Exploring Tumor Microenvironment Interactions: Tools and Workflows for ER+ Breast Cancer

Kiara Monahan, Nikita Chauhan, Lijun Yao, Alejandro Chibly

UC Berkeley Genentech

Introduction

In a cell signaling pathway, as shown on the right, a ligand sent from one cell interacts with a receptor of the target cell, triggering a cascade of events that ultimately results in a response in the target cell, as seen in Figure 1. Cell-cell communication analysis involves utilizing single-cell RNA sequencing data to study how different cell types interact with each other through examining expression levels of genes encoding ligands, which are signaling molecules, and especially useful in the study of cancer, where there is often abnormal cell-cell communication. Our project aims to develop a streamlined workflow to implement existing cell-cell communication (CCC) inference methods on single-cell RNA-seq data. We use data from ER+ breast cancer tissues [1] specifically examining interactions between tumor cells and fibroblasts to model these analyses and develop our workflows.

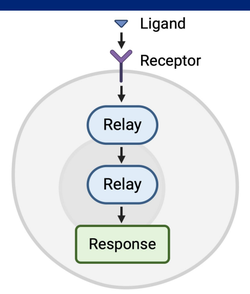

Figure 1: Ligand-receptor signaling pathway, showing the cascade from ligand binding to cellular response

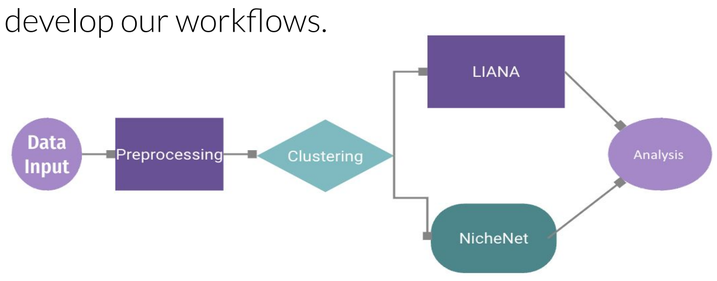

Figure 2: Diagram illustrating the streamlined workflow for analyzing cell-cell communication in ER+ breast cancer data, integrating preprocessing, clustering, and comparative analysis using LIANA and NicheNet.

Methods

To perform the analyses, we employed Seurat, SeuratObject, LIANA, and NicheNet libraries in R using RStudio on High Performance Computing (HPC) resources. LIANA is a framework that integrates multiple CCC inference methods to predict ligand-receptor interactions, and NicheNet is a method to predict how ligands affect target gene expression in receiver cells. Workflows were developed using RMarkdown. LIANA predicts interactions between ligands and receptors and is useful for more general analyses of cell-cell communication between cell types. NicheNet is focused around the downstream effects of cell-cell interactions. This method requires a downstream set of genes. With this set of genes, it is possible to examine ligand-receptor interactions that have the potential to affect their expression. NicheNet is useful for examining cell-cell communication differences between two groups, as differentially expressed genes between the two groups can be used as the geneset of interest for the analysis.

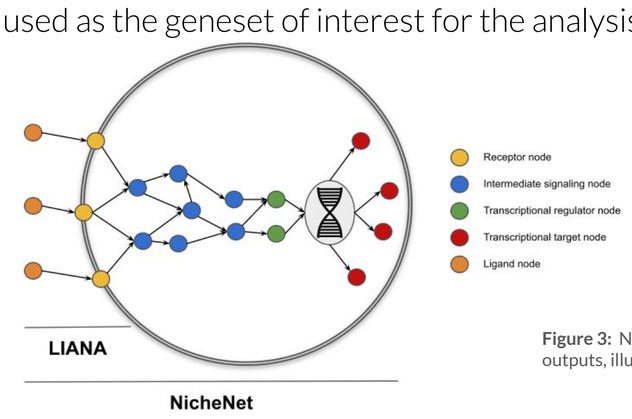

Figure 3 displays the differences between LIANA and NicheNet and their intended uses for cell-cell communication analysis.

Figure 3: Network diagram comparing LIANA and NicheNet outputs, illustrating receptor, ligand, and transcriptional targets.

Results

After quality control checks were conducted and plotted using the base Seurat Object file, a UMAP diagram was constructed, and cell types were labeled according to their gene expression profiles, which were plotted using heatmaps of differentially expressed genes and UMAP plots of specific marker expression. Some markers and cell type labels came from the original paper that contained the dataset[2]. Tumor cells are the largest group in this plot, as shown on the right.

To check for batch effect, we plotted UMAP with coloring by patient ID, shown in Figure 5. It appears that the various samples were grouped relatively similarly.

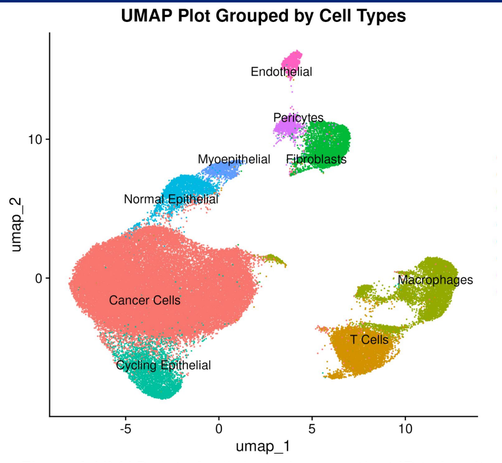

Figure 4: UMAP plot of cell clustering by patient ID to idenity patient-specific cell populations

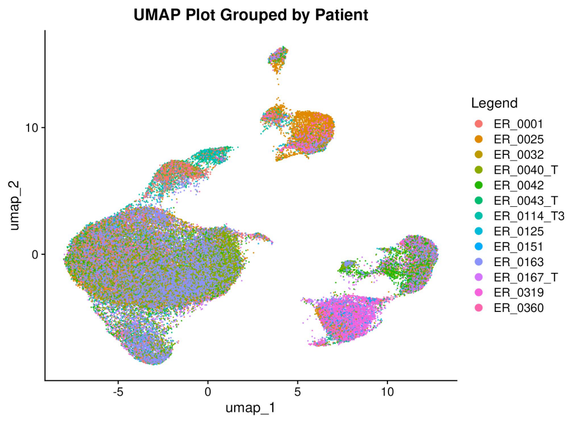

Figure 5: UMAP plot of cell identities clusterest at resolution 0.5, highlighting distinct cell types

With LIANA, we examined the overall patterns of potential cell-cell interactions between cell types. The heatmap in Figure 6 demonstrates that fibroblasts were the most frequent sender, with fibroblasts, endothelial, and myoepithelial cells as the most common receivers.

More specifically, we examined the interactions between cycling epithelial cells, cancer cells, and fibroblasts with LIANA.

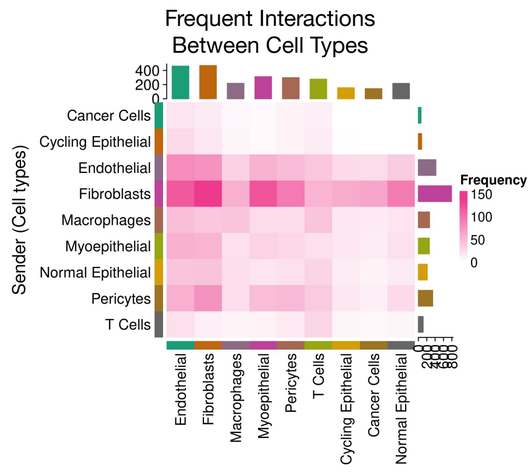

Figure 6: Generalized heatmap of frequent interactions between cell types according to LIANA

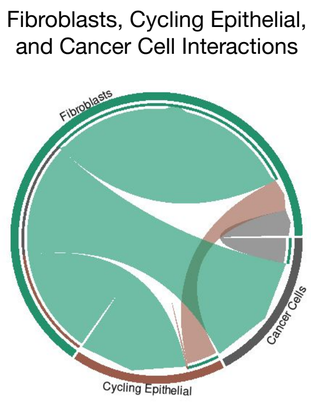

Figure 7: Chord diagram showing overall volume of interactions between fibroblasts, cancer cells, and cycling epithelial cells, using LIANA

The chord diagram in Figure 7 displays in general the amount of potential interactions between the three cell types. Larger chords mean more interactions between the cell types.

On the right are displayed in more detail using a dot plot the strongest potential interactions between these cell types. Interactions being sent from fibroblasts had the highest expression magnitude, and few potential interactions were found coming from cancer or cycling epithelial cells.

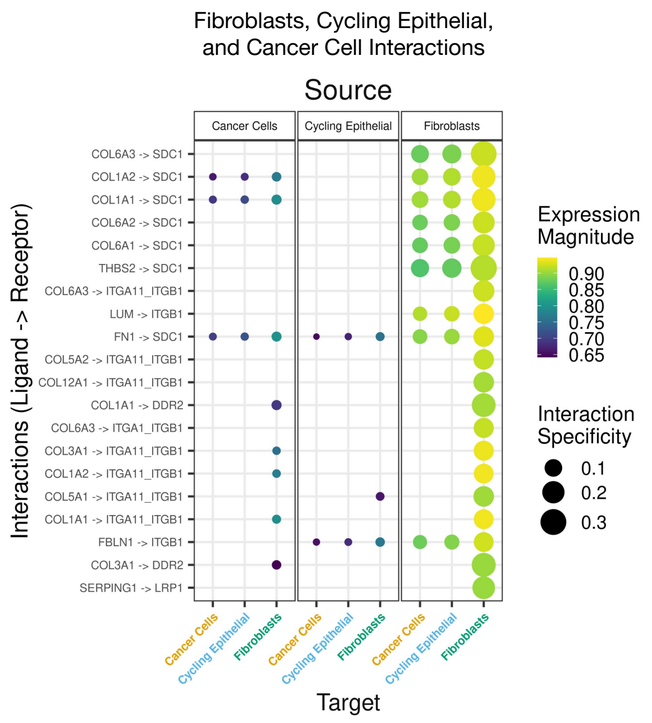

Figure 8: Dot plot of strongest potential interactions between fibroblasts, cancer cells, and cycling epithelial cells, created with LIANA.

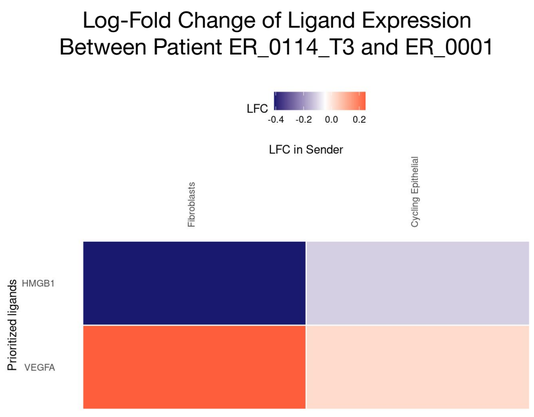

Figure 9: Log-fold change plot comparing VEGFA and HMGB1 expression between two patients using NicheNet.

NicheNet was utilized to examine differences in cell-cell communication between two patients, using differentially expressed genes between the two patients as the geneset of interest. In Figure 9,, we see moderate differences in expression between two prioritized ligands. ER_0001 has higher expression of VEGFA and lower expression of HMGB1 than ER_0114_T3

Conclusions

Our study included LIANA and NicheNet to simplify the analysis of cell-cell communication in ER+ breast cancer data. Tumor-fibroblast coupling revealed novel tracks through which the tumor microenvironment operates on cellular pathways, critical in cancer progression. Integrating these tools provided a comprehensive perspective on ligand-receptor interactions and downstream effects, offering new ways for targeting and biomarkers in ER+ breast cancer.

Future Applications

We have developed streamlined workflows for these analyses that can be applied to other datasets, cell types, and conditions of interest. By altering a YAML, users will be able to easily conduct complete analyses and generate HTML output files using our .rmd files. This simplifies the analysis process and will allow researchers to spend less time learning, writing, and editing code.

References

(1) Chen, Yunshun; Smyth, Gordon (2022). Data, R code and output Seurat Objects for single cell RNA-seq analysis of human breast tissues. figshare. Dataset. https://doi.org/10.6084/m9.figshare.17058077.v1

(2) Rolando Vegliante, Ievgenia Pastushenko, Cédric Blanpain, Deciphering functional tumor states at single-cell resolution, The EMBO Journal, 10.15252/embj.2021109221, 41, 2, (2022).