Scale-dependent landscape of semi-nested community structures of 3D chromosome contact networks

Dolores Bernenko,¹ Sang Hoon Lee,^{2,*} Per Stenberg,³ and Ludvig Lizana¹

¹Department of Physics, Integrated Science Lab, Umeå University, SE-901 87 Umeå, Sweden ²Department of Physics and Research Institute of Natural Science, Gyeongsang National University, Jinju 52828, Korea ³EMG, Umeå University, SE-901 87 Umeå, Sweden

Mammalian DNA folds into 3D structures that facilitate and regulate genetic processes such as transcription, DNA repair, and epigenetics. Several insights derive from chromosome capture methods, such as Hi-C, which allow researchers to construct contact maps depicting 3D interactions among all DNA segment pairs. To better understand the organizing principles, several groups analyzed Hi-C data assuming a Russian-doll-like nested hierarchy where DNA regions of similar sizes merge into larger and larger structures. However, while successful, this model is incompatible with the two competing mechanisms that seem to shape a significant part of the chromosomes' 3D organization: loop extrusion and phase separation.

The first part of our work [1] aims to map out the chromosome's actual folding hierarchy from empirical data, by treating the measured DNA-DNA interactions by Hi-C as a weighted network. From such a network, we extract 3D communities using the generalized Louvain algorithm with an adjustable resolution parameter, which allows us to scan seamlessly through the community size spectrum, from A/B compartments to topologically associated domains (TADs) [2]. By constructing a hierarchical tree connecting these communities (as illustrated in Fig. 1), we find that chromosomes are more complex than a perfect hierarchy. Analyzing how communities nest relative to a simple folding model, we find that chromosomes exhibit a significant portion of nested and non-nested community pairs alongside considerable randomness. In addition, by examining nesting and chromatin types, we discover that nested parts are often associated with actively transcribed chromatin.

Another reoccurring issue that seems to reflect the fundamental limitation of community detection in the case of stochastic algorithms is the possibility of inconsistent detection results (the same community detection method may disagree with itself) [3, 4]. If too strong, such inconsistencies may cause problems if the data interpretation relies too heavily on a specific community structure when there are others equally feasible. This is a fundamental problem pertaining to any data clustering scheme that cannot be solved using better community detection algorithms. In the second part of our work, we investigate the inconsistency of 3D communities in Hi-C data. We utilize an inconsistency metric [4], map out the community spectrum at different scales of the Hi-C contact network, and quantify where the community separation is most inconsistent. As a result, we find that the nodal inconsistency or functional flexibility [3] are also related to the local chromatin activity as in the nestedness analysis.

- D. Bernenko, <u>SHL</u>, P. Stenberg, and L. Lizana, Mapping the semi-nested community structure of 3D chromosome contact networks, bioRxiv 10.1101/2022.06.24.497560 (2022), https://www.biorxiv.org/content/early/2022/06/28/2022.06.24.497560.full.pdf.
- [2] <u>SHL</u>, Y. Kim, S. Lee, X. Durang, P. Stenberg, J.-H. Jeon, and L. Lizana, Mapping the spectrum of 3D communities in human chromosome conformation capture data, Scientific Reports 9, 6859 (2019).
- [3] H. Kim and <u>SHL</u>, Relational flexibility of network elements based on inconsistent community detection, Phys. Rev. E 100, 022311 (2019).
- [4] D. Lee, SHL, B. J. Kim, and H. Kim, Consistency landscape of network communities, Phys. Rev. E 103, 052306 (2021).



FIG. 1. Illustration of the semi-nested community structure of the Hi-C chromsome contact network.

^{*} Presenter: lshlj82@gnu.ac.kr