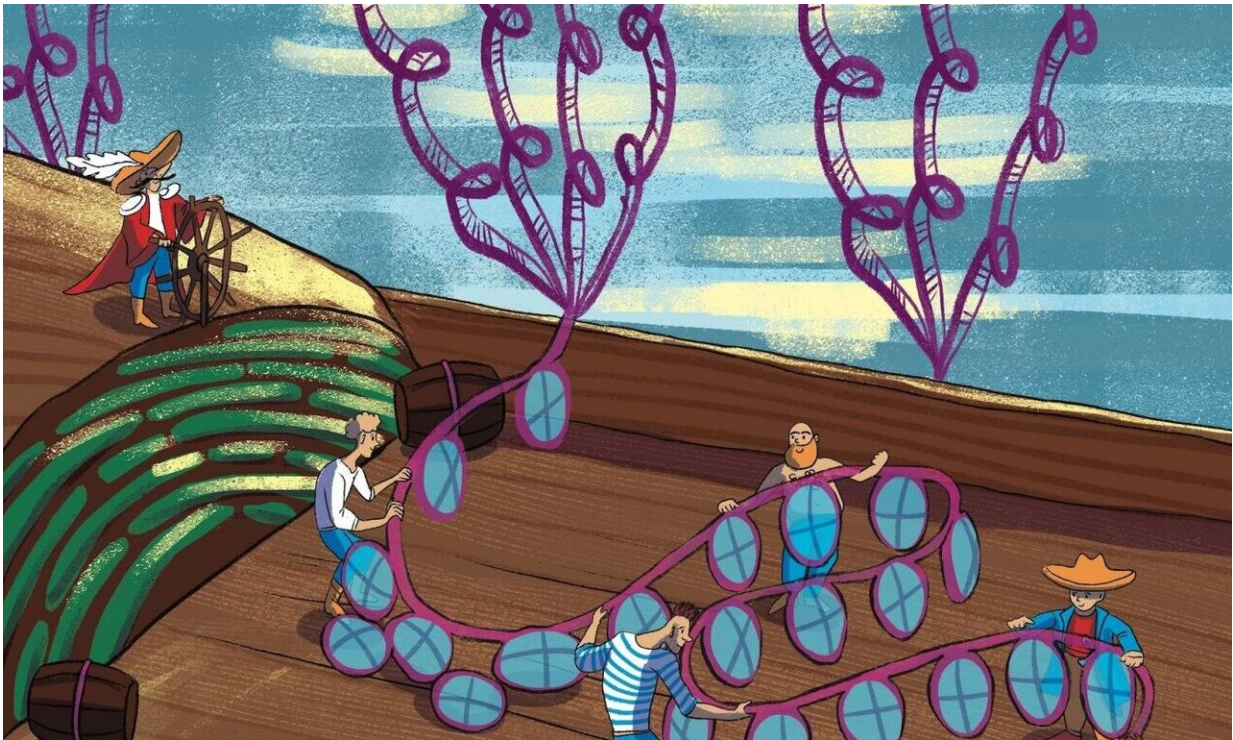


Exploring a genome's 3-D organization through a social network lens

February 20 2020



Much as groups of sailors work together with ropes to accomplish tasks aboard ship, collections of transcription factor proteins work with sections of chromosomes as a community in the nucleus of a cell to carry out cell functions. Computational biologists at Carnegie Mellon University have developed algorithms for identifying these communities in cell nuclei. Credit: Ella Marushchenko

Computational biologists at Carnegie Mellon University have taken an algorithm used to study social networks, such as Facebook communities, and adapted it to identify how DNA and proteins are interconnected into communities within the cell nucleus.

Jian Ma, associate professor in CMU's Computational Biology Department, said scientists have come to appreciate that DNA, proteins and other components within the nucleus appear to form structurally and functionally important communities. The behavior of these communities may prove key to understanding basic cellular processes and disease mechanisms, such as aging and cancer development.

Figuring out how to identify these communities among the tens of thousands of genes, proteins and other components of the cell is daunting, however. An important factor is proximity—both in terms of genes being controlled by the same regulatory proteins called [transcription factors](#) and in terms of spatial arrangement, with the complex folding and packing of DNA putting certain genes close to each other.

In many cases, the relationships are similar to many Facebook communities, with some members located near each other, while others who may be far apart are nevertheless drawn together through shared interests.

In a paper featured on the cover of the February issue of the journal *Genome Research*, lead authors Dechao Tian, a post-doctoral researcher, and Ruochi Zhang, a Ph.D. student in computational biology, explain how they developed a [new algorithm](#), MOCHI, to subdivide the interwoven nuclear components into communities.

MOCHI was inspired by an [algorithm](#) originally developed by the laboratory of computer scientist Jure Leskovec. Beginning as a Ph.D.

student at CMU and continuing as a faculty member at Stanford University, Leskovec has specialized in the analysis of large social and information networks.

The MOCHI algorithm looks at the spatial arrangement of all the genes and transcription factor proteins in a nucleus based on genome-wide chromosome interactions and global gene regulatory networks. Viewing this information as a 3-D graph, the algorithm looks for certain subgraphs or "motifs," within it. A motif might be, say, a triangular shape, as is typical in social network analysis, or a four-node subgraph, which MOCHI uses for analyzing complex networks in the cell nucleus. The algorithm then clusters, or subdivides, the graph in a way that minimizes disruption of these motifs.

They tested MOCHI by applying it to five different cell types. Just as the original algorithm has proved adept at identifying communities within a large mass of social network data, MOCHI identified what appear to be hundreds of communities within the nuclei of these cell types.

As of yet, the researchers don't know what each community might do, but they say they have reason to believe the subdivisions made by MOCHI are valid. For instance, Ma said that the algorithm identified communities that seem to be common to all of the cell types used in this study. It also identified some communities that appear to be unique to a particular cell type. In addition, Ma said they found "enrichment" of disease related genes within the communities.

Much more work will be necessary to identify the function and behavior of each of these communities, Ma said, but the MOCHI algorithm gives researchers a starting point for study.

"There's a reason why these communities are formed in the nucleus," he said. "We just don't know the formation mechanisms of these

communities yet." Understanding them might help researchers delineate fundamental cellular processes and suggest possible ways to better understand disease development.

The researchers also plan to include additional [cell nucleus](#) components, such as RNAs and other types of proteins, into their analysis.

In addition to Ma, Tian and Zhang, authors of the paper include Yang Zhang and Xiaopeng Zhu, a research associate and a project scientist, respectively, in the Computational Biology Department. The National Institutes of Health, including its 4-D Nucleome Program, and the National Science Foundation supported this research.

More information: Dechao Tian et al, MOCHI enables discovery of heterogeneous interactome modules in 3D nucleome, *Genome Research* (2020). [DOI: 10.1101/gr.250316.119](https://doi.org/10.1101/gr.250316.119)

Provided by Carnegie Mellon University

Citation: Exploring a genome's 3-D organization through a social network lens (2020, February 20) retrieved 18 April 2024 from <https://phys.org/news/2020-02-exploring-genome-d-social-network.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--