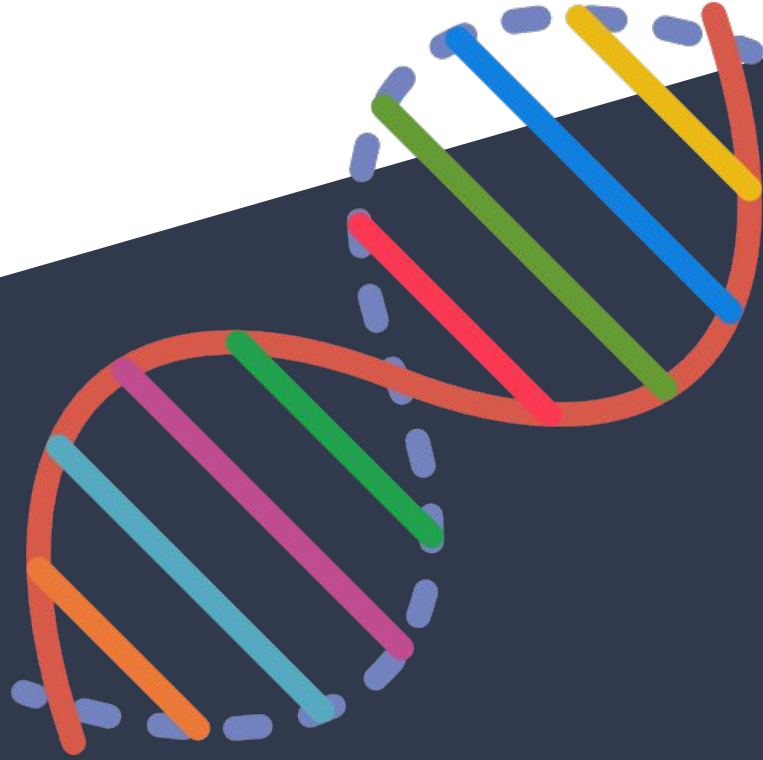
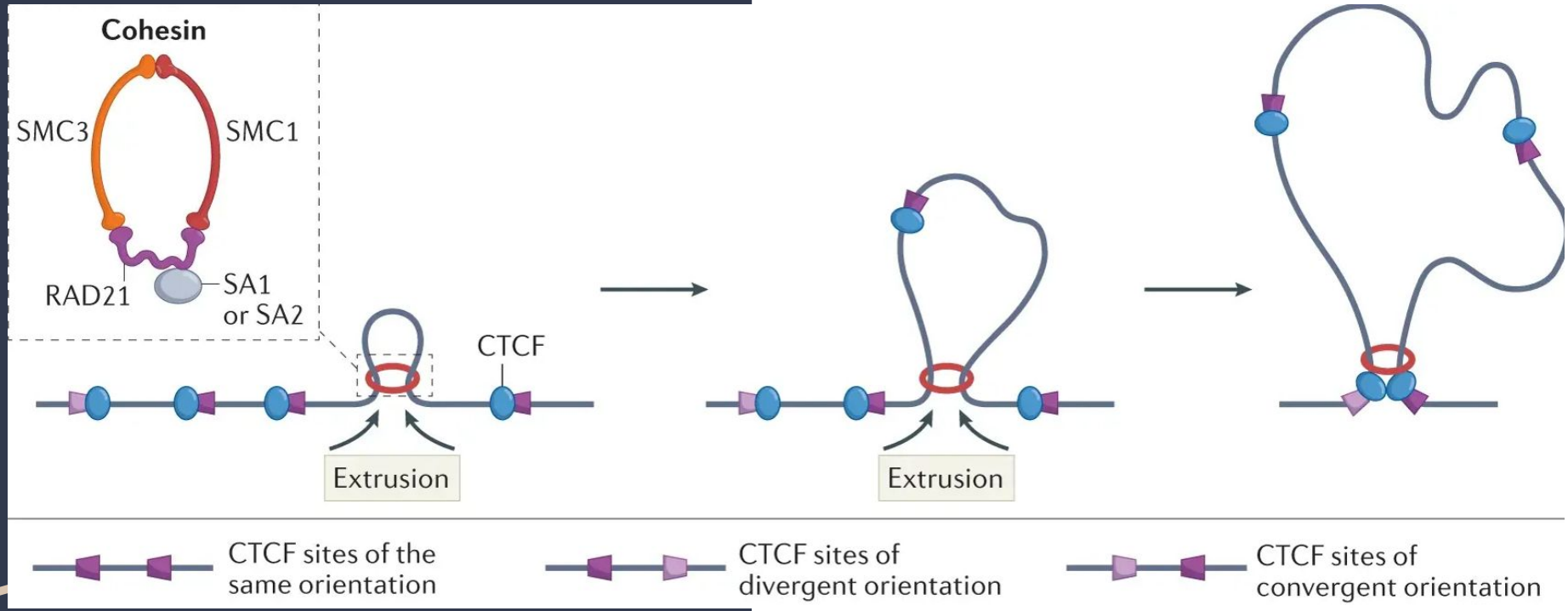
Utilizing Machine Learning to Identify Time Asymmetry of DNA Loop Extrusion

Anna Du



Loop Extrusion

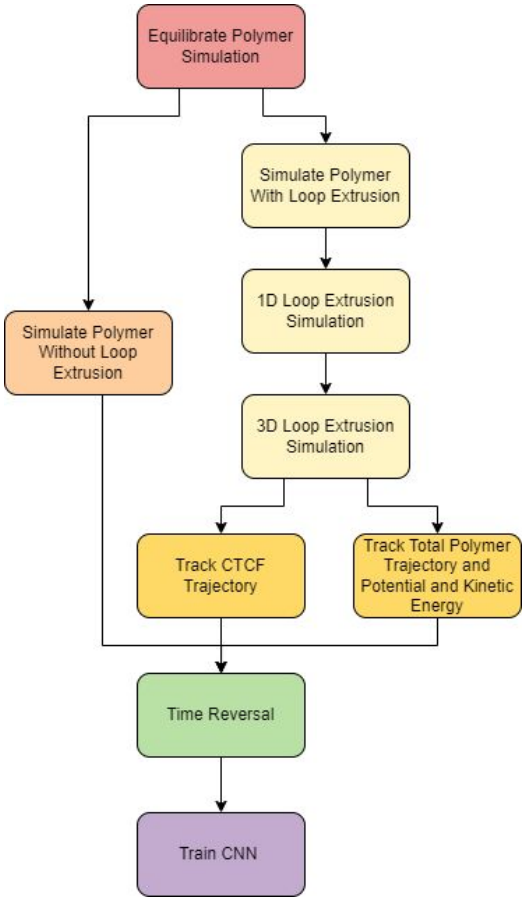


(Zhang, 2022)

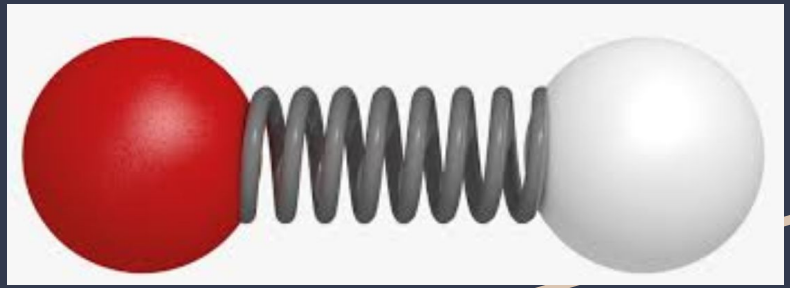
Goal

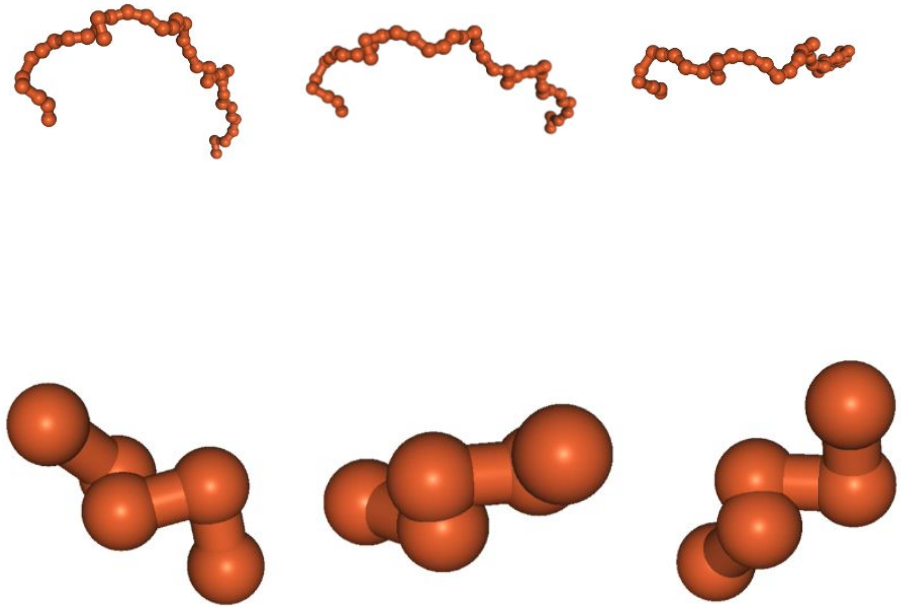
Detect loop extrusion in living systems by comparing “movies" of DNA with loop extrusion that are played forward and backward in time.

Workflow

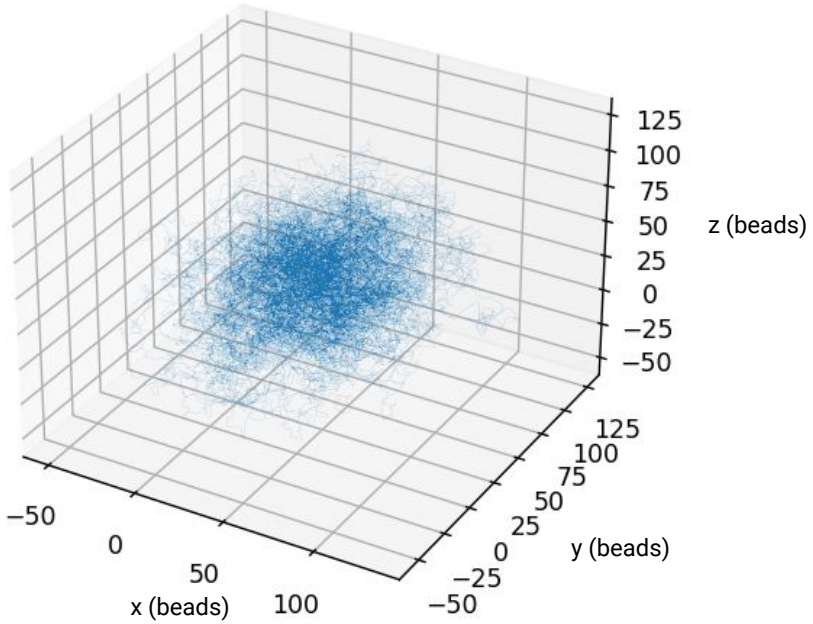


Polychrom Synthetic Dataset



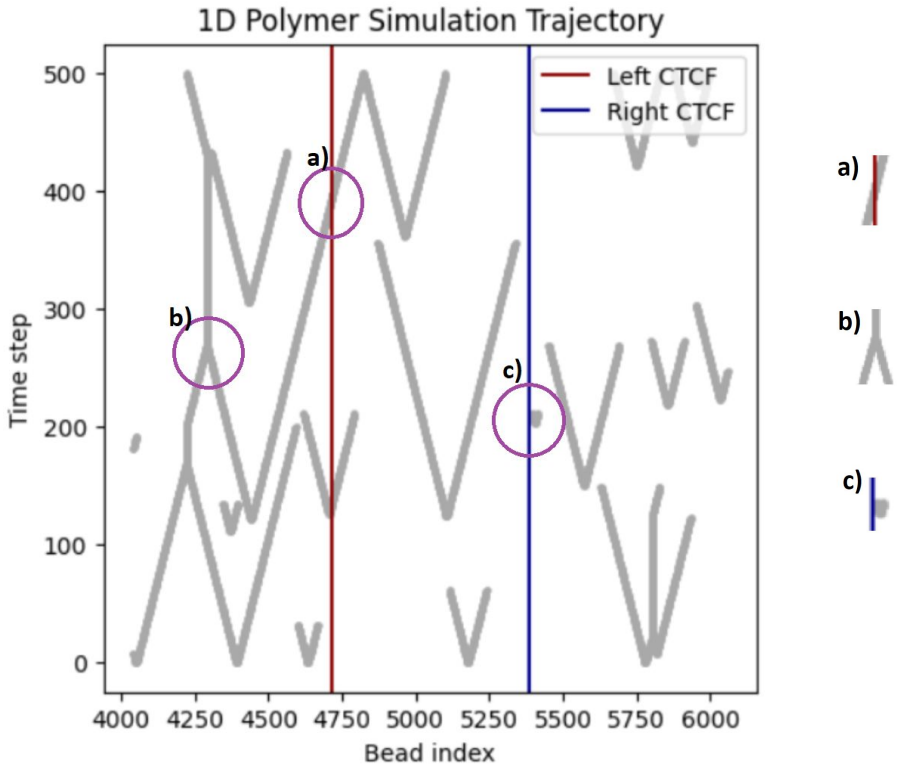


Equilibrated Polychrom Simulations

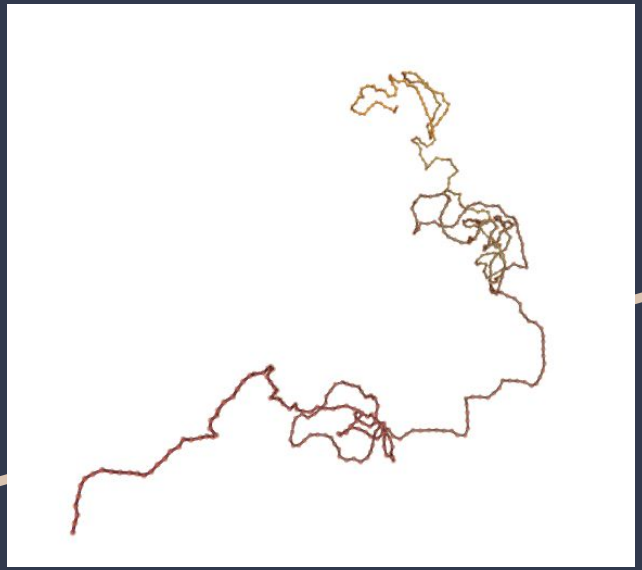


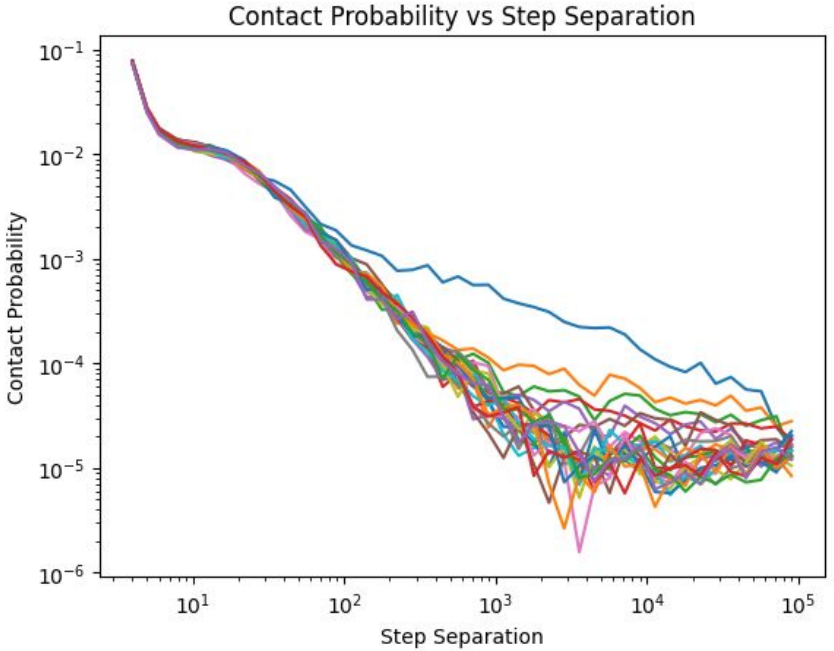
Representation of the Polychrom generated polymer using MatPlotLib

1D Loop Extrusion Simulations

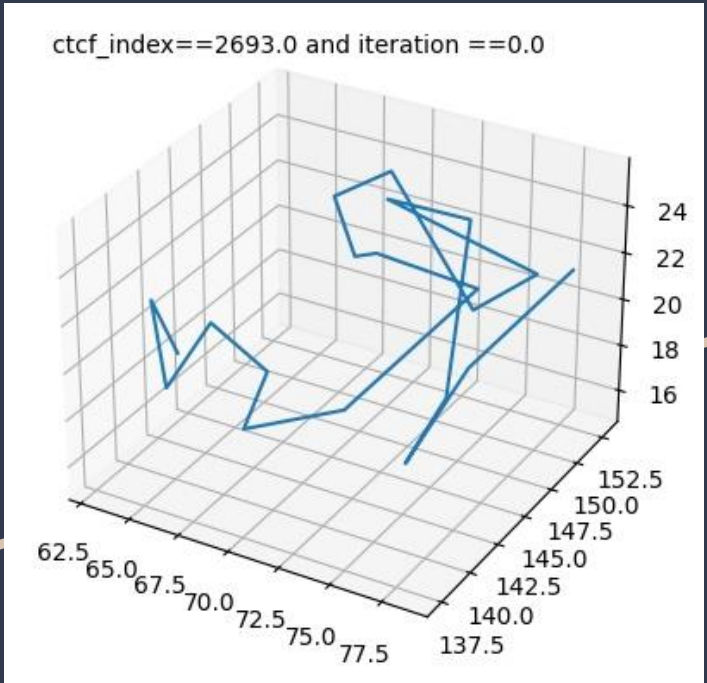


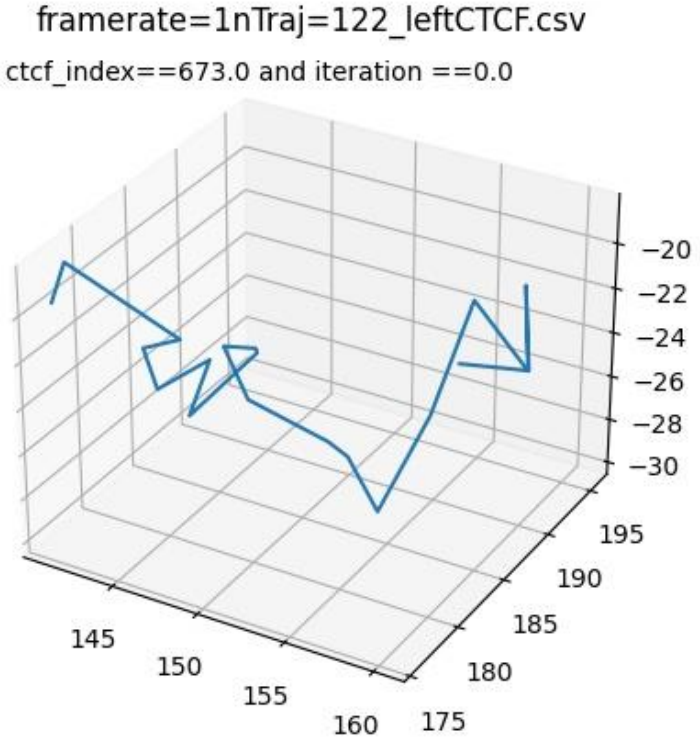
3D Loop Extrusion Simulations





CTCF Location Tracking

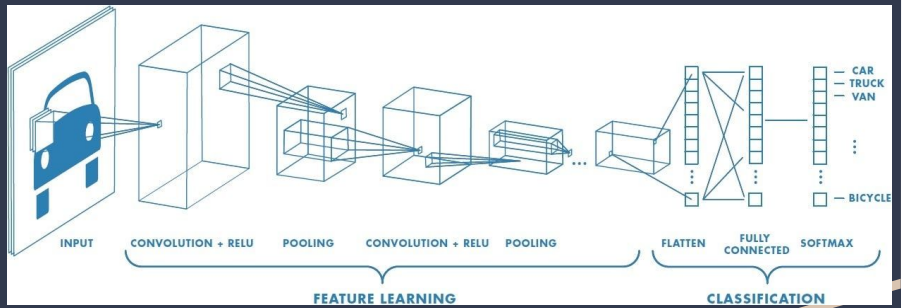




Time Reversal

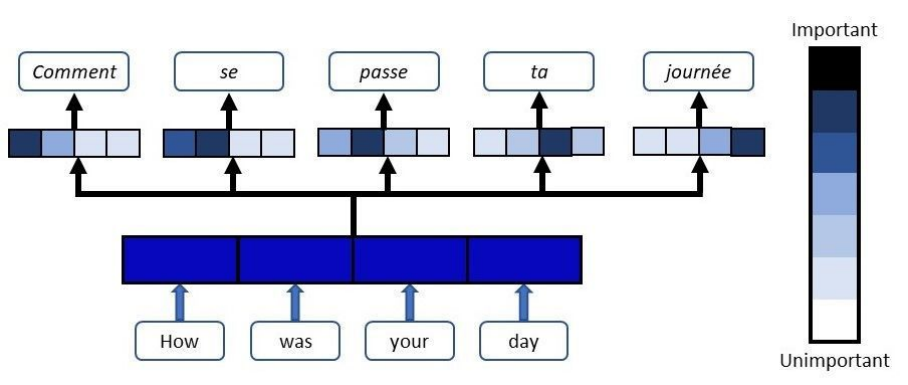
- Create a synthetic dataset for assessing the potential for loop extrusion detection
- DNA motion movies exhibit distinct behaviors when played forward compared to when played backward due to the influence of loop extrusion
- Quantitative method to distinguish between reversed and non-reversed DNA movement.

CNN



- Time-dependence relates to forward and backward propagation in neural networks.
- Create more efficient pathways to generate numerical training data both time dependent and multi-variant.
- Seek correlations both forward and backward through time.
- Help observe changes in genetic systems that can form a potential mutation event, which could form cancer, diabetes, or other diseases.

Alternative Strategies



- Transformers can be used as an alternative model to run simulations
  - Utilizes attention to focus on the entire sequence as a whole
  - Requires fewer training data to form a more efficient and accurate model
  - Handle sequences of variable lengths

Current Applications

- Used as an alternative to live imaging cells on the sub-nuclear level
- Understand the fundamental nature of the structure of sub-nuclear arrangements of genetic material
- Identify targeted therapies for cancers and diseases, based on driver and passenger mutation

Acknowledgement

- MIT PRIMES organizers and coordinators
- Professor Leonid Mirny, Mr. Henrik Pinholt, and Dr. Aleksandra Galitsyna from the MIT Mirny Lab

References

Gabriele, M., Brandão, H. B., Grosse-Holz, S., Jha, A., Dailey, G. M., Cattoglio, C., Hsieh, T.-H. S., Mirny, L., Zechner, C., & Hansen, A. S. (2022). *Dynamics of CTCF and cohesin mediated chromatin looping revealed by live-cell imaging*. Science. http://dx.doi.org/10.1101/2021.12.12.472242

Lupiáñez, D. G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J. M., Laxova, R., Santos-Simarro, F., Gilbert-Dussardier, B., Wittler, L., Borschiwer, M., Haas, S. A., Osterwalder, M., Franke, M., Timmermann, B., Hecht, J., … Mundlos, S. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell*, *161*(5), 1012–1025. https://doi.org/10.1016/j.cell.2015.04.004

Popay, T. M., & Dixon, J. R. (2022). Coming full circle: On the origin and evolution of the looping model for enhancer–promoter communication. *Journal of Biological Chemistry*, *298*(8), 102117. https://doi.org/10.1016/j.jbc.2022.102117

Sabaté, T., Lelandais, B., Bertrand, E., & Zimmer, C. (2023). Polymer simulations guide the detection and quantification of chromatin loop extrusion by imaging. *Nucleic Acids Research*, *51*(6), 2614–2632. https://doi.org/10.1093/nar/gkad034

Thank You!