, D.M.; Covo, S.; Li, L. Characterization of constitutive CTCF/cohesin loci: A possible role in establishing topological domains in mammalian genomes. *BMC Genomics* **2013**, doi:10.1186/1471-2164-14-553.

41. Le Gall, A.; Valeri, A.; Nollmann, M. Roles of chromatin insulators in the formation of long-range contacts. *Nucleus* **2015**, *17*, 1–5.

42. Van Bortle, K.; Nichols, M.H.; Li, L.; Ong, C.T.; Takenaka, N.; Qin, Z.S.; Corces, V.G. Insulator function and topological domain border strength scale with architectural protein occupancy. *Genome Biol.* **2014**, doi:10.1186/gb-2014-15-5-r82.

43. 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.; *et al.* Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* **2013**, *153*, 1281–1295.

44. Berlivet, S.; Paquette, D.; Dumouchel, A.; Langlais, D.; Dostie, J.; Kmita, M. Clustering of tissue-specific sub-tads accompanies the regulation of hoxa genes in developing limbs. *PLoS Genet.* **2013**, *9*, e1004018.

45. Guelen, L.; Pagie, L.; Brasset, E.; Meuleman, W.; Faza, M.B.; Talhout, W.; Eussen, B.H.; de Klein, A.; Wessels, L.; de Laat, W.; *et al.* Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **2008**, *453*, 948–951.

46. Peric-Hupkes, D.; Meuleman, W.; Pagie, L.; Bruggeman, S.W.; Solovei, I.; Brugman, W.; Graf, S.; Flicek, P.; Kerkhoven, R.M.; van Lohuizen, M.; *et al.* Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol. Cell* **2010**, *38*, 603–613.

47. Nemeth, A.; Conesa, A.; Santoyo-Lopez, J.; Medina, I.; Montaner, D.; Peterfia, B.; Solovei, I.; Cremer, T.; Dopazo, J.; Langst, G. Initial genomics of the human nucleolus. *PLoS Genet.* **2010**, *6*, e1000889.

48. Vietri Rudan, M.; Barrington, C.; Henderson, S.; Ernst, C.; Odom, D.T.; Tanay, A.; Hadjur, S. Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell Rep.* **2015**, *10*, 1297–1309.

49. Liu, M.; Maurano, M.T.; Wang, H.; Qi, H.; Song, C.Z.; Navas, P.A.; Emery, D.W.; Stamatoyannopoulos, J.A.; Stamatoyannopoulos, G. Genomic discovery of potent chromatin insulators for human gene therapy. *Nat. Biotechnol.* **2015**, *33*, 198–203.

50. Nora, E.P.; Dekker, J.; Heard, E. Segmental folding of chromosomes: A basis for structural and regulatory chromosomal neighborhoods? *Bioessays* **2013**, *35*, 818–828.

51. Mousavi, S.M.J.; Yeganeh, F. Role of chromatin insulators in gene regulation and diseases. *Res. Mol. Med.* **2013**, *1*, 1–5.

52. Salem, T.; Gomard, T.; Court, F.; Moquet-Torcy, G.; Brockly, F.; Forné, T.; Piechaczyk, M. Chromatin loop organization of the junb locus in mouse dendritic cells. *Nucleic Acids Res.* **2013**, *41*, 8908–8925.

53. Andrey, G.; Montavon, T.; Mascrez, B.; Gonzalez, F.; Noordermeer, D.; Leleu, M.; Trono, D.; Spitz, F.; Duboule, D. A switch between topological domains underlies hoxd genes collinearity in mouse limbs. *Science* **2013**, doi:10.1126/science.1234167.

54. Court, F.; Miro, J.; Braem, C.; Lelay-Taha, M.N.; Brisebarre, A.; Atger, F.; Gostan, T.; Weber, M.; Cathala, G.; Forné, T. Modulated contact frequencies at gene-rich loci support a statistical helix model for mammalian chromatin organization. *Genome Biol.* **2011**, doi:10.1186/gb-2011-1112-1185-r1142.

55. Braem, C.; Recolin, B.; Rancourt, R.C.; Angiolini, C.; Barthes, P.; Branchu, P.; Court, F.; Cathala, G.; Ferguson-Smith, A.C.; Forné, T. Genomic matrix attachment region and chromosome conformation capture quantitative real time pcr assays identify novel putative regulatory elements at the imprinted dlk1/gtl2 locus. *J. Biol. Chem.* **2008**, *283*, 18612–18620.

56. Giorgetti, L.; Galupa, R.; Nora, E.P.; Piolot, T.; Lam, F.; Dekker, J.; Tiana, G.; Heard, E. Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. *Cell* **2014**, *157*, 950–963.

57. Kleckner, N. Chiasma formation: Chromatin/axis interplay and the role(s) of the synaptonemal complex. *Chromosoma* **2006**, *115*, 175–194.

58. Phillips-Cremins, J.E. Unraveling architecture of the pluripotent genome. *Curr. Opin. Cell Biol.* **2014**, *28*, 96–104.

59. Kleckner, N.; Zickler, D.; Jones, G.H.; Dekker, J.; Padmore, R.; Henle, J.; Hutchinson, J. A mechanical basis for chromosome function. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 12592–12597.