Cell Type-specific Gene Regulatory Network Atlas: A Repository of Cell-type Networks with Disease and Pathway Annotation

RUOSHUI LIU, Michael Cheng, Julie Tran, Xia Yang

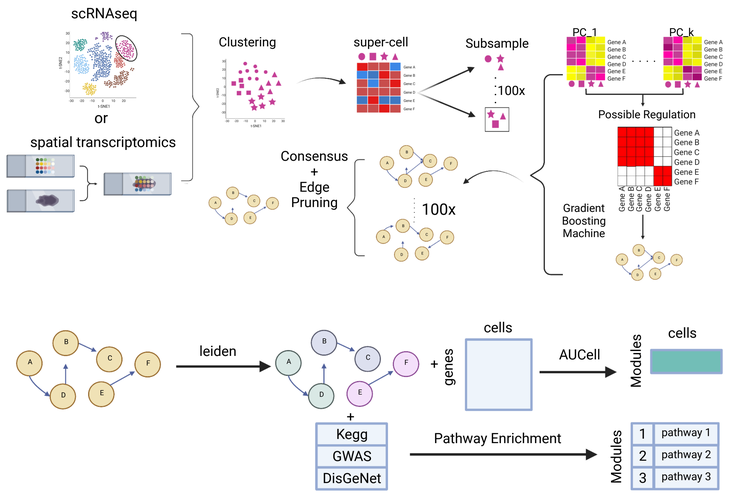
UCLA

YANG LAB

Abstract

- Gene regulatory networks (GRN) help understand physiology, but conventional tissue-level GRN inference lacks the resolution to capture cell-specific contributions in disease development; current cell-specific GRN algorithms are suboptimal in accuracy and speed.
- We developed Single Cell INtegrative Gene regulatory network inference (SCING), as it can capture cell-type characteristics.
- We applied SCING to create a network atlas for different cell types and revealed that the networks exhibit a highly connected architecture adhering to a power-law distribution.
- We conducted Key Driver Analysis (KDA) on a hepatocype network along with Non-alcoholic fatty liver disease (NAFLD) genes, where the key driver hubs recapitulated known pathways significantly altered in NAFLD.

SCING Method

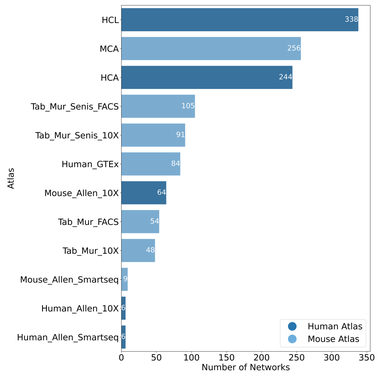


- SCING mitigates scRNAseq gene sparsity by aggregating cells based on similar expression.
- SCING bootstraps from supercells to build GRNs to account for technical variation between networks.
- SCING trains a gradient boosting regressor on each gene and selects predictors/features based on K nearest neighbors in the principal component space.
- SCING creates the final GRN by retaining edges present in at least 20% percent of bootstrap networks.
- SCING uses three resolutions to create modules of different sizes.
- SCING conducts pathway enrichment based on each module and filters the most frequent pathways.

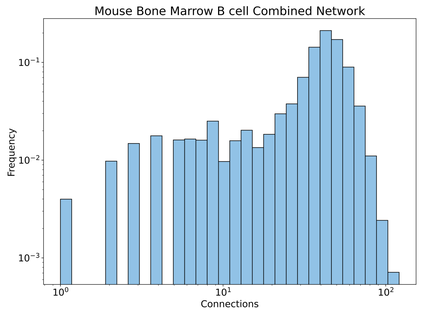
Result: Network Statistics

Overall Statistics

- We performed SCING inference on 12 human and mouse cell atlases and built 1305 networks from 91 tissues and 179 cell types.
- The networks have around 10,000 to 15,000 nodes and 40,000 to 60,000 edges on average.

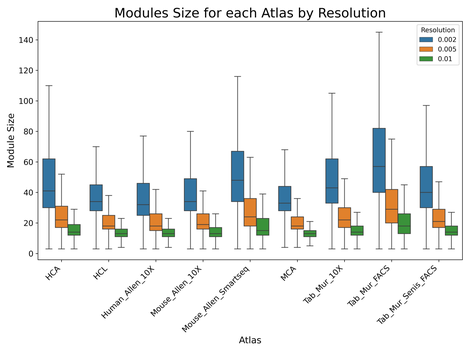


Combined Networks Characteristics



- We combined atlas networks based on species, tissue, and cell type and formed 152 combined networks.
- The networks are sparse on the left side and follow a power-law distribution on the right side.

Result: Module Analysis

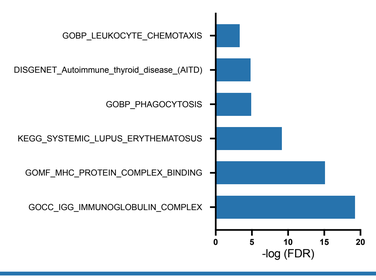


Module Properties

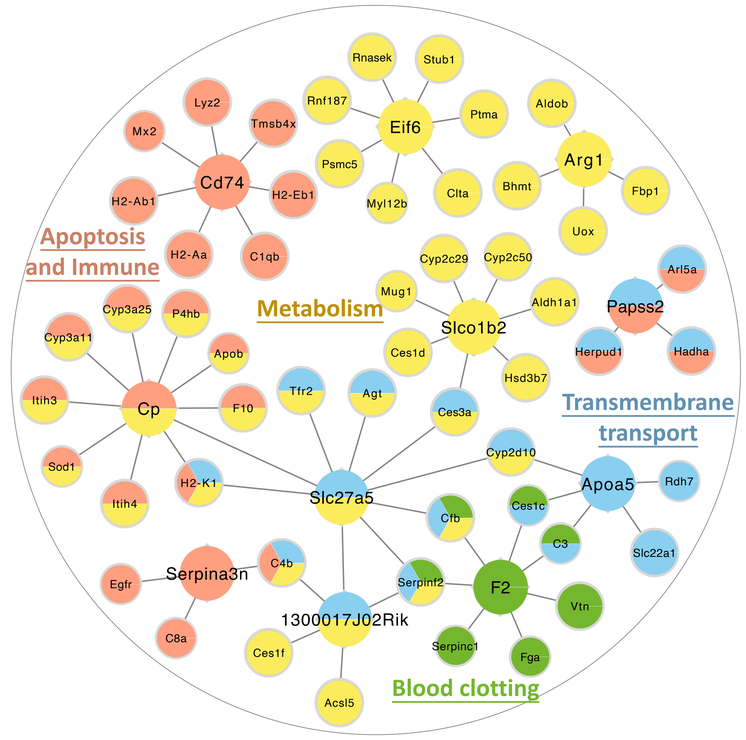
- We applied 3 module resolutions to each atlas network.
- Each module of the same resolution covers a different set of genes.
- Different resolutions cover different module sizes.

Pathway Annotation

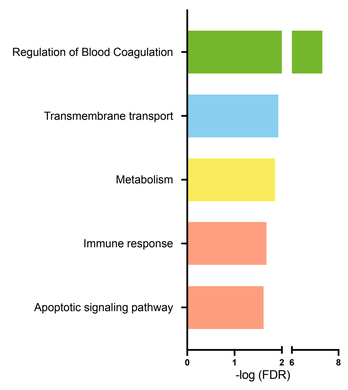
- We conducted pathway enrichment for the modules in each network.
- Pathways from thyroid b cells are highly related to immune mechanisms and diseases.



Result: Disease Annotation



- We performed KDA analysis on mouse hepatocyte network along with the differentially expressed genes (DEGs) extracted from a non-alcoholic fatty liver disease (NAFLD) study.
- We conducted pathway enrichment on each hub.
- Key driver hubs belong to pathways in apoptosis, metabolism, transmembrane transport, and blood clotting, which are consistent with the original study.



Reference

- Littman, Russell, et al. "SCING: Inference of Robust, Interpretable Gene Regulatory Networks from Single Cell and Spatial Transcriptomics." *iScience* (2023).
- Kurt, Zeyneb, et al. "Tissue-specific pathways and networks underlying sexual dimorphism in non-alcoholic fatty liver disease." *Biology of sex differences* 9.1 (2018): 1-14.