6-Gene Ontology enrichment

BAI Qiang*

$2021\hbox{-}10\hbox{-}06\ 20\hbox{:}29\hbox{:}29\ +0200$

Contents

| 1 | Description | 2 |
|--------------|--|-----------|
| 2 | Load data and packages | 2 |
| 3 | DE genes and GO enrichment | 2 |
| | 3.1 Cluster 1 DE analysis and GO enrichment | |
| | 3.2 Cluster 2 DE analysis and GO enrichment | |
| | 3.3 Cluster 3 DE analysis | 7 |
| | 3.4 Cluster 4 DE analysis | 7 |
| | $3.5~{ m GO/KEGG}$ enrichment and DE analysis in subpopulations of Cluster 3 after re-clustering | 7 |
| 4 | Session information | 12 |
| \mathbf{R} | eferences | 14 |

^{*}University Liege, mail qiang.bai@uliege.be

1 Description

The DE gene lists used for enrichment analyses were calculated using Seurat function FindMarkers[1] with default arguments except only.pos = TRUE in order to output only positively regulated genes and logfc.threshold of 0.2. KEGG and Gene Ontology enrichment analyses were made using enrichKEGG and enrichGO functions from package clusterProfiler package[2] with default arguments.

2 Load data and packages

```
library(Seurat)
library(ggplot2)
library(dplyr)
results <- readRDS(file = "../3-Merge and cell typing/so.merged_clusters.
    seuratObject.Rds")</pre>
```

3 DE genes and GO enrichment

3.1 Cluster 1 DE analysis and GO enrichment

The Cluster 1 mentioned in manuscript refers to cluster 0 in the codes following. The same for cluster 2-4 in manuscript, they are cluster 1-3 in the codes.

Calculate the DE genes.

```
de_ident0 <- FindMarkers(results, ident.1 = "0", only.pos = TRUE, logfc.
    threshold = 0.2)
head(de_ident0)</pre>
```

```
## # A tibble: 6 x 5
##
     p_val avg_log2FC pct.1 pct.2 p_val_adj
##
     <dbl>
                 <dbl> <dbl> <dbl>
## 1
         0
                 0.689 1
                              0.935
                                                                                       4
                                                                                       5
## 2
         0
                 0.682 0.994 0.892
                                              0
                                                                                       6
## 3
         0
                 0.622 0.998 0.912
                                              0
                                                                                       7
## 4
         0
                 0.379 0.993 0.941
                                              0
                                                                                       8
## 5
         0
                 0.733 0.477 0.107
                                              0
## 6
                 0.670 0.994 0.955
```

Save data

```
write.csv(de_ident0, file = "./Results_lists/DE_cluster1.csv")
```

KEGG/GO enrichment:

```
library(clusterProfiler)
source("../../R/entrez2symbol.R")
source("../../R/replaceEntrezID.R")

de_entrez.ident0 <- bitr( geneID = rownames(de_ident0), fromType = "SYMBOL",
    toType = "ENTREZID", OrgDb = "org.Hs.eg.db", drop = TRUE ) $ ENTREZID
result.enrichKEGG.de_ident0 <- enrichKEGG(de_entrez.ident0, organism = "hsa",
    keyType = "ncbi-geneid")
result.enrichKEGG.de_ident0 <- replaceEntrezID(result.enrichKEGG.de_ident0)
7</pre>
```

result.enrichKEGG.de_ident0@result

```
## # A tibble: 199 x 9
##
               Description GeneRatio BgRatio pvalue p.adjust qvalue geneID
   Count
##
      <chr>>
               <chr>>
                            <chr>>
                                       <chr>>
                                                 <dbl>
                                                          <dbl>
                                                                  <dbl> <chr>
                                                                                 3
   <int>
                                       96/8096 1.52e-9 3.02e-7 2.60e-7 C1QA/~
##
    1 hsa05150 Staphylococ~ 12/99
    2 hsa04940 Type I diab~ 8/99
                                       43/8096 3.83e-8 3.81e-6 3.28e-6 CPE/H~
##
    3 hsa05330 Allograft r~ 7/99
                                       38/8096 3.05e-7 1.91e-5 1.65e-5 HLA-C~
##
##
    4 hsa04612 Antigen pro~ 9/99
                                       78/8096 3.84e-7 1.91e-5 1.65e-5 CD74/~
    5 hsa05332 Graft-versu~ 7/99
                                       42/8096 6.27e-7 2.50e-5 2.15e-5 HLA-C~
##
##
    6 hsa04610 Complement ~ 9/99
                                       85/8096 8.09e-7 2.68e-5 2.31e-5 C1QA/~
    7 hsa05320 Autoimmune ~ 7/99
                                      53/8096 3.21e-6 9.13e-5 7.87e-5 HLA-C~
                                                                                 10
##
                                       60/8096 7.51e-6 1.87e-4 1.61e-4 HLA-C~
##
    8 hsa05416 Viral myoca~ 7/99
                                                                                 11
                                      149/80~ 1.26e-5 2.80e-4 2.41e-4 HLA-C~
    9 hsa04514 Cell adhesi~ 10/99
                                                                                 12
## 10 hsa04145 Phagosome
                            10/99
                                       152/80~ 1.51e-5 3.00e-4 2.59e-4 MARCO~
                                                                                 13
      10
## # ... with 189 more rows
                                                                                 14
```

```
result.enrichGO.de_identO <- enrichGO(gene = de_entrez.identO, OrgDb = "org.Hs
    .eg.db", ont = "BP")
result.enrichGO.de_identO <- replaceEntrezID(result.enrichGO.de_identO) 2
result.enrichGO.de_identO@result 3</pre>
```

```
## # A tibble: 2,919 x 9
##
             Description GeneRatio BgRatio
                                              pvalue p.adjust
                                                                 qvalue geneID
   Count
                                                                                 3
##
      <chr> <chr>
                         <chr>
                                    <chr>
                                               <dbl>
                                                        <dbl>
                                                                  <dbl> <chr>
   <int.>
##
    1 GO:00~ neutrophil~ 29/166
                                    487/18~ 2.98e-16 5.11e-13 3.77e-13 CAMP/M~
    2 GO:00~ neutrophil~ 29/166
                                    490/18~ 3.50e-16 5.11e-13 3.77e-13 CAMP/M~
                                    91/188~ 4.52e-10 4.40e- 7 3.24e- 7 PPARG/~
    3 GO:00~ interferon~ 11/166
    4 GO:00~ lipid loca~ 20/166
                                    440/18~ 2.03e- 9 1.48e- 6 1.09e- 6 PPARG/~
##
##
    5 GO:00~ reactive o~ 16/166
                                    288/18~ 5.55e- 9 3.24e- 6 2.39e- 6 AIF1/C~
```

```
6 GO:00~ response t~ 13/166
                                   202/18~ 2.84e- 8 1.38e- 5 1.02e- 5 PPARG/~
      13
   7 GO:19~ regulation~ 12/166
                                   171/18~ 3.80e- 8 1.58e- 5 1.17e- 5 PPARG/~
##
                                                                                10
   8 GO:00~ lipid tran~ 17/166
                                   393/18~ 7.28e- 8 2.45e- 5 1.81e- 5 PPARG/~
                                                                                11
      17
   9 GO:00~ cellular r~ 12/166
                                   182/18~ 7.56e- 8 2.45e- 5 1.81e- 5 PPARG/~
## 10 GO:00~ regulation~ 13/166
                                   240/18~ 2.13e- 7 6.18e- 5 4.55e- 5 IGFBP2~
                                                                                13
      13
## # ... with 2,909 more rows
                                                                                14
```

3.2 Cluster 2 DE analysis and GO enrichment

The Cluster 2 mentioned in manuscript refers to cluster 1 in the codes following.

Calculate the DE genes.

Save data

```
write.csv(de_ident1, file = "./Results_lists/DE_cluster2.csv")
write.csv(de_ident1_vs0, file = "./Results_lists/DE_cluster2_vs_cluster1.csv")
```

KEGG/GO enrichment:

```
de_entrez.ident1 <- bitr( geneID = rownames(de_ident1), fromType = "SYMBOL",
    toType = "ENTREZID", OrgDb = "org.Hs.eg.db", drop = TRUE ) $ ENTREZID
result.enrichKEGG.de_ident1 <- enrichKEGG(de_entrez.ident1, organism = "hsa",
    keyType = "ncbi-geneid")
result.enrichKEGG.de_ident1 <- replaceEntrezID(result.enrichKEGG.de_ident1)</pre>
```

```
de_entrez.ident1_vs0 <- bitr( geneID = rownames(de_ident1_vs0), fromType = "
    SYMBOL", toType = "ENTREZID", OrgDb = "org.Hs.eg.db", drop = TRUE ) $
    ENTREZID

result.enrichKEGG.de_ident1_vs0 <- enrichKEGG(de_entrez.ident1_vs0, organism = 2
    "hsa", keyType = "ncbi-geneid")

result.enrichKEGG.de_ident1_vs0 <- replaceEntrezID(result.enrichKEGG.de_ident1 3
    _vs0)</pre>
```

```
result.enrichKEGG.de_ident1@result
```

```
## # A tibble: 200 x 9
##
      TD
               Description
                             GeneRatio BgRatio pvalue p.adjust qvalue geneID
   Count
##
      <chr>
                             <chr>>
                                        <chr>
                                                   <dbl>
                                                            <dbl>
                                                                    <dbl> <chr>
               <chr>>
   <int>
##
    1 hsa04142 Lysosome
                             11/93
                                        128/80~ 2.04e-7 4.07e-5 3.75e-5 DNASE~
      11
                                        152/80~ 5.74e-5 5.74e-3 5.28e-3 NCF1/~
##
    2 hsa04145 Phagosome
                             9/93
    3 \text{ hsa} 04966 \text{ Collecting} \sim 4/93
                                        27/8096 2.34e-4 1.19e-2 1.09e-2 ATP6V~
##
    4 hsa05110 Vibrio chol~ 5/93
                                        50/8096 2.53e-4 1.19e-2 1.09e-2 ATP6A~
##
    5 hsa05171 Coronavirus~ 10/93
                                        232/80~ 3.09e-4 1.19e-2 1.09e-2 CYBB/~
##
    6 hsa00030 Pentose pho~ 4/93
                                        30/8096 3.56e-4 1.19e-2 1.09e-2 PGD/F~
##
    7 hsa03010 Ribosome
                             8/93
                                        158/80~ 4.40e-4 1.26e-2 1.16e-2 RPL39~
                                                                                   10
##
##
    8 hsa05323 Rheumatoid ~ 6/93
                                        93/8096 6.67e-4 1.67e-2 1.54e-2 CSF1/~
                                                                                   11
    9 hsa05120 Epithelial ~ 5/93
                                        70/8096 1.21e-3 2.68e-2 2.47e-2 ATP6A~
##
## 10 hsa04979 Cholesterol~ 4/93
                                        50/8096 2.51e-3 5.02e-2 4.62e-2 LIPA/~
                                                                                   13
## # ... with 190 more rows
                                                                                   14
```

```
write.csv(result.enrichKEGG.de_ident1, file = "./Results_lists/enrichKEGG_DE_
    cluster2.csv")
write.csv(result.enrichKEGG.de_ident1_vs0, file = "./Results_lists/enrichKEGG_
DE_cluster2_vs_cluster1.csv")
```

GO analysis on the upregulated genes in Cluster 2.

```
result.enrichGO.de_ident1 <- enrichGO(gene = de_entrez.ident1, OrgDb = "org.Hs
    .eg.db", ont = "BP")
result.enrichGO.de_ident1 <- replaceEntrezID(result.enrichGO.de_ident1)
result.enrichGO.de_ident1@result</pre>
```

```
## # A tibble: 2,543 x 9
                          GeneRatio BgRatio pvalue p.adjust qvalue geneID
##
      ID
             Description
   Count
                           <chr>
                                     <chr>
                                               <dbl>
                                                         <dbl>
                                                                 <dbl> <chr>
                                                                                 3
##
      <chr>
             <chr>
   <int>
                                     343/18~ 1.64e-8 4.17e-5 3.51e-5 FABP3/H~
##
    1 GO:00~ lipid catab~ 15/131
      15
    2 GO:00~ cell redox ~ 7/131
                                     59/188~ 1.67e-7 1.52e-4 1.28e-4 PRDX1/N~
##
    3 GO:00~ neutrophil ~ 16/131
##
                                     487/18~ 2.75e-7 1.52e-4 1.28e-4 FGR/HEX~
##
    4 GO:00~ bone resorp~ 7/131
                                     64/188~ 2.96e-7 1.52e-4 1.28e-4 ATP6AP1~
```

```
5 GO:00~ neutrophil ~ 16/131
                                    490/18~ 2.99e-7 1.52e-4 1.28e-4 FGR/HEX~
    6 GO:00~ tissue remo~ 10/131
                                    178/18~ 4.68e-7 1.98e-4 1.67e-4 ATP6AP1~
##
                                    192/18~ 9.36e-7 3.40e-4 2.86e-4 RPL39/P~
    7 GO:00~ translation~ 10/131
                                                                                10
    8 GO:00~ nuclear-tra~ 8/131
                                    120/18~ 1.91e-6 5.97e-4 5.02e-4 RPL39/P~
                                                                                11
##
    9 GO:00~ tissue home~ 11/131
                                    261/18~ 2.11e-6 5.97e-4 5.02e-4 PRDX1/C~
                                                                                12
## 10 GO:00~ positive re~ 6/131
                                    58/188~ 3.01e-6 7.64e-4 6.42e-4 CSF1/TR~
                                                                                13
## # ... with 2,533 more rows
                                                                                14
```

GO analysis on the regulated genes in cluster 2 vs cluster 1.

```
result.enrichGO.de_ident1_vs0 <- enrichGO(gene = de_entrez.ident1_vs0, OrgDb = 1
    "org.Hs.eg.db", ont = "BP")
result.enrichGO.de_ident1_vs0 <- replaceEntrezID(result.enrichGO.de_ident1_vs0 2
    )
result.enrichGO.de_ident1_vs0@result</pre>
```

```
## # A tibble: 3,414 x 9
             Description GeneRatio BgRatio
                                               pvalue p.adjust qvalue geneID
##
      TD
   Count
##
      <chr>
             <chr>
                          <chr>>
                                     <chr>
                                                <dbl>
                                                         <dbl>
                                                                  <dbl> <chr>
                                                                                 3
   <int>
##
   1 GO:00~ neutrophil ~ 29/230
                                     487/18~ 1.76e-12 3.50e-9 2.84e-9 FGR/HE~
                                     490/18~ 2.05e-12 3.50e-9 2.84e-9 FGR/HE~
##
    2 GO:00~ neutrophil ~ 29/230
##
    3 GO:00~ translation~ 16/230
                                     192/18~ 1.82e- 9 2.07e-6 1.68e-6 RPS8/R~
    4 GO:00~ cotranslati~ 12/230
                                     109/18~ 8.95e- 9
                                                      6.65e-6 5.40e-6 RPS8/R~
##
                                     36/188~ 9.74e- 9
    5 GO:00~ transferrin~ 8/230
                                                      6.65e-6 5.40e-6 ATP6V0~
                                                                                 8
##
##
    6 GO:00~ nuclear-tra~ 12/230
                                     120/18~ 2.68e- 8 1.08e-5 8.74e-6 RPS8/R~
                                                                                 9
                                                                                 10
    7 GO:00~ protein tar~ 12/230
                                     120/18~ 2.68e- 8 1.08e-5 8.74e-6 RPS8/R~
##
                                     58/188~ 3.22e- 8 1.08e-5 8.74e-6 ATP6V0~
##
    8 GO:00~ pH reduction 9/230
                                                                                 11
                                                                                 12
    9 GO:00~ phagosome a~ 7/230
                                     28/188~ 3.48e- 8 1.08e-5 8.74e-6 ATP6V0~
## 10 GO:00~ cell redox ~ 9/230
                                     59/188~ 3.76e- 8 1.08e-5 8.74e-6 PRDX1/~
                                                                                 13
## # ... with 3,404 more rows
                                                                                 14
```

```
write.csv(result.enrichGO.de_ident1, file = "./Results_lists/enrichGOBP_DE_
    cluster2.csv")
write.csv(result.enrichGO.de_ident1_vs0, file = "./Results_lists/enrichGOBP_DE 2
    _cluster2_vs_cluster1.csv")
```

3.3 Cluster 3 DE analysis

The Cluster 3 mentioned in manuscript refers to cluster 2 in the codes following.

Calculate the DE genes.

```
de_ident2 <- FindMarkers(results, ident.1 = "2", only.pos = TRUE, logfc.
    threshold = 0.2)
    nrow(de_ident2)</pre>
```

```
## [1] 660
```

Save data

```
write.csv(de_ident2, file = "./Results_lists/DE_cluster3.csv")
```

3.4 Cluster 4 DE analysis

The Cluster 4 mentioned in manuscript refers to cluster 3 in the codes following.

Calculate the DE genes.

```
de_ident3 <- FindMarkers(results, ident.1 = "3", only.pos = TRUE, logfc.
    threshold = 0.2)
head(de_ident3)</pre>
```

```
## # A tibble: 6 x 5
##
     p_val avg_log2FC pct.1 pct.2 p_val_adj
                                                                                       2
##
     <dbl>
                 <dbl> <dbl> <dbl>
                                         <dbl>
                                                                                       3
## 1
                 0.244 0.274 0.028
                                              0
                                                                                       4
         0
## 2
         0
                 0.208 0.175 0.002
                                              0
                                                                                       5
                                                                                       6
## 3
                 2.34 0.983 0.684
                                              0
         0
                 0.770 0.473 0.014
                                              0
## 4
         0
                                              0
                                                                                      8
## 5
         0
                 0.319 0.225 0.009
                 0.885 0.352 0.003
                                              0
```

Save data

```
write.csv(de_ident3, file = "./Results_lists/DE_cluster4.csv")
```

3.5 GO/KEGG enrichment and DE analysis in subpopulations of Cluster 3 after re-clustering

Load Cluster 3 object:

```
results.c2 <- readRDS("../4-Functional characterization of clustered populations/cluster2_clustered.seuratObject.Rds")
```

REMIND: the subpopulations cluster 1-4 (after reclustering the cluster 3) in the manuscript refer to the cluster 0-3 in the following codes.

Calculate the DE genes.

DE genes for each subpopulations:

```
de_subpop0.c2 <- FindMarkers(object = results.c2, ident.1 = "0", only.pos =
    TRUE, logfc.threshold = 0.2, verbose = FALSE)

de_subpop1.c2 <- FindMarkers(object = results.c2, ident.1 = "1", only.pos =
    TRUE, logfc.threshold = 0.2, verbose = FALSE)

de_subpop2.c2 <- FindMarkers(object = results.c2, ident.1 = "2", only.pos =
    TRUE, logfc.threshold = 0.2, verbose = FALSE)

de_subpop3.c2 <- FindMarkers(object = results.c2, ident.1 = "3", only.pos =
    TRUE, logfc.threshold = 0.2, verbose = FALSE)</pre>
```

Save data

```
write.csv(de_subpop0.c2, file = "./Results_lists/DE_subpop1_c3.csv")
write.csv(de_subpop1.c2, file = "./Results_lists/DE_subpop2_c3.csv")
write.csv(de_subpop2.c2, file = "./Results_lists/DE_subpop3_c3.csv")
write.csv(de_subpop3.c2, file = "./Results_lists/DE_subpop4_c3.csv")
4
```

Since the subpopulations 2 and 3 are very similar. We would like also to have the DE genes in both subpopulations 2 & 3 comparing to subpopulation 1 (we excluded cluster 4 because it's high similarity to DCs).

```
de_subpop1_2.c2 <- FindMarkers(object = results.c2, ident.1 = c("1", "2"),
   ident.2 = ("0"), only.pos = TRUE, logfc.threshold = 0.2, verbose = FALSE)
write.csv(de_subpop1_2.c2, file = "./Results_lists/DE_subpop2_and_3_of_c3.csv" 2
)</pre>
```

KEGG enrichment:

```
de_entrez.subpop0.c2 <- bitr( geneID = rownames(de_subpop0.c2), fromType = "</pre>
   SYMBOL", toType = "ENTREZID", OrgDb = "org.Hs.eg.db", drop = TRUE ) $
   ENTREZID
de_entrez.subpop1.c2 <- bitr( geneID = rownames(de_subpop1.c2), fromType = "</pre>
   SYMBOL", toType = "ENTREZID", OrgDb = "org.Hs.eg.db", drop = TRUE ) $
   ENTREZID
de entrez.subpop2.c2 <- bitr( geneID = rownames(de subpop2.c2), fromType = "</pre>
   SYMBOL", toType = "ENTREZID", OrgDb = "org.Hs.eg.db", drop = TRUE ) $
de_entrez.subpop3.c2 <- bitr( geneID = rownames(de_subpop3.c2), fromType = "</pre>
   SYMBOL", toType = "ENTREZID", OrgDb = "org.Hs.eg.db", drop = TRUE ) $
   ENTREZID
de_entrez.subpop1_2.c2 <- bitr( geneID = rownames(de_subpop1_2.c2), fromType =</pre>
    "SYMBOL", toType = "ENTREZID", OrgDb = "org.Hs.eg.db", drop = TRUE ) $
   ENTREZID
result.enrichKEGG.de_entrez.subpop0.c2 <- enrichKEGG(de_entrez.subpop0.c2,
   organism = "hsa", keyType = "ncbi-geneid")
result.enrichKEGG.de_entrez.subpop1.c2 <- enrichKEGG(de_entrez.subpop1.c2,</pre>
                                                                                  8
   organism = "hsa", keyType = "ncbi-geneid")
result.enrichKEGG.de_entrez.subpop2.c2 <- enrichKEGG(de_entrez.subpop2.c2,
                                                                                  9
   organism = "hsa", keyType = "ncbi-geneid")
result.enrichKEGG.de_entrez.subpop3.c2 <- enrichKEGG(de_entrez.subpop3.c2,</pre>
                                                                                  10
   organism = "hsa", keyType = "ncbi-geneid")
result.enrichKEGG.de_entrez.subpop1_2.c2 <- enrichKEGG(de_entrez.subpop1_2.c2,
                                                                                  11
    organism = "hsa", keyType = "ncbi-geneid")
                                                                                  12
```

```
result.enrichGO.de_entrez.subpop0.c2 <- enrichGO(gene = de_entrez.subpop0.c2,
   OrgDb = "org.Hs.eg.db", ont = "BP")
result.enrichGO.de_entrez.subpop1.c2 <- enrichGO(gene = de_entrez.subpop1.c2,
                                                                                14
   OrgDb = "org.Hs.eg.db", ont = "BP")
result.enrichGO.de_entrez.subpop2.c2 <- enrichGO(gene = de_entrez.subpop2.c2,
                                                                                15
   OrgDb = "org.Hs.eg.db", ont = "BP")
result.enrichGO.de_entrez.subpop3.c2 <- enrichGO(gene = de_entrez.subpop3.c2,</pre>
   OrgDb = "org.Hs.eg.db", ont = "BP")
result.enrichGO.de_entrez.subpop1_2.c2 <- enrichGO(gene = de_entrez.subpop1_2.
   c2, OrgDb = "org.Hs.eg.db", ont = "BP")
                                                                                 18
result.enrichKEGG.de_entrez.subpop0.c2 <- replaceEntrezID(result.enrichKEGG.de | 19
   _entrez.subpop0.c2)
result.enrichKEGG.de_entrez.subpop1.c2 <- replaceEntrezID(result.enrichKEGG.de | 20
   _entrez.subpop1.c2)
result.enrichKEGG.de_entrez.subpop2.c2 <- replaceEntrezID(result.enrichKEGG.de
   _entrez.subpop2.c2)
result.enrichKEGG.de_entrez.subpop3.c2 <- replaceEntrezID(result.enrichKEGG.de
   _entrez.subpop3.c2)
result.enrichKEGG.de_entrez.subpop1_2.c2 <- replaceEntrezID(result.enrichKEGG.
   de_entrez.subpop1_2.c2)
                                                                                24
                                                                                25
result.enrichGO.de_entrez.subpopO.c2 <- replaceEntrezID(result.enrichGO.de_
   entrez.subpop0.c2)
                                                                                26
result.enrichGO.de_entrez.subpop1.c2 <- replaceEntrezID(result.enrichGO.de_
   entrez.subpop1.c2)
result.enrichGO.de_entrez.subpop2.c2 <- replaceEntrezID(result.enrichGO.de_
                                                                                27
   entrez.subpop2.c2)
result.enrichGO.de_entrez.subpop3.c2 <- replaceEntrezID(result.enrichGO.de_
                                                                                28
   entrez.subpop3.c2)
                                                                                29
result.enrichGO.de_entrez.subpop1_2.c2 <- replaceEntrezID(result.enrichGO.de_
   entrez.subpop1_2.c2)
write.csv(result.enrichKEGG.de_entrez.subpop0.c2, file = "./Results_lists/
   enrichKEGG_DE_subpop1_c3.csv")
write.csv(result.enrichKEGG.de_entrez.subpop1.c2, file = "./Results_lists/
                                                                                2
   enrichKEGG_DE_subpop2_c3.csv")
write.csv(result.enrichKEGG.de_entrez.subpop2.c2, file = "./Results_lists/
                                                                                3
   enrichKEGG_DE_subpop3_c3.csv")
write.csv(result.enrichKEGG.de_entrez.subpop3.c2, file = "./Results_lists/
                                                                                 4
   enrichKEGG_DE_subpop4_c3.csv")
write.csv(result.enrichKEGG.de_entrez.subpop1_2.c2, file = "./Results_lists/
                                                                                5
   enrichKEGG_DE_subpop2_3.c3.csv")
```

```
write.csv(result.enrichGO.de_entrez.subpop0.c2, file = "./Results_lists/
    enrichGOBP_DE_subpop1_c3.csv")
write.csv(result.enrichGO.de_entrez.subpop1.c2, file = "./Results_lists/
    enrichGOBP_DE_subpop2_c3.csv")
write.csv(result.enrichGO.de_entrez.subpop2.c2, file = "./Results_lists/
    enrichGOBP_DE_subpop3_c3.csv")
write.csv(result.enrichGO.de_entrez.subpop3.c2, file = "./Results_lists/
    enrichGOBP_DE_subpop4_c3.csv")
write.csv(result.enrichGO.de_entrez.subpop1_2.c2, file = "./Results_lists/
    enrichGOBP_DE_subpop4_c3.csv")
write.csv(result.enrichGO.de_entrez.subpop1_2.c2, file = "./Results_lists/
    enrichGOBP_DE_subpop2_3_C3.csv")
```

6 7

8

10

11

Head 10 lines in GO:

For subpop 1:

```
head(result.enrichGO.de_entrez.subpopO.c2@result, 10)
```

```
# A tibble: 10 x 9
##
                                                                                     2
      ID
             Description GeneRatio BgRatio
                                                pvalue p.adjust
                                                                    qvalue geneID
   Count
##
      <chr>
             <chr>>
                           <chr>>
                                     <chr>
                                                 <dbl>
                                                           <dbl>
                                                                     <dbl> <chr>
                                                                                     3
   <int>
                                     487/18~ 3.27e-36 1.38e-32 1.07e-32 CYBB/M~
##
   1 GO:00~ neutrophil~ 91/749
                                     490/18~ 5.45e-36 1.38e-32 1.07e-32 CYBB/M~
    2 GO:00~ neutrophil~ 91/749
##
##
    3 \text{ GO:} 00^{\circ} \text{ oxidative } \sim 37/749
                                     149/18~ 1.25e-19 2.12e-16 1.63e-16 RHOA/U~
                                     311/18~ 2.51e-15 3.19e-12 2.45e-12 FBP1/T~
    4 GO:00~ ATP metabo~ 47/749
##
                                     382/18~ 6.65e-15 6.76e-12 5.21e-12 ITGAM/~
##
    5 GO:00~ phagocytos~ 52/749
    6 GO:00~ electron t~ 33/749
                                     184/18~ 2.93e-13 2.48e-10 1.91e-10 CYBB/S~
##
    7 GO:00~ mitochondr~ 23/749
                                     97/188~ 2.97e-12 2.16e- 9 1.66e- 9 UQCRH/~
                                                                                     10
##
    8 GO:00~ ATP synthe~ 23/749
                                     98/188~ 3.74e-12 2.38e- 9 1.83e- 9 UQCRH/~
                                                                                     11
                                     360/18~ 9.28e-12 5.24e- 9 4.04e- 9 RGCC/S~
    9 GO:00~ regulation~ 45/749
                                                                                     12
## 10 GO:00~ cellular r~ 31/749
                                     187/18~ 1.31e-11 6.66e- 9 5.13e- 9 UQCRH/~
                                                                                     13
      31
```

For subpop 2:

```
head(result.enrichGO.de_entrez.subpop1.c2@result, 10)
```

```
## # A tibble: 10 x 9
##
      ΤD
             Description GeneRatio BgRatio
                                              pvalue p.adjust
                                                                  qvalue geneID
   Count
##
      <chr>
             <chr>
                          <chr>>
                                    <chr>>
                                                <dbl>
                                                         <dbl>
                                                                   <dbl> <chr>
                                                                                  3
   <int>
##
   1 GO:00~ neutrophil~ 39/299
                                    487/18~ 7.13e-17 1.58e-13 1.21e-13 MME/FT~
    2 GO:00~ neutrophil~ 39/299
                                    490/18~ 8.78e-17 1.58e-13 1.21e-13 MME/FT~
##
##
    3 GO:00~ cellular r~ 21/299
                                    182/18~ 1.35e-12 1.63e- 9 1.25e- 9 CCL18/~
                                    202/18~ 1.03e-11 9.28e- 9 7.11e- 9 CCL18/~
    4 GO:00~ response t~ 21/299
    5 GO:00~ neutrophil~ 14/299
                                    103/18~ 9.05e-10 6.53e- 7 5.01e- 7 CCL18/~
    6 GO:00~ response t~ 20/299
                                    250/18~ 3.43e- 9 1.81e- 6 1.39e- 6 NUPR1/~
      20
```

For subpop 3:

```
head(result.enrichGO.de_entrez.subpop2.c2@result, 10)
```

```
## # A tibble: 10 x 9
             Description GeneRatio BgRatio
                                               pvalue p.adjust
##
      ID
                                                                  qvalue geneID
   Count
                                                <dbl>
                                                                   <dbl> <chr>
##
      <chr>
             <chr>
                          <chr>
                                     <chr>>
                                                          <dbl>
                                                                                   3
   <int>
##
   1 GO:00~ neutrophil~ 70/526
                                     487/18~ 3.13e-30 1.07e-26 7.63e-27 CTSB/P~
                                     490/18~ 4.62e-30 1.07e-26 7.63e-27 CTSB/P~
##
    2 GO:00~ neutrophil~ 70/526
                                     447/18~ 2.37e-11 3.67e- 8 2.61e- 8 HIF1A/~
##
    3 G0:00^{\circ} positive r^{\circ} 41/526
    4 GO:00~ regulation~ 27/526
                                     212/18~ 4.77e-11 5.54e- 8 3.94e- 8 LGMN/A~
##
                                     94/188~ 9.84e-11 9.14e- 8 6.51e- 8 SPP1/A~
    5 GO:19~ positive r~ 18/526
##
##
    6 GO:00~ maintenanc~ 33/526
                                     324/18~ 1.49e-10 1.10e- 7 7.85e- 8 ABCA1/~
                                     208/18~ 1.66e-10 1.10e- 7 7.85e- 8 ABCA1/~
    7 GO:00~ cellular r~ 26/526
                                                                                   10
##
                                     334/18~ 3.31e-10 1.92e- 7 1.37e- 7 ABCA1/~
##
    8 GO:00~ response t~ 33/526
                                                                                   11
    9 GO:00~ cellular r~ 26/526
                                     222/18~ 6.99e-10 3.25e- 7 2.31e- 7 ABCA1/~
                                                                                   12
## 10 GO:00~ myeloid le~ 26/526
                                     222/18~ 6.99e-10 3.25e- 7 2.31e- 7 LGMN/A~
                                                                                   13
```

For subpop 4:

```
head(result.enrichGO.de_entrez.subpop3.c2@result, 10)
```

```
## # A tibble: 10 x 9
##
                                              pvalue p.adjust
      ID
             Description GeneRatio BgRatio
                                                                 qvalue geneID
   Count
##
      <chr>
             <chr>
                          <chr>
                                    <chr>
                                               <dbl>
                                                         <dbl>
                                                                  <dbl> <chr>
                                                                                  3
   <int>
   1 GO:00~ SRP-depend~ 79/926
                                    105/18~ 1.35e-81 6.99e-78 5.90e-78 RPL18/~
    2 GO:00~ cotranslat~ 80/926
                                    109/18~ 4.03e-81 1.05e-77 8.85e-78 RPL18/~
    3 GO:00~ protein ta~ 81/926
                                    120/18~ 2.86e-77 4.95e-74 4.18e-74 RPL18/~
      81
```

```
4 GO:00~ establishm~ 81/926
                                    124/18~ 1.81e-75 2.35e-72 1.99e-72 RPL18/~
##
    5 GO:00~ translatio~ 97/926
                                    192/18~ 3.70e-75 3.84e-72 3.25e-72 RPL18/~
##
                                    120/18~ 5.07e-74 4.38e-71 3.70e-71 RPL18/~
    6 GO:00~ nuclear-tr~ 79/926
   7 GO:00~ protein lo~ 84/926
                                    152/18~ 1.97e-69 1.46e-66 1.23e-66 RPL18/~
                                                                                 10
##
    8 GO:00~ nuclear-tr~ 90/926
                                    210/18~ 1.02e-61 6.60e-59 5.58e-59 RPL18/~
                                                                                 11
    9 GO:00~ viral gene~ 87/926
                                    195/18~ 1.69e-61 9.77e-59 8.25e-59 RPL18/~
                                                                                 12
## 10 GO:00~ viral tran~ 81/926
                                    178/18~ 4.02e-58 2.09e-55 1.76e-55 RPL18/~
                                                                                 13
      81
```

For subpop 2,3:

```
head(result.enrichGO.de_entrez.subpop1_2.c2@result, 10)
```

```
## # A tibble: 10 x 9
##
      ID
             Description GeneRatio BgRatio
                                              pvalue p.adjust
                                                                 qvalue geneID
##
      <chr>
            <chr>
                          <chr>
                                    <chr>
                                               <dbl>
                                                         <dbl>
                                                                  <dbl> <chr>
                                                                                  3
   <int>
                                    487/18~ 8.26e-24 2.24e-20 1.57e-20 PSAP/C~
##
    1 GO:00~ neutrophil~ 51/353
                                    490/18~ 1.09e-23 2.24e-20 1.57e-20 PSAP/C~
##
    2 GO:00~ neutrophil~ 51/353
                                    94/188~ 1.27e-13 1.73e-10 1.22e-10 PLTP/A~
##
    3 GO:19~ positive r~ 18/353
                                    425/18~ 2.48e-11 2.54e- 8 1.78e- 8 MGLL/N~
    4 GO:00~ regulation~ 32/353
##
                                    202/18~ 3.36e-11 2.75e- 8 1.93e- 8 NR1H3/~
##
    5 GO:00~ response t~ 22/353
##
    6 GO:00~ regulation~ 33/353
                                    470/18~ 7.61e-11 4.47e- 8 3.14e- 8 SERPIN~
                                    393/18~ 7.65e-11 4.47e- 8 3.14e- 8 PSAP/P~
                                                                                  10
    7 GO:00~ lipid tran~ 30/353
    8 GO:00~ cholestero~ 13/353
                                    64/188~ 1.54e-10 7.86e- 8 5.52e- 8 PLTP/A~
                                                                                  11
    9 GO:00~ cellular r~ 20/353
                                    182/18~ 2.28e-10 9.99e- 8 7.02e- 8 NR1H3/~
                                                                                  12
                                    246/18~ 2.64e-10 9.99e- 8 7.02e- 8 ABCA1/~
                                                                                  13
## 10 GO:00~ cellular r~ 23/353
      23
```

4 Session information

R session:

```
sessionInfo()
```

```
## R version 4.0.3 (2020-10-10)
## Platform: x86_64-pc-linux-gnu (64-bit)
```

```
## Running under: Ubuntu 20.04.3 LTS
##
                                                                                     4
                                                                                     5
## Matrix products: default
           /usr/lib/x86_64-linux-gnu/openblas-pthread/libblas.so.3
                                                                                     6
                                                                                     7
## LAPACK: /usr/lib/x86_64-linux-gnu/openblas-pthread/liblapack.so.3
                                                                                     8
##
                                                                                     9
## locale:
    [1] LC_CTYPE=en_US.UTF-8
                                                                                     10
##
                                     LC NUMERIC=C
##
    [3] LC_TIME=en_GB.UTF-8
                                     LC_COLLATE=en_US.UTF-8
                                                                                     11
                                                                                     12
##
    [5] LC_MONETARY=en_GB.UTF-8
                                     LC_MESSAGES=en_US.UTF-8
    [7] LC_PAPER=en_GB.UTF-8
                                     LC_NAME=C
                                                                                     13
                                     LC_TELEPHONE=C
                                                                                     14
    [9] LC_ADDRESS=C
                                                                                     15
  [11] LC_MEASUREMENT=en_GB.UTF-8 LC_IDENTIFICATION=C
                                                                                     16
##
                                                                                     17
## attached base packages:
## [1] parallel
                  stats4
                                                                        datasets
                                                                                     18
                             stats
                                        graphics
                                                  grDevices utils
                                                                                     19
  [8] methods
                  base
                                                                                     20
##
                                                                                     21
## other attached packages:
                                                                                     22
    [1] org.Hs.eg.db 3.12.0
                                 AnnotationDbi 1.52.0
                                                          IRanges 2.24.1
##
    [4] S4Vectors_0.28.1
                                 Biobase_2.50.0
                                                          BiocGenerics_0.36.1
                                                                                     23
                                                                                     24
    [7] clusterProfiler_3.18.1 dplyr_1.0.7
                                                          ggplot2_3.3.5
                                                                                     25
                                 Seurat_4.0.3
## [10] SeuratObject 4.0.2
                                                                                     26
                                                                                     27
  loaded via a namespace (and not attached):
##
     [1] shadowtext 0.0.8
                                 fastmatch 1.1-3
                                                         plyr_1.8.6
                                                                                     28
##
     [4] igraph_1.2.6
                                 lazyeval_0.2.2
                                                                                     29
                                                         splines_4.0.3
                                                                                     30
##
     [7] BiocParallel_1.24.1
                                 listenv_0.8.0
                                                         scattermore_0.7
                                                                                     31
##
    [10] digest_0.6.27
                                 htmltools_0.5.1.1
                                                         GOSemSim_2.16.1
                                                                                     32
##
    [13] viridis_0.6.1
                                 GO.db_3.12.1
                                                         fansi_0.5.0
                                                                                     33
##
    [16] magrittr_2.0.1
                                 memoise_2.0.0
                                                         tensor_1.5
##
    [19] cluster_2.1.0
                                 ROCR_1.0-11
                                                         limma_3.46.0
                                                                                     34
                                                                                     35
##
    [22] graphlayouts_0.7.1
                                 globals_0.14.0
                                                         matrixStats_0.60.0
                                                                                     36
##
    [25] spatstat.sparse_2.0-0 enrichplot_1.10.2
                                                         colorspace_2.0-2
                                                                                     37
##
    [28] blob 1.2.2
                                 ggrepel_0.9.1
                                                         xfun 0.24
##
                                                                                     38
    [31] crayon_1.4.1
                                 jsonlite_1.7.2
                                                         scatterpie_0.1.6
##
    [34] spatstat.data_2.1-0
                                 survival 3.2-7
                                                         zoo 1.8-9
                                                                                     39
##
    [37] glue_1.4.2
                                                                                     40
                                 polyclip_1.10-0
                                                         gtable_0.3.0
                                                                                     41
##
    [40] leiden_0.3.9
                                 future.apply_1.7.0
                                                         abind_1.4-5
                                                                                     42
##
    [43] scales_1.1.1
                                 DOSE_3.16.0
                                                         DBI_1.1.1
                                                                                     43
    [46] miniUI 0.1.1.1
                                 Rcpp_1.0.7
                                                         viridisLite 0.4.0
##
    [49] xtable 1.8-4
                                 reticulate 1.20
                                                                                     44
                                                         spatstat.core_2.3-0
                                                                                     45
##
    [52] bit 4.0.4
                                 htmlwidgets_1.5.3
                                                         httr 1.4.2
##
                                                                                     46
    [55] fgsea_1.16.0
                                 RColorBrewer_1.1-2
                                                         ellipsis_0.3.2
                                                                                     47
##
    [58] ica_1.0-2
                                 farver_2.1.0
                                                         pkgconfig_2.0.3
##
    [61] uwot_0.1.10.9000
                                                                                     48
                                 deldir_0.2-10
                                                         utf8_1.2.2
                                                                                     49
##
    [64] tidyselect_1.1.1
                                 rlang_0.4.11
                                                         reshape2_1.4.4
                                                                                     50
##
    [67] later_1.2.0
                                 munsell_0.5.0
                                                         tools_4.0.3
##
    [70] cachem_1.0.5
                                 downloader_0.4
                                                         cli_3.0.1
                                                                                     51
                                                                                     52
##
    [73] generics_0.1.0
                                 RSQLite_2.2.7
                                                         ggridges_0.5.3
##
                                                                                     53
    [76] evaluate_0.14
                                 stringr_1.4.0
                                                         fastmap_1.1.0
                                                                                     54
##
    [79] yaml 2.2.1
                                 goftest_1.2-2
                                                         knitr 1.33
##
    [82] bit64_4.0.5
                                 fitdistrplus_1.1-5
                                                                                     55
                                                         tidygraph_1.2.0
                                                                                     56
    [85] purrr_0.3.4
                                 RANN_2.6.1
                                                         ggraph_2.0.5
```

```
##
    [88] pbapply_1.4-3
                                 future_1.21.0
                                                        nlme 3.1-152
                                                                                    57
                                                                                    58
##
    [91] mime_0.11
                                 DO.db_2.9
                                                        compiler_4.0.3
                                                                                    59
##
    [94] rstudioapi 0.13
                                 plotly_4.9.4.1
                                                        png_0.1-7
    [97] spatstat.utils_2.2-0
                                 tibble_3.1.3
                                                        tweenr_1.0.2
                                                                                    60
                                                                                    61
   [100] stringi_1.7.3
                                 lattice_0.20-41
                                                        Matrix_1.3-4
  [103] vctrs 0.3.8
                                 pillar 1.6.2
                                                        lifecycle_1.0.0
                                                                                    62
  [106] BiocManager 1.30.16
                                 spatstat.geom_2.2-2
                                                        1mtest 0.9-38
                                                                                    63
  [109] RcppAnnoy_0.0.19
                                                                                    64
                                 data.table_1.14.0
                                                        cowplot_1.1.1
   [112]
         irlba_2.3.3
                                 httpuv_1.6.1
                                                        patchwork_1.1.1
                                                                                    65
  [115] qvalue_2.22.0
                                 R6_2.5.0
                                                        promises_1.2.0.1
                                                                                    66
  [118] KernSmooth_2.23-20
                                 gridExtra_2.3
                                                        parallelly_1.27.0
                                                                                    67
  [121] codetools_0.2-18
                                 MASS_7.3-53
                                                        assertthat_0.2.1
                                                                                    68
                                                        mgcv_1.8-33
   [124] withr_2.4.2
                                                                                    69
                                 sctransform_0.3.2
                                                                                    70
  [127]
        grid_4.0.3
                                 rpart_4.1-15
                                                        tidyr_1.1.3
## [130]
        rvcheck_0.1.8
                                 rmarkdown_2.9
                                                        Rtsne_0.15
                                                                                    71
                                                                                    72
   [133]
         ggforce_0.3.3
                                 shiny_1.6.0
```

References

- Hao Y, Hao S, Andersen-Nissen E, Mauck III WM, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zagar M, Hoffman P, Stoeckius M, Papalexi E, Mimitou EP, Jain J, Srivastava A, Stuart T, Fleming LB, Yeung B, Rogers AJ, McElrath JM, Blish CA, Gottardo R, Smibert P, Satija R. Integrated analysis of multimodal single-cell data. Cell [Internet] 2021; Available from: https://doi.org/10.1016/j.cell.2021.04.048.
- 2. Yu G, Wang L-G, Han Y, He Q-Y. cluster Profiler: An r package for comparing biological themes among gene clusters. *OMICS: A Journal of Integrative Biology* 2012; 16: 284–287.