Table of Contents

[1. scRNA-seq of B-cell development stages 1](#_Toc145447391)

[1.1.data collection: download HCA 1](#_Toc145447392)

[1.2.network building 1](#_Toc145447393)

[1.3. infer activity 1](#_Toc145447394)

[1.4. web portal 1](#_Toc145447395)

[2. scMultiome profiling of B-ALL cases 1](#_Toc145447396)

[3. NetBID2 analysis of Asp pharmacotyping cohort 1](#_Toc145447397)

[3.1.data preprocess 1](#_Toc145447398)

[3.2.network building 2](#_Toc145447399)

[3.3. infer activity 2](#_Toc145447400)

[Data 2](#_Toc145447401)

# 1. scRNA-seq of B-cell development stages

## 1.1.data collection: download HCA

we utilized the biggest healthy human BM single-cell RNA sequencing dataset in human cell atlas Atlas Census of Immune Cells project and analyzed it by Seurat.

See scripts: 1.1.downloadHCA.R

## 1.2.network building

we used the SJARACNe algorithm on scRNA-seq profile of normal human BM B cells above to computationally reverse engineer cell-type-specific interactomes for each of the major cell types (HSC, CLP, Pre-pro-B, Pro-B, Pre-B, Immature B, Mature B, and Plasma cells). tf\_sigs\_hg.RData is a predefined driver list including human signaling proteins (SIG) and transcription factors (TF).

See scripts: 1.2.1.generateSJARACNeInput.R; 1.2.2.generateSJARACNenetwork.sh

## 1.3. infer activity

we inferred their network activities in each cell from the scRNA-seq data by using the NetBID2 (data-driven Network-based Bayesian Inference of Drivers) algorithm with the interactome of the corresponding cell type.

See scripts: 1.3.GetActivityFromSJARACNe.R

## 1.4. web portal

We also built a visual web portal for scRNA-seq expression and activity data of human BM B-cell development stages.

# 2. scMultiome profiling of B-ALL cases

Cell Ranger ARC pipeline (version 2.0.0) from 10x Genomics was used to generate barcoded count matrices of gene expression and ATAC data. For each sample, count matrices were loaded in Seurat v4.1.0.

See scripts: 01\_run\_mkfastq\_ATAC.sh; 02\_run\_mkfastq\_GEX.sh; 03\_run\_Count.sh;

# 3. NetBID2 analysis of Asp pharmacotyping cohort

## 3.1.data preprocess

We set up RNAseq data analysis pipelines using RSEM or Salmon package. Here we used Salmon as NetBID2 provides function load.exp.RNASeq.demoSalmon() for loading an expression dataset derived from RNA-Seq.

See scripts: 3.1.1.quantification\_Salmon.sh; 3.1.2.load.exp.RNASeq.demoSalmon.R

## 3.2.network building

due to the heterogeneity among B-ALL subtypes, we reverse-engineered B-ALL subtype-specific gene-gene interactomes (B-ALLi) for each subtype from the RNA-seq dataset of a published B-ALL cohort with 1986 cases by using the SJARACNe algorithm.

See scripts: 3.2.generateSJARACNeInput.R

## 3.3. infer activity

Overlaying B-ALLi to the RNA-seq data of the pharmacotyping, we then inferred activity of each hub gene or driver on the basis of expression of its targets weighted by the interaction strength of each hub–target pair.

See scripts: 3.3. cal.Activity.R

# Data

1.2.tf\_sigs\_hg.RData; GEO\_table.xlsx; SampleSheet.csv;

1.4.1.Expression\_matrix.csv; 1.4.2.Activity\_matrix.csv

To be uploaded