Published in final edited form as:

Epigenomics. 2015 ; 7(4): 523–526. doi:10.2217/epi.15.25.

Breaking TADs: insights into hierarchical genome organization

Pedro P. Rocha¹, Ramya Raviram¹, Richard Bonneau^{2,3,4}, and Jane A. Skok^{1,*}

¹Department of Pathology, New York University School of Medicine, New York, NY 10016, USA

²Department of Biology, New York University, New York, New York, 10003, USA

³Department of Computer Science, Courant Institute of Mathematical Sciences, New York, New York, 10003, USA

⁴Simons Center for Data Analysis, New York, New York, 10010, USA

The 3D organization of chromosomes enables cells to balance the biophysical constraints of the crowded nucleus with the functional dynamics of gene regulation. Physical contacts between genes and their regulatory elements are essential for proper transcriptional control and maintenance of these interactions is critical for preventing aberrations in physiological processes that could manifest as disease states. The first insights into global nuclear organization came from imaging studies using FISH (fluorescent in-situ hybridization) analyses, which demonstrated that chromosomes occupy individual territories in the nucleus with minimal intermingling between them [1]. The development of chromosome conformation capture (3C) in which chromatin fragments in close physical proximity can be detected enabled the characterization of molecular interactions between different loci [2]. When 3C-based techniques incorporated massive parallel sequencing (such as in Hi-C) the description of molecular chromatin interactions at a genome-wide scale was finally possible [3]. Hi-C was the first unbiased approach aimed at capturing all interactions in the nucleus thereby providing a snapshot of nuclear organization at the global scale. The first Hi-C study revealed that each chromosomal territory is further divided into large domains of 5–10Mb that physically separate two compartments (A and B), which strongly correlate with active and inactive chromatin, respectively [3]. Furthermore, this study demonstrated that interactions between loci in the same compartment occur at a higher frequency than between loci in different compartments [3]. With the progressive decrease in sequencing costs, higher-resolution Hi-C revealed a new level of nuclear organization where compartments A and B can be further divided into “topologically associated domains” (TADs) [4–6]. In mammalian cells these domains range in size from a few 100kbs to 5Mbs in size (with an average of 1MB). Since they exhibit a high degree of conservation between cell types and species it was proposed that TADs represent the fundamental unit of physical organization of the genome [5].

TAD's were defined as regions with distinct boundaries containing loci that interact at a higher frequency with other regions in the same TAD than with regions in different TADs [4–6]. The organization of the genome into these structures contributes to establishing

*To whom correspondence should be addressed: Tel: 212-263-0504, jane.skok@med.nyu.edu.

regulatory interactions (enhancer-promoter), which occur almost exclusively within the same TAD [5, 7, 8]. In fact, changes in transcriptional programs or cellular identity, predominantly lead to changes in interactions occurring between elements in the same TAD without any major alterations in TAD boundaries [9]. Furthermore, genes within the same TAD are more likely to be co-regulated and to respond similarly to transcriptional stimuli [10]. It is not clear however if the TAD structure facilitates this co-regulation or if the shared transcriptional program and epigenetic status of nearby genes underlies the establishment of the TAD structure.

The regions between TADs show an abrupt drop in Hi-C interaction signal. These regions, defined as TAD boundaries, are responsible for limiting the interactions between neighboring TADs thereby leading to a high level of self-interaction within the domains. Although it is not yet fully clear how these boundaries are established it is known that the insulator protein CTCF, the Cohesin complex, housekeeping and tRNA genes as well as some transposable elements are present at a higher than expected frequency and their presence may contribute to the separation of individual TADs [4–6]. Finally, sharp transitions between active and inactive epigenetic signatures are commonly found at TAD boundaries and likely contribute to their formation [4–6]. In-depth analysis of Hi-C data revealed that the strength of these borders as well as their size is not homogenous [11]. While some boundaries with very strong insulating activity show a clear demarcation, others are more permissive for inter-TAD interactions. An attractive hypothesis that could explain these observations is that the strength of each border is a reflection of the degree of enrichment of insulating features and enrichment of boundary elements that are present [11]. It is also possible that some borders are less well defined because of lower resolution Hi-C data in some areas of the genome resulting from technical limitations of the experimental approach.

The seemingly invariant nature of TAD borders suggests a fundamental level of chromosomal organization common to all cell types. However, we know that there are certain features in cells (transcriptional programs, genomic accessibility and epigenetic signatures) that dictate unique cellular identity. The unique identity of individual cell types should therefore also be reflected by small changes in the location of TAD boundaries, which would allow for the formation of new regulatory interactions within adjacent TADs. A good example of such a change occurs during embryonic development within the murine *Hox* clusters [12]. Activation of *Hox* genes occurs in a collinear fashion and two distinct physical chromatin domains are formed separating the active from inactive genes in two separate adjacent TADs. As development progresses, the border of the active domain shifts to incorporate newly activated *Hox* genes. This leads to a change in the location of the border between the two TADs that potentially permits transitioning genes to be exposed to different regulatory elements within the respective TAD [12]. Interestingly, deletion of a single CTCF binding motif that marks the boundary between active and inactive genes of the *Hoxa* and *Hoxc* clusters in differentiated motor neurons leads to an expansion of the active domain and activation of a *Hox* locus that is normally repressed [13]. These examples clearly demonstrate that changes in epigenetic landscape and repositioning of TAD borders

are linked. Nonetheless, the full molecular mechanism underlying this relationship is far from clear.

One of the most surprising features of TADs is the high level of conservation of borders across syntenic regions of the genome in different mammalian species. A recent Hi-C study using liver cells from four different species (mouse, dog, rabbit and macaque) observed that the overall chromosome architecture of these distantly related species at the TAD level is well conserved [14]. Furthermore, conserved CTCF sites are predominantly found at TAD borders while the majority of divergent sites are located within TADs. This is consistent with the finding that TAD borders are highly conserved between species and supports a role for CTCF as an insulator protein that separates adjacent TADs. CTCF is also responsible for promoter-enhancer contacts that tend to be species-specific [15]. Therefore it is expected that these interactions, which typically are represented by intra-TAD associations, will be less conserved.

To study how chromosome architecture changes during cellular differentiation the Ren lab performed Hi-C in ES cells differentiated into four distinct states [9]. Their data demonstrate that nuclear compartments can switch between the active (A) and inactive (B) states and that there is an increase in switching to the B compartment as cells differentiate. Compartments that change tend to coincide with single or multiple adjacent TADs suggesting that although TADs are isolated from their neighbors, in some instances the activity of several adjacent TADs can be altered by the same differentiation cues. In line with previous studies, they find that TAD boundaries are stable between cell types, but changes in the internal structure of the TAD are detected following differentiation such that intra-TAD interactions are increased when there is a change from an inactive to an active domain and reduced when a domain is repressed. These changing interactions provide insight into the way in which the three-dimensional organization of the genome reflects alterations in lineage and stage specific transcriptional programs that govern cell fate. However, further work using high-resolution approaches is required to better characterize intra-TAD interactions and the regulation and functional ramifications of these structures.

Currently all methods for exploring chromosomal structure and structure-function relationships are limited by a trade-off between resolution and the proportion of the genome that can be analyzed. For example use of 5C provides high-resolution information for localized interactions, while Hi-C generates lower resolution data for genome wide associations. A study using 5C allowed for the characterization of TADs at a finer scale and demonstrated that these units can be further partitioned into smaller sub-megabase domains [15]. The structure of sub-TADs in ESCs changes upon differentiation into neural progenitors. Studies such as these raise the question of whether TADs are indeed the fundamental unit of genome organization, or if there are yet smaller structures that will be identified with higher resolution approaches of chromosome conformation capture. Indeed, by increasing the depth of sequencing in Hi-C, genome-wide scale interactions can be identified at 1kb resolution [16]. This study confirmed all previously identified levels of nuclear organization (such as chromosome territories, A / B compartments, and TADs) and additionally showed that small 100kb domains can be found throughout the genome. These findings demonstrate that, as expected, the size of the smallest identifiable chromosomal

domain depends on the sequencing depth and the resolution of the 3C-based approach. It is also important to remember that we do not fully understand how these computationally defined domains (that are derived from a convolution of contacts from a large population of cells) translate into actual physical chromatin structures in single cells and whether intra-TAD interactions are representative of stable structures as opposed to dynamic interactions [17]. Moreover, it is clear from Hi-C data that there are hierarchical interaction-based structures occurring beyond the level of TADs suggesting the presence of important long-range interactions undetectable by current Hi-C methods [18]. Physical contacts measured by Hi-C seem to be limited to small linear distances while failing to detect interactions with other chromosomes that are reproducibly captured by other techniques such as FISH and 4C-Seq [19, 20], suggesting hybrid experimental designs are required to connect chromosome structure with regulatory function. Finally, the recent NIH 4D-Nucleome RFA that focuses on tool development with the aim of obtaining a high resolution map of interactions across the genome as a function of development/disease progression is well timed to support efforts to generate new approaches that tackle the issues raised in this article.

Acknowledgments

PR is supported by an ASH fellowship. RB is supported by the Simons Foundation. JAS is a Leukemia & Lymphoma Society (LLS) scholar and is supported by NIH grants R01 GM086852, NIAID R56 A1099111 and R01GM112192.

References

1. Cremer T, Cremer C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nature reviews Genetics*. 2001; 2(4):292–301.
2. Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. *Science*. 2002; 295(5558):1306–1311. [PubMed: 11847345]
3. Lieberman-Aiden E, Van Berkum NL, Williams L, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*. 2009; 326(5950):289–293. [PubMed: 19815776]
4. Nora EP, Lajoie BR, Schulz EG, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*. 2012; 485(7398):381–385. [PubMed: 22495304]
5. Dixon JR, Selvaraj S, Yue F, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*. 2012; 485(7398):376–380. [PubMed: 22495300]
6. Sexton T, Yaffe E, Kenigsberg E, et al. Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell*. 2012; 148(3):458–472. [PubMed: 22265598]
7. Schoenfelder S, Furlan-Magaril M, Mifsud B, et al. The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements. *Genome research*. 2015
8. Downen JM, Fan ZP, Hnisz D, et al. Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell*. 2014; 159(2):374–387. [PubMed: 25303531]
9. Dixon JR, Jung I, Selvaraj S, et al. Chromatin architecture reorganization during stem cell differentiation. *Nature*. 2015; 518(7539):331–336. [PubMed: 25693564]
10. Le Dily F, Bau D, Pohl A, et al. Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes & development*. 2014; 28(19): 2151–2162. [PubMed: 25274727]
11. Van Bortle K, Nichols MH, Li L, et al. Insulator function and topological domain border strength scale with architectural protein occupancy. *Genome biology*. 2014; 15(6):R82. [PubMed: 24981874]

12. Andrey G, Montavon T, Mascres B, et al. A switch between topological domains underlies HoxD genes collinearity in mouse limbs. *Science*. 2013; 340(6137):1234-1237. [PubMed: 23744951]
13. Narendra V, Rocha PP, An D, et al. Transcription. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science*. 2015; 347(6225):1017–1021. [PubMed: 25722416]
14. Vietri Rudan M, Barrington C, Henderson S, et al. Comparative Hi-C Reveals that CTCF Underlies Evolution of Chromosomal Domain Architecture. *Cell reports*. 2015; 10(8):1297–1309. [PubMed: 25732821]
15. Phillips-Cremins JE, Sauria ME, Sanyal A, et al. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell*. 2013; 153(6):1281–1295. [PubMed: 23706625]
16. Rao SS, Huntley MH, Durand NC, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*. 2014; 159(7):1665–1680. [PubMed: 25497547]
17. Giorgetti L, Galupa R, Nora EP, et al. Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. *Cell*. 2014; 157(4):950–963. [PubMed: 24813616]
18. Lajoie BR, Dekker J, Kaplan N. The Hitchhiker's guide to Hi-C analysis: practical guidelines. *Methods*. 2015; 72:65–75. [PubMed: 25448293]
19. Noordermeer D, De Wit E, Klous P, et al. Variegated gene expression caused by cell-specific long-range DNA interactions. *Nature cell biology*. 2011; 13(8):944–951. [PubMed: 21706023]
20. Chaumeil J, Micsinai M, Ntziachristos P, et al. The RAG2 C-terminus and ATM protect genome integrity by controlling antigen receptor gene cleavage. *Nature communications*. 2013; 4:2231.