Single Cell RNA-sequencing Analysis of Human Lung Tissue Samples

THE OHIO STATE UNIVERSITY
COLLEGE OF MEDICINE

Grace Xu, Qin Ma, Melanie Koenigshoff, Cankun Wang
Bioinformatics and Mathematical Biosciences Lab. https://u.osu.edu/bmbl/
Department of Biomedical Informatics

Background

The TriState SenNet Consortium seeks to comprehensively map senescent cells in the human heart and lung using a multi-modality approach involving various unbiased approaches on whole tissues, organ slices, and isolated cells in order to better characterize the senescent cell population and to identify the physiologically relevant triggers for senescent cell formation.

Utilizing Single-cell RNA sequencing (scRNA-seq), individual lung cells are isolated, and their gene expression profiles are decoded. This vast dataset is analyzed to reveal unique gene expression patterns, uncovering genetic diversity within various cell types.

Objectives

- Filter 20 human lung tissue data sets only to include cells of high-quality
- Perform downstream analysis to develop a UMAP of scRNA-seq data, clustering cells to reveal distinct cell population types

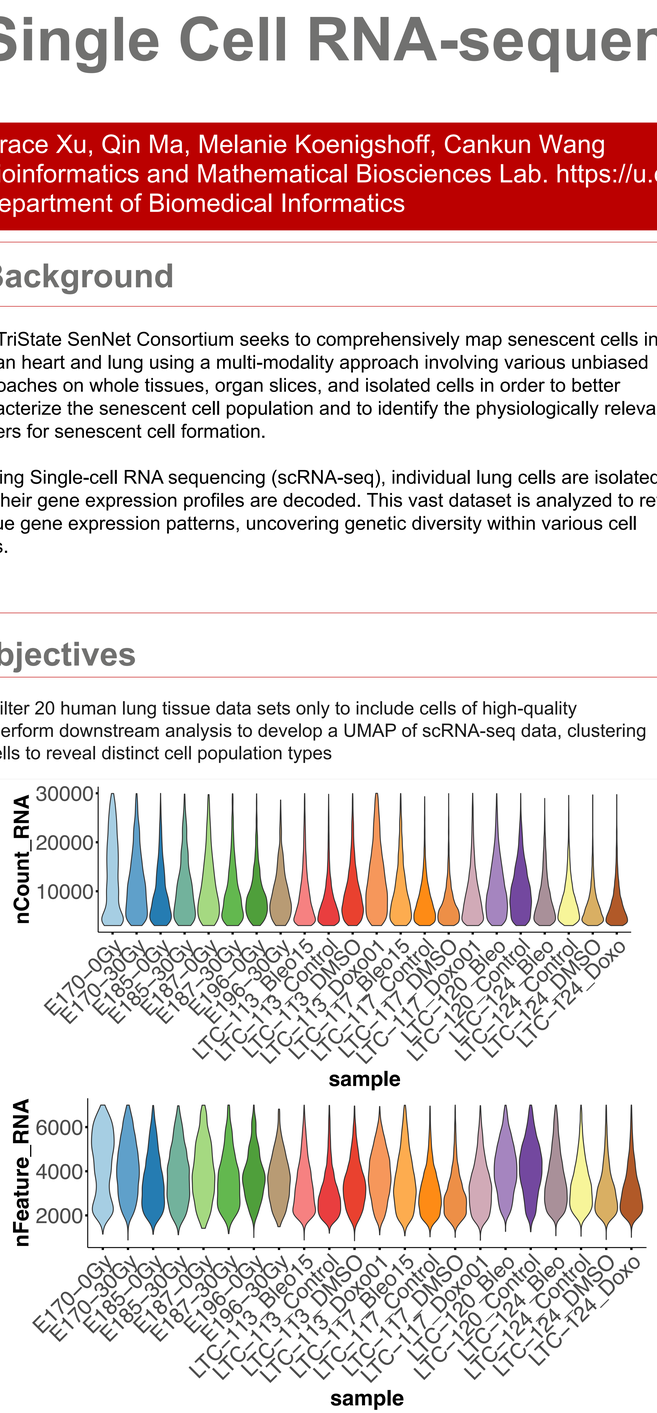

Figure 1. Violin plot displaying cutoff after cells with less than 6,500 genes and less than 30,000 total counts are removed.

Methods

- 20 precision-cut lung slice scRNA-seq data samples were obtained from TriState SenNet
- Utilized RStudio version 4.2.1 and Seurat package version 4.3.0.1 to perform downstream analysis
- Established specific cutoffs to filter out low-quality cells:
  - Cells with less than 200 genes or more than 30,000 genes
  - Cells with over 40% ribosomal reads or above 20% mitochondrial reads
  - Genes detected in fewer than three cells
- Subsequent dimension reduction was performed to generate UMAP plot, which displays clusters of cells with similar gene expression patterns.

Results

- Assisted Cankun to generate a portal containing scRNA-seq analysis results
- https://bmblx.bmi.osumc.edu/sennet_pcls_scrnaseq/

Summary

- Thorough quality control measures enhanced the reliability of data, resulting in a refined dataset with 156,761 high-quality cells.
- Dimension reduction techniques, represented by the UMAP plot, revealed distinct cell populations with unique gene expression patterns.

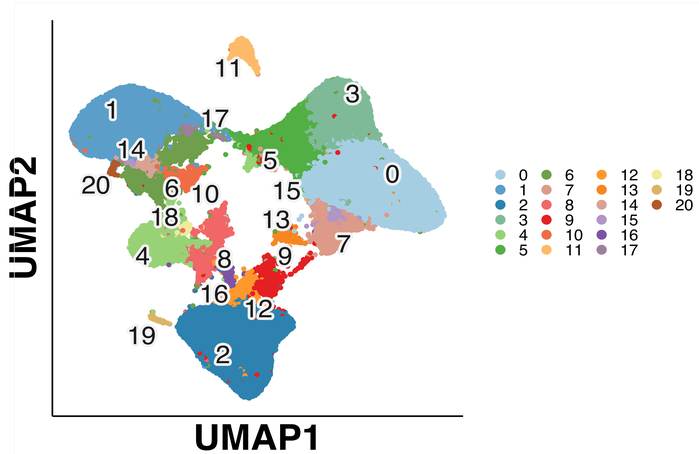

Figure 2. UMAP Plot of Preprocessed Lung Tissue scRNA-seq Clusters.

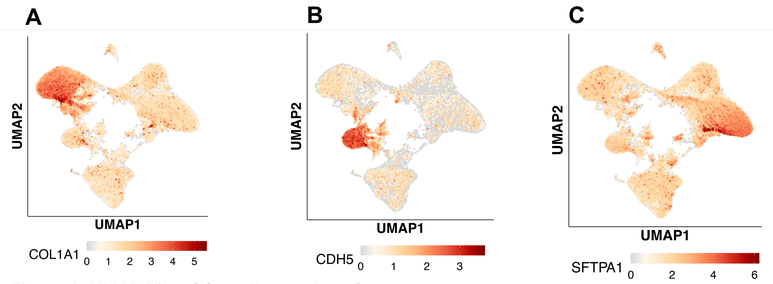

Figure 3. UMAP Plot of Gene Expression of
(A) COL1A1 (Fibroblasts)
(B) CDH5 (Smooth Muscle Cells)
(C) SFTPA1 (AT2)

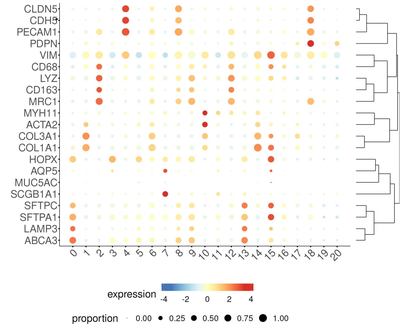

Figure 4. Bubble Plot of Marker Gene Expression Patterns in Seurat Clusters.

Acknowledgments and People

Grace Xu would like to extend her heartfelt gratitude to Cankun Wang and Dr. Qin Ma for their invaluable guidance and mentorship throughout the course of this research. Their insights and support have been instrumental in the progression and success of this work.

Qin Ma
The Ohio State University

Melanie Koenigshoff
University of Pittsburgh

Cankun Wang
The Ohio State University

References

1. Seurat: Integrated analysis of multimodal single-cell data
2. ShinyCell: simple and sharable visualization of single-cell gene expression data