



HAL
open science

Of Dots and Stripes: The Morse Code of Micro-C Reveals the Ultrastructure of Transcriptional and Architectural Mammalian 3D Genome Organization

Li-Hsin Chang, Daan Noordermeer

► **To cite this version:**

Li-Hsin Chang, Daan Noordermeer. Of Dots and Stripes: The Morse Code of Micro-C Reveals the Ultrastructure of Transcriptional and Architectural Mammalian 3D Genome Organization. *Molecular Cell*, 2020, 78 (3), pp.376–378. 10.1016/j.molcel.2020.04.021 . hal-02995878

HAL Id: hal-02995878

<https://hal.science/hal-02995878v1>

Submitted on 9 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Of Dots and Stripes: The Morse Code of Micro-C Reveals the Ultrastructure of Transcriptional and Architectural Mammalian 3D Genome Organization

Li-Hsin Chang¹ and Daan Noordermeer^{1,*}

¹ Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), Gif- sur-Yvette, France

* Correspondence: daan.noordermeer@i2bc.paris-saclay.fr

Abstract (*40 words*)

Using Micro-C, Hsieh et al. (2020) and Krietenstein et al. (2020) investigated 3D chromatin folding in human and mouse cells at unprecedented resolution, to uncover ultra-fine-scale chromatin structures that provide direct links between genome architecture, gene expression and gene regulation.

Main text (1056 words)

Hi-C, a genome-wide 3C (Chromosome Conformation Capture) assay, has been instrumental to characterize global 3D chromosome organization in mammalian cells (Lieberman-Aiden et al. 2009). Stepwise improvements to the resolution of Hi-C, by enhancing the protocol and by increasing sequencing depth, have consistently revealed new layers of genome structure. The first Hi-C study uncovered multi-Megabase (Mb) chromosome compartments, which separate the genome according to global transcriptional activity (Lieberman-Aiden et al. 2009). Next, sub-Mb Topologically Associating Domains (TADs) were identified (Dixon et al. 2012). TADs are shaped by a process of Cohesin-mediated ‘loop extrusion’ which is paused by the CTCF insulator protein (Fudenberg et al. 2016). The resulting insulated domains, with about two-fold enrichment of intra-domain interactions, function as regulatory neighborhoods by restricting enhancer-promoter loops (see e.g. Chang et al., 2020). Subsequent studies revealed hundred-kb-scale contact domains, associated with histone modifications, and several thousands of DNA loops linking various genomic elements (Bonev et al. 2017; Rao et al. 2014). The resolution of Hi-C, like other conventional 3C assays, is limited by the use of restriction enzymes for DNA fragmentation, which generates fragments of irregular size and independent of other aspects of chromatin structure (Figure 1A). Micro-C and its further development Micro-C XL—both derivatives of the Hi-C assay—overcome these limitations of fragmentation by using micrococcal nuclease (MNase) digestion to generate mononucleosomes (Hsieh et al. 2016). In this issue of *Molecular Cell*, two articles present the first ultrastructural Micro-C studies of human and mouse chromatin architecture, which constitute the currently highest-resolution genome-wide investigations of mammalian 3D chromosome organization (Hsieh et al., 2020; Krietenstein et al. 2020).

By generating Micro-C contact matrices in human foreskin fibroblasts, as well as in human and mouse embryonic stem cells, both studies recapitulated the known features of mammalian genome organization in Hi-C maps: chromosome compartments, TADs and DNA loops. More importantly, the improved signal-to-noise ratio of Micro-C provided access to new levels of chromatin organization: an increased ability to identify chromatin loops, boundaries and self-interacting domains, and the measurement of fine-scale structure at ultra-short distance (200 bp – 20 kb) (Figure 1A). As a result, this allowed the identification of subtle and dynamic chromatin structures that appear functionally related to gene activity and genome architecture. Both studies

identified thousands of boundaries that delimit small (5 - 10 kb) and often nested self-interacting domains that remained undetected in previous Hi-C studies due to their relatively moderate insulation (Figure 1B). These boundaries can be separated into multiple clusters based on the presence of TAD-associated CTCF and Cohesin (Dixon et al., 2012), the overlap with RNA polymerase II (Pol II), the H3K4me3 histone modification and the YY1-factor that are all associated with promoter activity, and the binding of both global and specific transcription factors that suggest links with enhancers. Large numbers of boundaries serve as anchors for so-called ‘stripes’, which were previously linked to (asymmetric) loop extrusion at TADs (Chang et al., 2020; Fudenberg et al., 2016). Moreover, along these stripes, tens of thousands of ‘dots’ are visible, thereby revealing the widespread presence of chromatin loops in all three cell types (Figure 1B).

The analysis of Hsieh and colleagues focused on promoter-promoter and enhancer-promoter stripes (P-P and E-P stripes). Compared to stripes found at TAD boundaries, P-P and E-P stripes are smaller (10 - 50 kb), localize at active promoters and enhancers, and generally do not overlap with CTCF (Figure 1B). This intriguing observation suggests that active promoters can maintain contacts with the gene body during transcription and beyond. Large numbers of nested dots are visible where P-P and E-P stripes intersect, suggesting that collisions between proteins bound to these elements result in stabilized interactions. Short-term transcriptional inhibition experiments support these observations, as P-P stripes and (to a lesser extent) E-P stripes are reduced, whereas both P-P and E-P dots appear more stable.

Krietenstein and colleagues instead mostly concentrated on CTCF-associated stripes and dots (Figure 1B). Stripes do not only emanate from the convergent CTCF-bound sites at the boundaries of TADs, but large numbers are detected from CTCF sites within TADs as well. Dots are visible where these stripes collide, albeit often with relatively weak signal enrichment. These interactions therefore appear to constitute relatively weak ‘pause sites’ for loop extrusion, which provides further support to the notion that TADs and their boundaries are of a dynamic nature (see e.g. Chang et al., 2020).

Both studies used Micro-C to quantitatively determine the fine-scale chromatin structure of nucleosomal interactions (Figure 1C). In both human and mouse cells, interaction frequencies show a distinctive pattern whereby decay is not gradual but rather by pairs of adjacent nucleosomes. Such a pattern is compatible with previous *in-vitro* Cryo-EM results that identified

the organization into small ‘clutches’, where tetranucleosomes are stacked into a typical two-start zig-zag motif (Song et al., 2014).

The unmatched detail of these Micro-C studies raises a number of important questions. The identification of prevalent P-P and E-P stripes brings up the issue if Pol II actively establishes interactions between promoters and enhancers, and if this involves a process similar to the loop extrusion that shapes TADs (Fudenberg et al., 2016). Hsieh and colleagues provide several hypotheses, including the reeling-in of promoters and enhancers into “regulatory hubs”, a potential association of Pol II with the Cohesin machinery, or the physical association of promoters and enhancers to transcribing Pol II. Additional investigation will be required to make this distinction or, alternatively, if the increased sensitivity of Micro-C has unearthed other biological or experimental characteristics associated with loop anchors. Similarly interesting is the identification of thousands of self-interacting domains with links to loop extrusion. The demarcation of these domains by dots and stripes indicates that they are formed like TADs, albeit at a smaller scale and with relatively weak signal enrichment (Fudenberg et al., 2016). Their identification suggests that loop extrusion slows down at weak pause sites within TADs before arresting at the strong loop anchors at TAD boundaries (Figure 1B), which supports previous genomics, imaging and polymer physics modeling studies that suggest that TADs are dynamic structures (see Chang et al., 2020). Future Micro-C measurements after acute depletion of CTCF or Cohesin will be instrumental for a truly quantitative understanding of loop extrusion kinetics and its interference with the transcription machinery. Both these studies confirmed that the resolution and sensitivity of Micro-C allow the characterization of such molecular effects with unprecedented detail.

References

- Bonev, B., Mendelson Cohen, N., Szabo, Q., Fritsch, L., Papadopoulos, G.L., Lubling, Y., Xu, X., Lv, X., Hugnot, J.P., Tanay, A., et al. (2017). Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell* 171, 557–572.
- Chang, L.-H., Ghosh, S., and Noordermeer, D. (2020). TADs and Their Borders: Free Movement or Building a Wall? *J. Mol. Biol.* 432, 643-652.
- 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.
- Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., and Mirny, L.A. (2016). Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep.* 15, 2038–2049.
- Hsieh, T.S., Fudenberg, G., Goloborodko, A., and Rando, O.J. (2016). Micro-C XL: assaying chromosome conformation from the nucleosome to the entire genome. *Nat. Methods* 13, 1009–1011.
- Hsieh, T.H.S., Cattoglio, C., Slobodyanyuk, E., Hansen, A.S., Rando, O.J., Tjian, R., and Darzacq, X. (2020). Resolving the 3D Landscape of Transcription-Linked Mammalian Chromatin Folding. *Molecular Cell*, this issue.
- Krietenstein, N., Abraham, S., Venev, S.V., Abdennur, N., Gibcus, J., Hsieh, T.H.S., Mohan Parsi, K., Yang, L., Maehr, R., Mirny, L.A., et al. (2020). Ultrastructural Details of Mammalian Chromosome Architecture. *Molecular Cell*, this issue.
- 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. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289–293.
- 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., and Aiden, E.L. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665–1680.

Song, F., Chen, P., Sun, D., Wang, M., Dong, L., Liang, D., Xu, R.-M., Zhu, P., and Li, G. (2014). Cryo-EM Study of the Chromatin Fiber Reveals a Double Helix Twisted by Tetranucleosomal Units. *Science* 344, 376–380.

Figure legend

Figure 1. Micro-C uncovers the ultrastructure of 3D chromatin organization and nucleosomal interactions.

- (A) Essential differences between Micro-C and Hi-C. Top: schematic electrophoresis image with size distributions of Micro-C and Hi-C material after fragmentation. In Micro-C, the majority of fragments are mononucleosomes (up to 90%, red arrowhead), whereas in Hi-C a wide range of fragments is generated (schematic pattern for the widely used MboI restriction enzyme). Bottom: schematic scaling of contact frequency $P(s)$ relative to the genomic distance between read pairs in Micro-C and Hi-C experiments. In the range of 1kb to 20 kb, Micro-C signal maintains its quantitative scaling, thus allowing the measurement of nucleosomal interactions (left red shading). At larger distances, Micro-C has an improved signal-to-noise ratio, as shown by the extended range of linear decay at genomic separations of over than 10 Mb (right red shading).
- (B) Cartoon of a Micro-C contact matrix illustrating chromatin ultrastructure within a TAD. Left: stripes emanate from both enhancers (E) and promoters (P), suggesting they act as anchors in a process that resembles one-sided loop extrusion. At the intersection of convergent stripes, the stabilized interactions appear as E-P and P-P dots. Genomic regions encapsulated within convergent stripes display moderately increased interactions that appear as E-P and P-P domains. Right: loop extrusion stripes appear from CTCF sites located within TADs, with dots appearing where convergent stripes collide. Compared to the dots at the far end of TADs, which represent strong extrusion pause sites, these intra-TAD CTCF dots have lower signal enrichment that suggests they act as weaker pause sites. Below the TAD, the schematic position of an enhancer, genes, and CTCF binding sites with their orientations are indicated.
- (C) Micro-C confirms a previously *in-vitro* identified two-start zig-zag model for tetranucleosome organization in mammalian cells (Song et al., 2014). As depicted in this schematic of the zig-zag model, a nucleosome (N) interacts with the two downstream N+2 and N+3 nucleosomes at roughly the same frequency.