RESEARCH HIGHLIGHT
3D genome encoded by LINE and SINE repeats

Zhengyu Liang[1] and Xiang-Dong Fu [1]

Retrotransposons are embedded in distinct genomic regions, many giving rise to repeat-enriched transcripts. In a recent paper of *Cell Research*, Lu et al. show that L1 and B1 retrotransposons are separately clustered in mammalian genomes where L1 transcription and transcripts play a vital role in orchestrating the formation of 3-dimensional genome.

Eukaryotic genome is hierarchically packaged in the nucleus as multiscale structural units, allowing gene transcription and DNA replication to take place in a spatially and temporally regulated fashion.[1] High-throughput mapping of DNA–DNA interactions reveals that the human genome is segregated into two compartment domains, named A and B, with distinct transcriptional activities, histone modifications and nuclear positioning.[2] Underlying the plaid pattern of A/B compartments are locally interacting sequences, corresponding to topologically associating domains (TADs), that are also engaged in preferential homotypic (i.e., A–A and B–B) long-distance interactions across the entire chromosome (Fig. 1).

Past research has been focused on the principle for the formation of these genomic structures. For example, inhibition of Pol II transcription or depletion of RNA has been found to weaken TAD bordering and reduce B-type compartmental interactions, although the majority of homotypic A–A and B–B interactions are largely preserved.[3] Therefore, while the genome is dynamically regulated, its basic organization units appear to be quite stable, raising the question of whether the genome architecture is encoded in its primary DNA sequences.

In a recent paper of *Cell Research*,[4] Xiaohua Shen and colleagues systematically addressed this fundamental question, finding that retrotransposons embedded in mammalian genomes, particularly Alu/B1 and L1 elements, are preferentially enriched in euchromatic A compartments and heterochromatic B compartments, respectively. In line with specific chromatin markers associated with these elements, short and long interspersed nuclear elements (SINEs and LINEs) are differentially distributed in the 3-dimensional (3D) space of the nucleus, with SINEs (particularly Alu in humans and B1 in mice) localized in the nuclear interior while LINEs (particularly L1 in both humans and mice) around the periphery of nucleus and nucleolus. This pattern is conserved across cell types in mammals, reconstructed during the cell cycle, and de novo established in early embryogenesis, indicating that these repeat-enriched elements play important roles in instructing genomic folding through homotypic clustering. An unsettled question is whether such homotypic interactions are mediated by DNA–protein interactions, RNA–protein interactions or both.

A subset of SINEs and LINEs are actively transcribing in mammalian genomes. Given the increasing evidence for RNA-facilitated genomic interactions at chromatin levels,[5] the role of L1-derived RNAs was investigated in the current study during cell cycle and early embryogenesis when chromatin undergoes dynamic reorganization. Blocking L1 transcripts with antisense morpholino (AMO) in mouse embryos clearly demonstrated the functional requirement of L1 transcripts in establishing homotypic clustering to segregate A and B compartments and to promote their differential localization in the nucleus. Although not all repeat-rich sequences in the genome are transcribing, a recent study showed that these RNAs are able to target related sequences both in cis and in trans[6] and resulting chromatin features, such as H3K9me3 and HP1α binding, may then drive homotypic interactions.

Compared to L1 transcripts, the functional contribution of SINE RNA to genome partitioning remains less clear. Technically, depletion of B1-derived transcripts leads to severe cell death, thus precluding functional evaluation of these RNAs in euchromatin formation. Theoretically, various coding and non-coding RNAs, including SINE transcripts, may work together to promote the formation of various transcriptional hubs as part of A compartments. In this regard, it is worth pointing out that SINE transcripts are able to directly bind Pol II, a subset of which (Alu, but not scAlu in humans, B2, but not B1 in mice) may even suppress Pol II transcription under heat shock conditions.[7] This may alter the status of euchromatin, especially when SINE transcripts are overexpressed. Given that various retrotransposons are induced first during zygotic genome activation,[8] it is conceivable that the formation of B compartments may take place ahead of A compartments, the latter of which may take shape in a semi self-organized fashion after the general framework of heterochromatic territories is established.[9]

Last, but not least, active retrotransposons embedded in mammalian genomes likely contribute to the organization of 3D genome via the formation of RNA-mediated, phase-separated droplets. This concept in fact applies to both heterochromatin and euchromatin to form distinct intranuclear compartments.[10] The functional distinction is that repeat-derived RNAs initially facilitate the formation of heterochromatin, which then becomes largely silent in transcription, whereas diverse coding and non-coding RNAs help nucleate the formation of euchromatin to drive and maintain active transcription. In both of these processes, nascent RNAs are not only products but also key architectural components of the 3D genome.

[1]Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, CA 92093-0651, USA
Correspondence: Xiang-Dong Fu (xdfu@ucsd.edu)
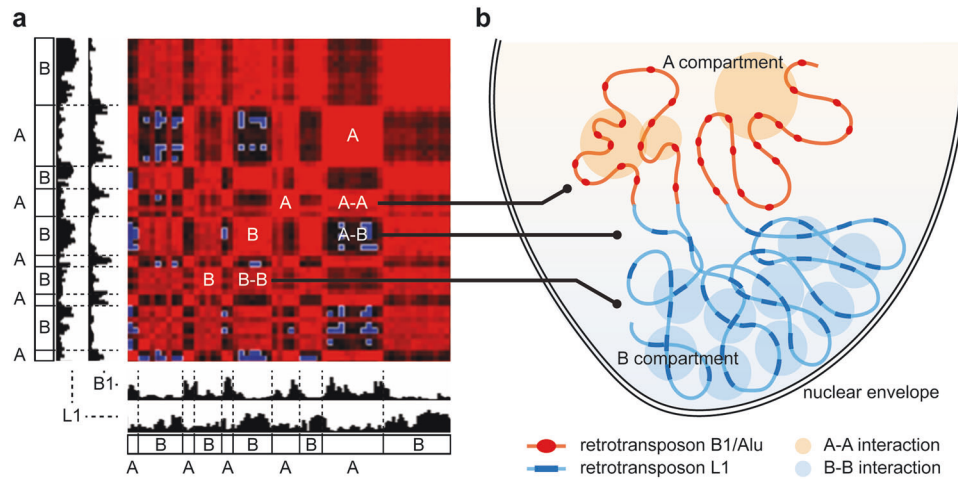
Fig. 1 L1 and B1 retrotransposons instruct genome folding. a Differential distribution of Alu/B1 and L1 elements in the mouse genome, with preferential localization of Alu/B1 in A compartments and L1 in B compartments in the Hi-C DNA–DNA map (derived from Fig. 1f in[4]). **b** Illustrated is homotypic clustering of these repeat-rich elements, resulting in the plaid pattern of frequent A–A and B–B interactions (red), but relatively rare A–B or B–A interactions (blue).

REFERENCES

1. Rowley, M. J. & Corces, V. G. *Nat. Rev. Genet.* **19**, 789–800 (2018).
2. Lieberman-Aiden, E. et al. *Science* **326**, 289–293 (2009).
3. Barutcu, A. R., Blencowe, B. J. & Rinn, J. L. *EMBO Rep.* **20**, e48068 (2019).
4. Lu, J. Y. et al. *Cell Res*. https://doi.org/10.1038/s41422-020-00466-6 (2021).
5. Li, X. & Fu, X. D. *Nat. Rev. Genet.* **20**, 503–519 (2019).
6. Hao, Y. et al. *Genome Res.* **30**, 1570–1582 (2020).
7. Mariner, P. D. et al. *Mol. Cell* **29**, 499–509 (2008).
8. Hendrickson, P. G. et al. *Nat. Genet.* **49**, 925–934 (2017).
9. Falk, M. et al. *Nature* **570**, 395–399 (2019).
10. Henninger, J. E. et al. *Cell* **184**, 207–225 (2021).