Submitter Name: Haohan Wang Submitter Email:haohanw@illinois.edu

Understanding Variations in Regulatory Networks Across Cell Types through Transformer Model with Knowledge on Regulatory Interactions

Xinyu Gu[1], Xin Li[2], and Haohan Wang[1,3,4]

[1]Department of Computer Science, University of Illinois Urbana-Champaign;
[2]School of Molecular & Cellular Biology, University of Illinois Urbana-Champaign;
[3]School of Information Sciences, University of Illinois Urbana-Champaign;
[4]Carl R. Woese Institute for Genomic Biology, University of Illinois Urbana-Champaign

Subtle variations in gene regulatory networks (GRNs) are pivotal in understanding cellular differentiation and disease pathogenesis. These networks, which govern gene expression, display significant differences across various cell types, influencing cellular functionality and behavior. The challenge in current research lies in decoding these variations at a granular, cell-type-specific level. Traditional approaches to GRN analysis often rely on broad, unsupervised network estimation methods, leading to an understanding that may overlook critical cell-specific nuances.

In this study, we address this gap by proposing an innovative approach that utilizes the Transformer model to predict cell types from transcriptomic data. The architecture in hypothesized to enhance the model's ability to capture the dynamic interplay of transcriptomics associated with specific cell types. This integration allows for a more nuanced understanding of GRN variations and their impact on cellular behavior.

Empirical results demonstrate a significant improvement in cell-type prediction accuracy, which suggests that our model effectively captures the intricate dynamics between transcriptomics and cell types. By embedding the regulatory network within the model, we enable dynamic updates to the GRN structure during training.

Our approach offers a powerful tool for researchers to unravel the complexities of gene regulation at a cellular level. The implications of this study, particularly in the realm of substance use disorders, are significant. By enhancing our understanding of cell-type-specific gene regulatory networks, this research could inform the development of more effective strategies for SUD treatment and prevention. It highlights the potential for personalized therapeutic approaches that consider individual genetic and epigenetic profiles.