

Evolutionary analysis of 3-dimensional chromatin structure

M. Nuriddinov^{1*}, V. Fishman^{1,2}

¹*Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*

²*Novosibirsk State University, Novosibirsk, Russia*

* e-mail: nuriddinov@bionet.nsc.ru

Key words: Hi-C, chromatin architecture, spatial chromosome organization, genome evolution

Motivation and Aim: Development of the genome-wide chromatin conformation capture techniques such as the Hi-C method allowed discovering of the new layer of spatial chromatin organization, formed by the closely interacting chromosomal regions, in other words, topologically associated domains (TADs). The TAD architecture appears to play important roles in gene regulation, transcription and replication [1]. Furthermore, chromatin domain structure reveals high conservation across both cell types within one organism and evolutionary lineages [1, 2]. Nevertheless, there have been a number of Hi-C data of various species, this are not used to investigate evolutionary changes of chromatin architecture in details. To bridge the gap, we have developed bioinformatic tool allowing cross evolution comparison of chromatin structure.

Methods and Algorithms: We developed a python-based software that uses standard output from HiC-Pro pipeline [3] to compute frequencies of 3-dimentional chromatin contacts in each specie. The software estimates variance of contact frequencies by binomial distribution and normalizes contact counts by vanilla coverage. To compare contact frequencies between species, our algorithm solves several problems:

- finding orthologues contacts of studied species;
- accounting for species-to-species changes of linear distance between contacting loci;
- computing expected contact frequency for one specie based on observed contact frequency for the other specie;
- estimating significance of difference or similarity between observed and expected contact frequencies.

Results: We developed the software that compares Hi-C data of two species and discriminating differential and conserved chromatin interactions. The software can visualize obtained results in a heatmap-like format. As a proof-of-principle, we compared chromatin architecture in mouse and human and showed conservation of spatial contacts around conserved TAD boundaries.

Conclusion: Described bioinformatic tool allows evolutionary comparison of 3-dimentional chromatin structure. This gives new approaches to process chromosome conformation data, exploration of evolution hot spots and identification of regulatory elements. This could give fresh insights to genome evolution and molecular mechanisms forming chromosomal architecture.

Acknowledgements: This study was supported by the RSF (grant 17-74-10143)

References

1. Dixon J.R. et al. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*. 485:376-380.
2. Battulin N. et al. (2015) Comparison of the 3D organization of sperm and fibroblast genomes using the Hi-C approach. *Genome Biol*. 16:77.
3. Servant N. et al. (2015) HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol*. 16:259.