Cell cell communication in COVID data

Yue Cao

2025-3-26

Introduction

In this analysis, we examine a COVID-19 single-cell RNA-sequencing data obtained from the study " COVID-19 severity correlates with airway epithelium–immune cell interactions identified by single-cell analysis" by Chua et al., published in Nature Biotechnology. The paper sequenced nasopharyngeal and bronchial samples to examine molecular mechanisms behind COVID-19 disease severity.

In this analysis, we aim to perform some basic data exploration follow by cell cell interaction analysis. Cell cell interaction infers the communication between individual cells. Given a single-cell RNA-sequencing data, it predicts the interaction between cell types and between ligand receptor pairs based on the gene expression of individual cells.

Basic data exploration

First, we subset the data to moderate and critical patients and explore the dataset.

```
chua <- load_data()

chua <- chua[, chua$severity %in% c("moderate" , "critical")]

meta_chua <- colData(chua)
meta_chua  <- meta_chua [!duplicated(meta_chua$sample), ]
rownames(meta_chua) <- meta_chua$sample

meta_chua_unique <- meta_chua[!duplicated(meta_chua$sample), ]

print("dimension of the data")
```

```
## [1] "dimension of the data"
```

```
dim( chua)
```

```
## [1]  16355 132618
```

```
print("number of samples in each severity class")
```

```
## [1] "number of samples in each severity class"
```

```
table(as.character( meta_chua_unique$severity) )
```

```
##
## critical moderate
##       13       14
```

Visualising the cell type composition of the samples. There are 13 cell types in total, B, Basal, Basophil/Mast , Ciliated , Dendritic, Fibroblast, Goblet, intermediate, Macrophage, Monocyte, Neutrophil, T , unassigned.

Neurophil, goblet, T cells are some of the major cell types.

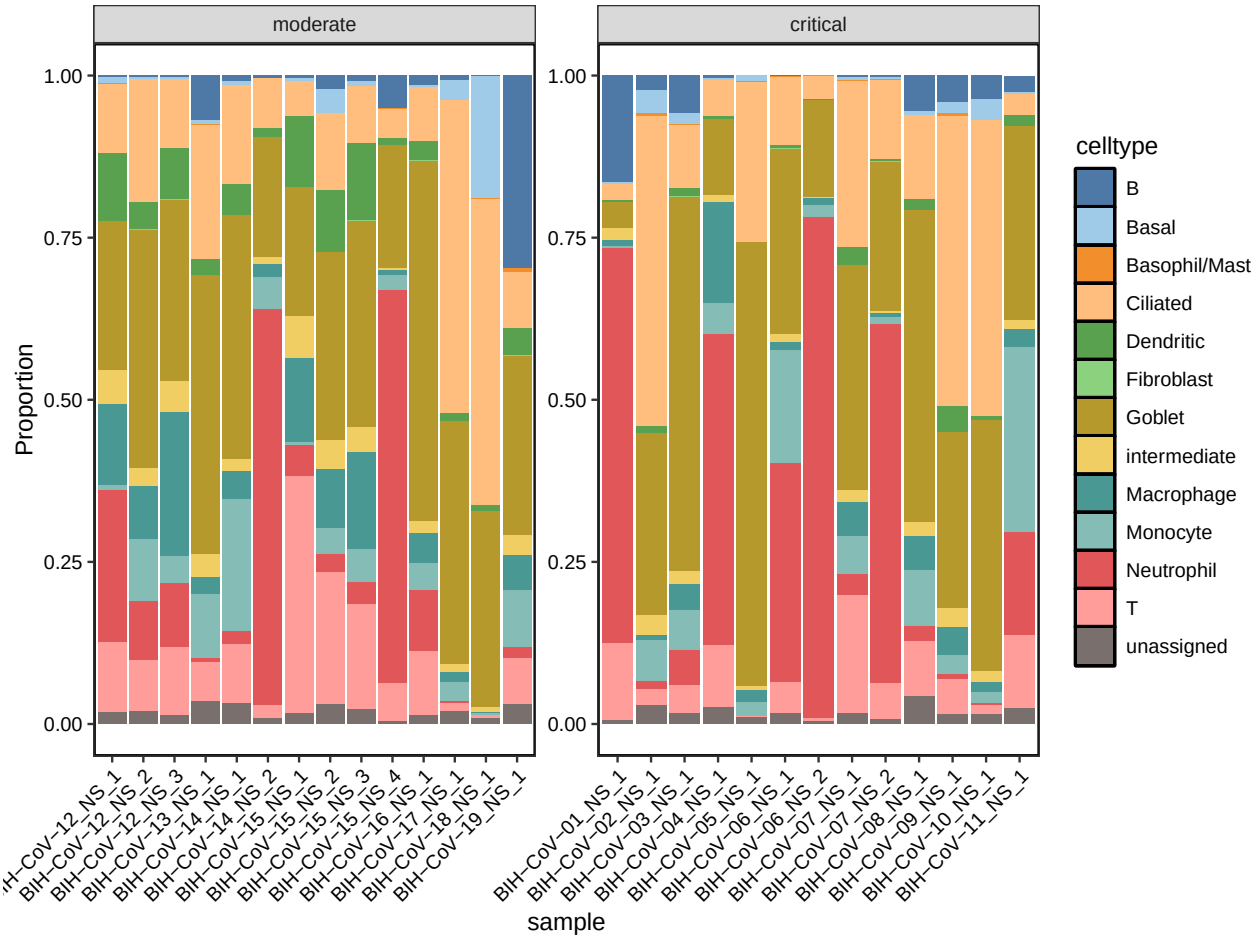
```
df <- data.frame( table(chua$sample, chua$pred_chua_scClassify) )

colnames(df) <- c("sample" , "celltype" , "count" )

df  <- suppressMessages(  df  %>%
  dplyr::group_by( sample, celltype) %>%
  dplyr::summarise(Total = sum(count)) %>%
  dplyr::mutate(Proportion = Total / sum(Total)) )

df$severity <- chua[ , match(df$sample, chua$sample )]$severity


ggplot( df, aes(x = sample, y = Proportion, fill = celltype)) +
  geom_col()  + facet_wrap(~severity,scale = "free") +
  scale_fill_tableau("Tableau 20")
```



Exploring cell cell communication

Now, we perform cell cell communication algorithm and explore the results. We first ran cell cell communication analysis using a popular method called CellChat. CellChat infers the cell cell communication within each individual. We then aggregates the outputs within moderate and within critical samples to compare the overall profile of the moderate and the critical samples.

```
cellchat_res_list <- run_CCI()
```

```
aff_mat_bySample <- summarise_CCI_sample(cellchat_res_list )


critical_patients <- rownames(meta_chua)[meta_chua$severity == "critical"]

aff_mat_critical <- Reduce("+", aff_mat_bySample[names(aff_mat_bySample) %in%
                                                   critical_patients])/length(critical_patients)


moderate_patients <- rownames(meta_chua)[meta_chua$severity == "moderate"]

aff_mat_moderate <- Reduce("+", aff_mat_bySample[names(aff_mat_bySample) %in%
                                                   moderate_patients])/length(moderate_patients)
```

Overall cell cell communication pattern between moderate and critical patients

After we obtain the output from the algorithm, we first examine the cell cell communication patterns between pairwise cell types and compare the patterns between moderate and critical samples.

The plot below visualises the difference between moderate and critical samples, where red indicates a stronger signal in critical samples, blue indicates greater signal in moderate samples.

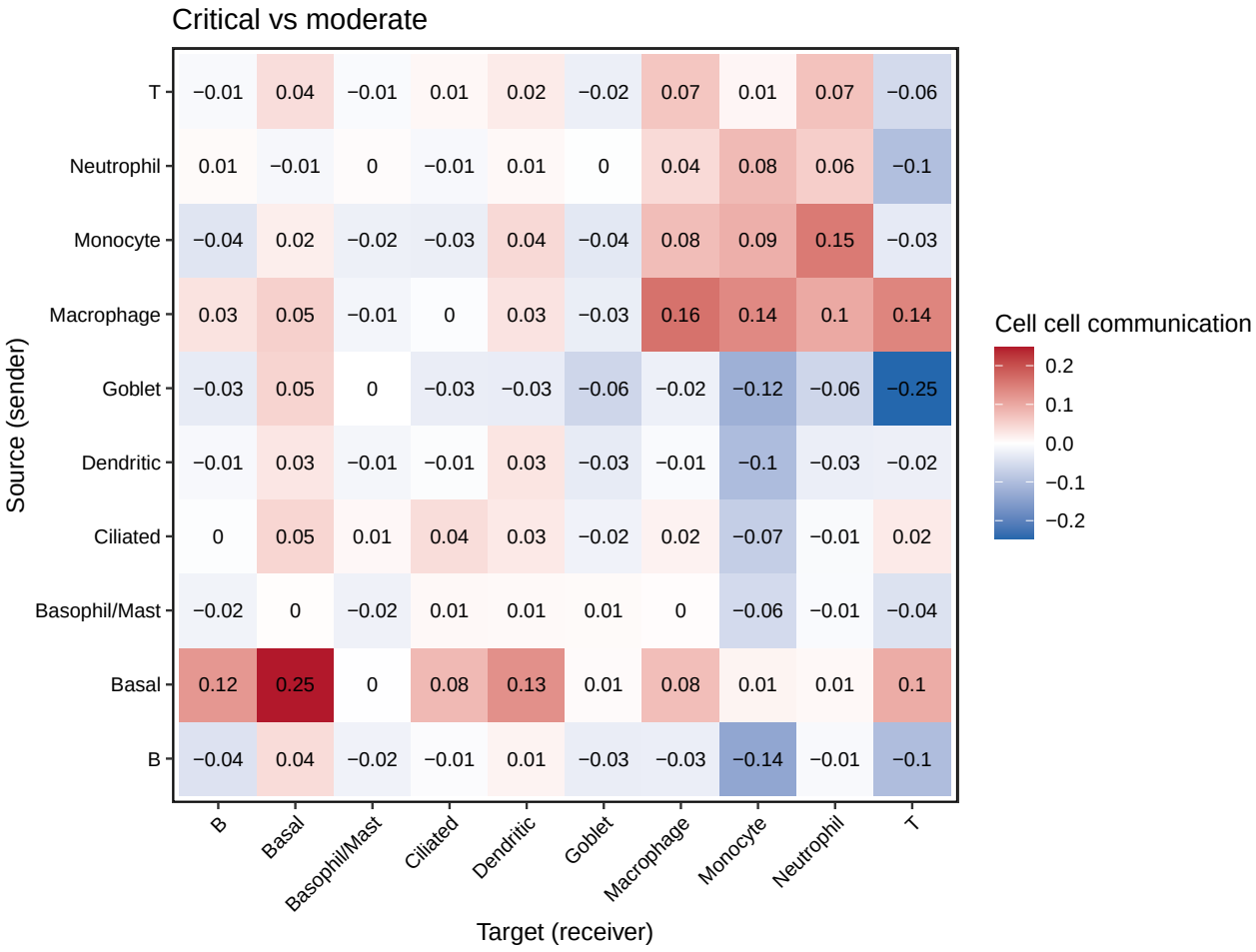
For example, in the mild patients, we see there is a stronger interaction between Goblet (sender cell type) with T (receiver cell type). In the critical patients, there is a stronger interaction between Monocyte (sender cell type) with neutrophil (receiver cell type).

```
mat <- aff_mat_critical - aff_mat_moderate



melted_data <- suppressMessages( melt(as.matrix(mat)) )
colnames(melted_data) <- c("Row", "Column", "Value")


ggplot(melted_data, aes(x = Column, y = Row, fill = Value)) +
  geom_tile() +
  scale_fill_gradient2(
    low = "#2166ac",
    mid = "white",
    high = "#b2182b",
    midpoint = 0,
    limits = c(min(melted_data$Value), max(melted_data$Value))
  ) +
  labs(fill = "Cell cell communication") +    geom_text(aes(label = round(Value,2)),
     color = "black", size = 3 ) + ylab("Source (sender)") + xlab("Target (receiver)") +
  ggtitle("Critical vs moderate")
```



Ligand receptor in the mild patients

Here we examine the top ligand receptor interactions in the mild patients. From the previous result, we noticed there is a greater interaction in the mild patients with goblet being the sender cell type. Therefore, we selectively focus on goblet being the sender, and macrophage, monocyte, neutrophil and T cells being the receiver and examine the top ligand receptor within these cell type pairs. We show the top 20 ligand receptor interactions, ranked by the interaction probability.

```
ligand_receptor <- get_ligand_receptor(cellchat_res_list )

ligand_receptor_moderate <-  ligand_receptor[ ligand_receptor$source %in% c("Goblet" ) &
               ligand_receptor$target %in%  c(  "Macrophage",  "Monocyte",
        "Neutrophil",  "T"),  ]


ligand_receptor_moderate <-  ligand_receptor_moderate[ ligand_receptor_moderate$sample %in%
          moderate_patients,  ]

ligand_receptor_moderate <-  suppressMessages(  ligand_receptor_moderate %>%
        dplyr::group_by( source.target, interaction_name_2) %>%
        dplyr::summarise(prob = mean(prob)) )


data.frame( ligand_receptor_moderate [order(ligand_receptor_moderate$prob ,
```

```
    decreasing = T ) , ][ 1:20, ] )
```

```
##           source.target  interaction_name_2      prob
## 1  Goblet -> Neutrophil        ANXA1 - FPR1 0.4914052
## 2    Goblet -> Monocyte        ANXA1 - FPR1 0.3883413
## 3           Goblet -> T        HLA-B - CD8A 0.3644147
## 4           Goblet -> T        HLA-A - CD8A 0.3570695
## 5           Goblet -> T        HLA-C - CD8A 0.3507049
## 6  Goblet -> Neutrophil        ANXA1 - FPR2 0.3315680
## 7  Goblet -> Macrophage   MIF - (CD74+CD44) 0.3154304
## 8    Goblet -> Monocyte        ANXA1 - FPR3 0.3142500
## 9           Goblet -> T        HLA-B - CD8B 0.3084079
## 10          Goblet -> T        HLA-A - CD8B 0.3030001
## 11          Goblet -> T        HLA-C - CD8B 0.2993934
## 12 Goblet -> Macrophage          APP - CD74 0.2965434
## 13   Goblet -> Monocyte ANXA1 - (FPR2+LXA4) 0.2913392
## 14 Goblet -> Macrophage        ANXA1 - FPR3 0.2827022
## 15   Goblet -> Monocyte          APP - CD74 0.2800994
## 16   Goblet -> Monocyte   MIF - (CD74+CD44) 0.2777662
## 17          Goblet -> T           MDK - NCL 0.2776401
## 18          Goblet -> T  HLA-E - CD94:NKG2C 0.2721093
## 19          Goblet -> T        HLA-E - CD8A 0.2702813
## 20 Goblet -> Neutrophil  MIF - (CD74+CXCR2) 0.2617477
```

ligand receptor in the critical patients

Here we examine the top ligand receptor interactions in the critical patients. From the previous result, we noticed there is a greater interaction in the critical patients with monocyte being the sender cell type and neutrophil being the receiver. Therefore, we examine the top ligand receptor within these cell type pairs. We show the top 20 ligand receptor interactions, ranked by the interaction probability.

```
ligand_receptor_critical <-  ligand_receptor[ ligand_receptor$source %in% c("Monocyte" ) &
               ligand_receptor$target %in%  c(  "Neutrophil"),  ]


ligand_receptor_critical <-  ligand_receptor_critical[ ligand_receptor_critical$sample %in%
 critical_patients,  ]

ligand_receptor_critical <- suppressMessages(  ligand_receptor_critical %>%
        dplyr::group_by( source.target, interaction_name_2) %>%
        dplyr::summarise(prob = mean(prob)) )


data.frame( ligand_receptor_critical[order(ligand_receptor_critical$prob ,
            decreasing = T ) , ][ 1:20, ] )
```

```
##             source.target    interaction_name_2      prob
## 1  Monocyte -> Neutrophil          ANXA1 - FPR1 0.3459013
## 2  Monocyte -> Neutrophil           CCL3 - CCR1 0.3429835
## 3  Monocyte -> Neutrophil         CXCL3 - CXCR2 0.3182781
## 4  Monocyte -> Neutrophil  SPP1 - (ITGA5+ITGB1) 0.2980590
## 5  Monocyte -> Neutrophil         CXCL2 - CXCR2 0.2936457
## 6  Monocyte -> Neutrophil          IL1B - IL1R2 0.2927104
## 7  Monocyte -> Neutrophil         CXCL3 - CXCR1 0.2868949
## 8  Monocyte -> Neutrophil          ANXA1 - FPR2 0.2812324
```

```
## 9  Monocyte -> Neutrophil NAMPT - (ITGA5+ITGB1) 0.2811578
## 10 Monocyte -> Neutrophil         CXCL5 - CXCR2 0.2806063
## 11 Monocyte -> Neutrophil           SPP1 - CD44 0.2779032
## 12 Monocyte -> Neutrophil         CXCL2 - CXCR1 0.2620764
## 13 Monocyte -> Neutrophil         LGALS9 - CD45 0.2409503
## 14 Monocyte -> Neutrophil         CXCL1 - CXCR1 0.2307019
## 15 Monocyte -> Neutrophil          PLAU - PLAUR 0.2290763
## 16 Monocyte -> Neutrophil         CXCL5 - CXCR1 0.2261698
## 17 Monocyte -> Neutrophil ICAM1 - (ITGAX+ITGB2) 0.2227672
## 18 Monocyte -> Neutrophil           CCL7 - CCR1 0.2222415
## 19 Monocyte -> Neutrophil         LGALS9 - CD44 0.2163627
## 20 Monocyte -> Neutrophil         CXCL1 - CXCR2 0.2151466
```