[BIOINFO 2025]

Cell-type-specific network inference leveraging single cell foundation models and graph neural networks

Yurim Jung[1] , Kyungwoo Song[2], and Insuk Lee[1*]

[1] *Department of Biotechnology, College of Life Science & Biotechnology, Yonsei University, Seoul 03722, Republic of Korea*

[2] *Department of Statistics and Data Science, Yonsei University, Seoul 03722, Republic of Korea*

**Corresponding author: insuklee@yonsei.ac.kr*

Single-cell network biology has proven useful for elucidating key genes and pathways involved in development and the identity of specific cell types. Recently, single-cell foundation models (scFMs) pretrained on millions of cells with gene expression data have been developed and shown to be effective for a range of downstream tasks in single-cell biology, including gene network inference. However, owing to the high noise and sparsity of single-cell gene expression data, gene regulatory relationships inferred from scFMs alone, while reasonably accurate, may exhibit variability and benefit from further refinement. To address this, we incorporated a graph neural network (GNN) into scFM-based network inference. For each cell population of the same type, we constructed a context-specific gene network by filtering a reference interactome, HumanNet, with coexpression patterns observed across cells. Initial node features were derived from scFM gene embeddings and aggregated by mean pooling across cells. The GNN was then trained for each cell-type using cross-validation designed to prevent data leakage between labeled and unlabeled edges, yielding interaction probabilities for gene pairs. Evaluation of the resulting cell-type-specific networks demonstrated high accuracy. Furthermore, we observed that pathways associated with each cell-type were significantly enriched, and that marker genes exhibited increased hubness within the inferred networks.