PAPER TITLE TO BE DEFINED (in common.yaml)

13-Compare_with_Mafb-KO_IM

2022-01-16 22:15:49 +0100

Abstract

Lung interstitium macrophages (IMs) are non-alveolar resident tissue macrophages which contribute to the lung homeostasis. These cells were reported to be heterogeneous by our group and other teams, which contains two main distinct subpopulations: CD206+ IMs and CD206- IMs. However, the exact origin of IMs and the transcriptional programs that control IM differentiation remains unclear. In recent report, we analyzed the refilled IMs in the course of time after induced IM depletion with single-cell RNA sequencing (10X Genomics Chromium) and bulk RNA sequencing.

Contents

1 Description

2 Load packages and data

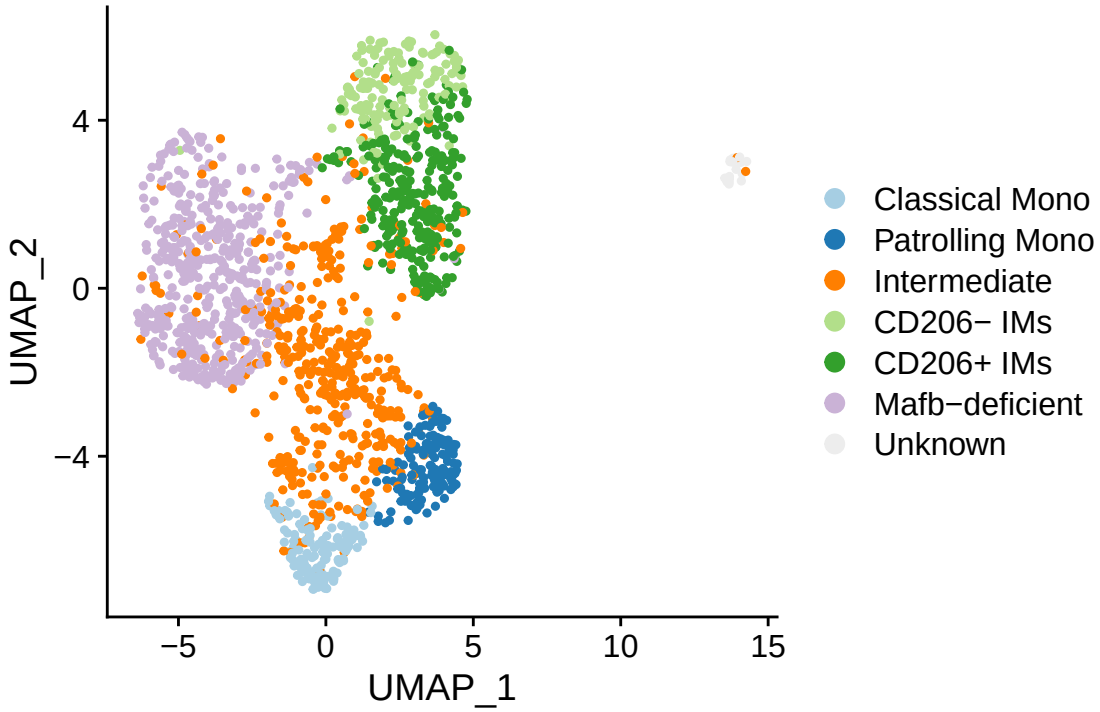
```
suppressMessages({
  library(Seurat)
  library(ggplot2)
})

results <- readRDS(file = "../12-cMAF_and_Mafb_deficient_IM/All_samples_
    Maf.seuratObject.Rds")

# we will work on control and Mafb-KO samples, so subset:
so <- subset(results, subset = group == c("HT5-Control", "HT7-MAFb-KO"))
```

3 Compare populations

```
pal3 <- c(
  "#A6CEE3", # cMo
  "#1F78B4", # pMo
  "#FF7F00", # Intermediate
  "#B2DF8A", # MHCII IM
  "#33A02C", # CD206 IM
  "#CAB2D6", # Mafb- neo
  "#ededed"# Unknown
)
DimPlot(so, cols = pal3)
```



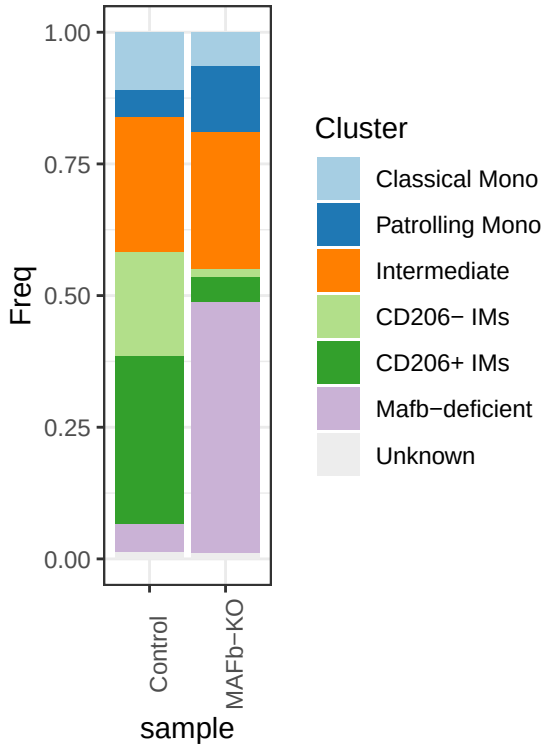

```
ggsave(filename = "../Figures/UMAPplot_Ctl_MafbKO_with_legend.pdf",
```

```
       width = 6, height = 4)
```

See population frequencies:

```
source("../R/SeuratFreqTable.R")
freq.celltype.list <- list(
  `Control` = Seurat2CellFreqTable(subset(so, subset = group == "HT5-
     Control"), slotName = "cell.type3"),
  `MAFb-KO` = Seurat2CellFreqTable(subset(so, subset = group == "HT7-MAFb-
     KO"), slotName = "cell.type3")
)

source("../R/barChart.R")
barChart(freq.celltype.list) + labs(fill = "Cluster") + scale_fill_manual(
   values = pal3) + theme(axis.text.x = element_text(angle = 90))
```

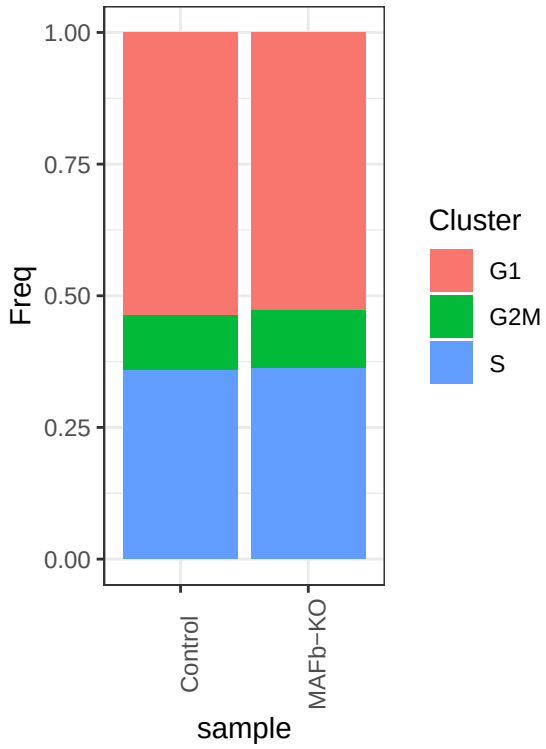



```
ggsave(filename = "../Figures/Barplot_Ctl_MafbKO_population_frequency.pdf"
   ,
       width = 3, height = 4)
```

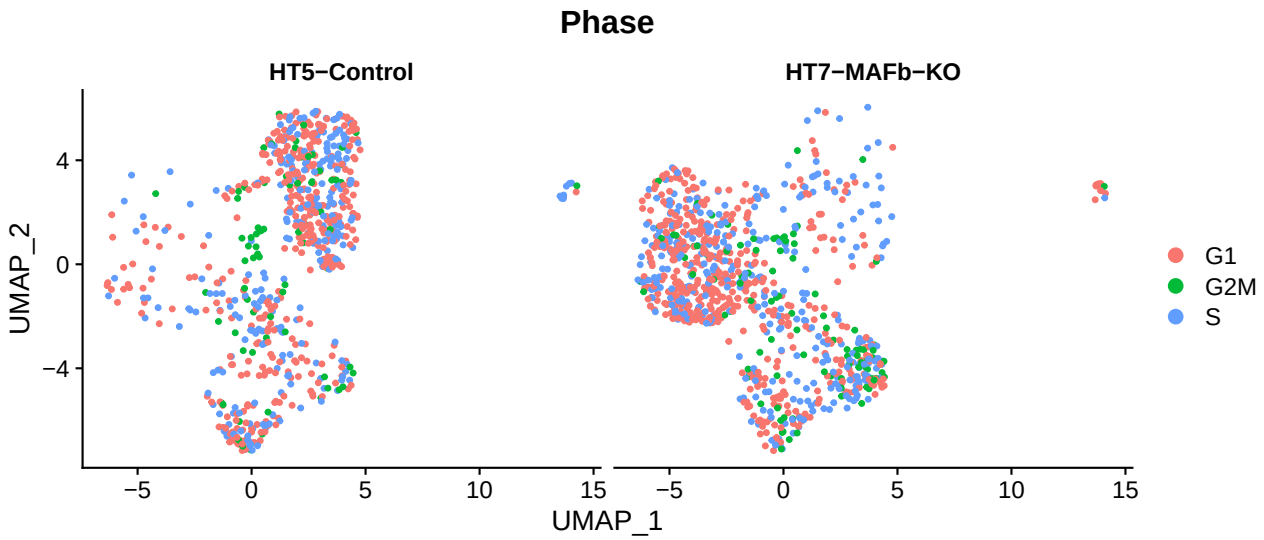
4 Proliferation comparison

```
freq.celltype.list <- list(
  `Control` = Seurat2CellFreqTable(subset(so, subset = group == "HT5-
     Control"), slotName = "Phase"),
  `MAFb-KO` = Seurat2CellFreqTable(subset(so, subset = group == "HT7-MAFb-
     KO"), slotName = "Phase")
)
```

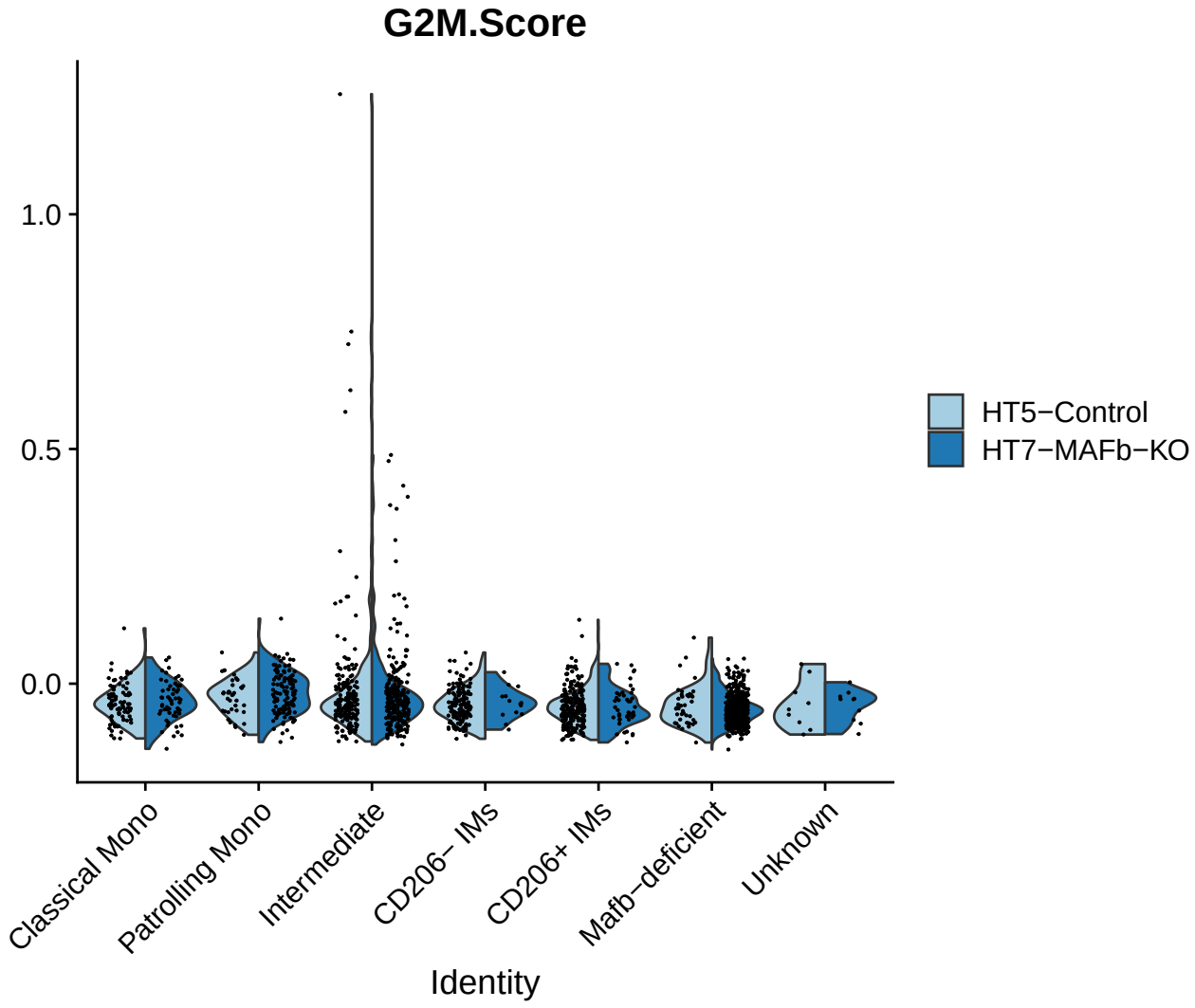
```
barChart(freq.celltype.list) + labs(fill = "Cluster") +  theme(axis.text.x
    = element_text(angle = 90))
```




```
DimPlot(so, group.by = "Phase", split.by = "group")
```




```
VlnPlot(so, features = "G2M.Score", split.by = "group", cols = pal3, split
    .plot = TRUE)
```



5 Comparison between Mafb-deficient population and IMs

Let's focus on the Mafb-deficient population in Mafb-deficient sample.

```
neo_IM <- subset(so, subset = cell.type3 %in% c("Mafb-deficient", "CD206-␣
    IMs", "CD206+␣IMs"))
```

5.1 DE genes between Mafb- neo and IM population

```
library(dplyr)
Neo_vs_IM <- FindMarkers(so,
                         ident.1 = "Mafb-deficient",
                         ident.2 = c("CD206-␣IMs", "CD206+␣IMs"),
                         logfc.threshold = 0,
                         verbose = FALSE)


# keep only adj p value < 0.05 and logFC > 0.5 as significant markers.
```

```
Neo_vs_IM.markers <- Neo_vs_IM[Neo_vs_IM$p_val_adj < 0.05 & abs(Neo_vs_IM$
   avg_log2FC) > 0.5, ]
Neo_vs_IM.markers <- Neo_vs_IM.markers[order(Neo_vs_IM.markers$avg_log2FC,
    decreasing = TRUE), ]
nrow(Neo_vs_IM.markers)
```

```
## [1] 225
```

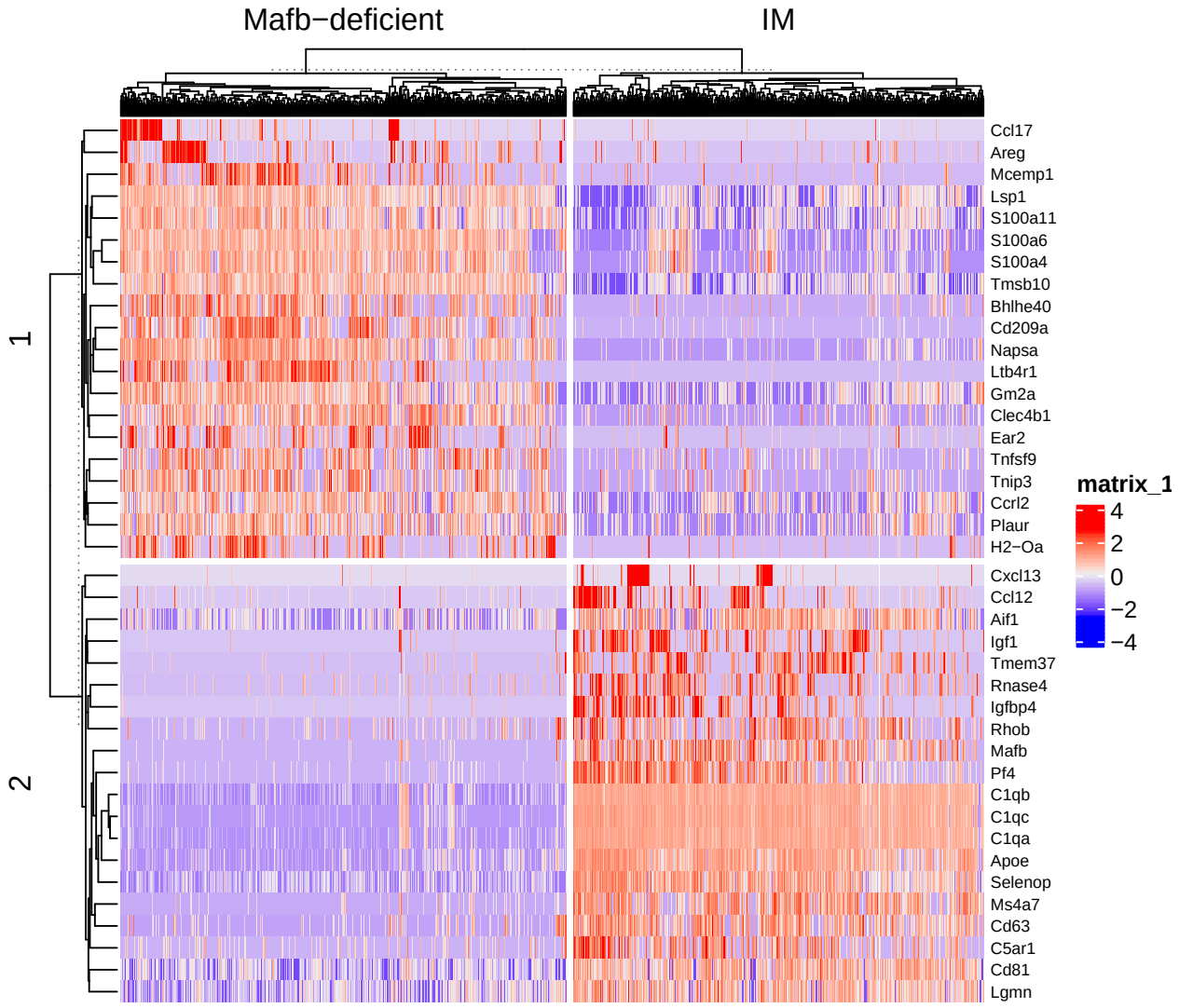
```
write.csv(Neo_vs_IM.markers ,file = "./Mafb-deficient_vs_IM.DEgenes.
   results.csv", quote = FALSE)
```

Let's show the top 20 of each side:

```
Neo_vs_IM.markers.top20 <- Neo_vs_IM.markers[c(1:20, (nrow(Neo_vs_IM.
   markers)-19):(nrow(Neo_vs_IM.markers))), ]
```

```
library(ComplexHeatmap)
mat <- GetAssayData(neo_IM)[rownames(Neo_vs_IM.markers.top20), ]
mat.scale <- t(scale(t(as.matrix(mat))))
```

```
Heatmap(mat.scale, show_row_names = TRUE, show_column_names = FALSE,
        row_names_gp = gpar ( fontsize = 7),
        column_split = factor(neo_IM$cell.type2),
        km = 2)
```



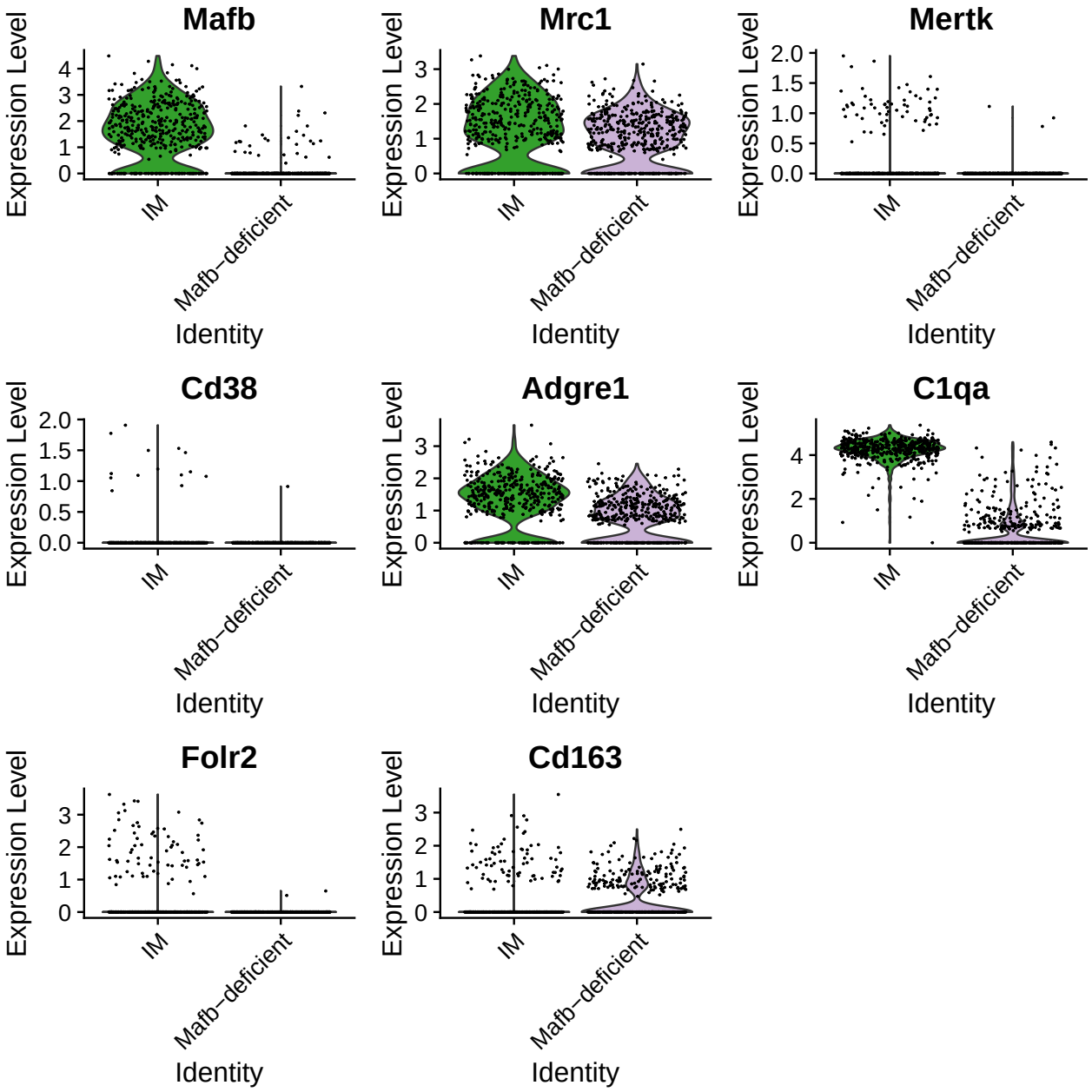

```
ggsave(filename = "../Figures/Heatmap_IM_vs_Mafb-neo.pdf", height = 10,
    width = 8)
```

1

Show in vlnplot

```
VlnPlot(neo_IM, features = c("Mafb", "Mrc1", "Mertk", "Cd38", "Adgre1", "
    C1qa", "Folr2", "Cd163"), group.by = "cell.type2", cols = c("#33A02C",
     "#CAB2D6"))
```

1



5.2 GO enrichment analysis with DE genes

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

5.2.1 GO enrichment analysis of up-regulated DE genes in Mafb-deficient population

```
DE.MafbKO <- Neo_vs_IM.markers[Neo_vs_IM.markers$avg_log2FC > 0, ]
symb <- rownames(DE.MafbKO)
de_entrez <- bitr( geneID = symb, fromType = "SYMBOL", toType = "ENTREZID"
    , OrgDb = "org.Mm.eg.db", drop = TRUE ) $ ENTREZID
result.enrichGO <- enrichGO(de_entrez, OrgDb = "org.Mm.eg.db", ont = "BP")
```

```
result.enrichGO <- replaceEntrezID(result.enrichGO, organism = "mmu")
write.csv(result.enrichGO@result, file = "./Results_enrichment/enrichGO_DE_Mafb_KO_vsIM.csv")
result.enrichGO@result
```

```
## # A tibble: 2,112 x 9
##    ID     Description GeneRatio BgRatio   pvalue p.adjust  qvalue geneID   Count
##    <chr>  <chr>       <chr>     <chr>      <dbl>    <dbl>   <dbl> <chr>    <int>
##  1 GO:00~ leukocyte ~ 14/106    360/23~ 9.91e-10  2.09e-6 1.49e-6 Ccl17/I~    14
##  2 GO:00~ myeloid le~ 11/106    219/23~ 4.94e- 9  5.22e-6 3.72e-6 Ccl17/C~    11
##  3 GO:00~ cell chemo~ 12/106    303/23~ 1.38e- 8  9.68e-6 6.89e-6 Ccl17/C~    12
##  4 GO:00~ leukocyte ~ 10/106    219/23~ 6.29e- 8  3.32e-5 2.36e-5 Ccl17/C~    10
##  5 GO:19~ neutrophil~ 8/106     124/23~ 9.90e- 8  4.18e-5 2.98e-5 Ccl17/L~     8
##  6 GO:00~ regulation~ 11/106    321/23~ 2.46e- 7  8.42e-5 5.99e-5 Cd209a/~    11
##  7 GO:00~ fever gene~ 4/106     13/233~ 2.79e- 7  8.42e-5 5.99e-5 Ptgs2/I~     4
##  8 GO:00~ neutrophil~ 7/106     99/233~ 3.46e- 7  9.13e-5 6.50e-5 Ccl17/L~     7
##  9 GO:00~ granulocyt~ 8/106     155/23~ 5.51e- 7  1.02e-4 7.25e-5 Ccl17/L~     8
## 10 GO:00~ regulation~ 12/106    427/23~ 5.66e- 7  1.02e-4 7.25e-5 Tnfsf9/~    12
## # ... with 2,102 more rows
```

5.2.2 GO enrichment analysis of up-regulated DE genes in IMs

```
DE.IM <- Neo_vs_IM.markers[Neo_vs_IM.markers$avg_log2FC < 0, ]
symb <- rownames(DE.IM)
de_entrez <- bitr( geneID = symb, fromType = "SYMBOL", toType = "ENTREZID", OrgDb = "org.Mm.eg.db", drop = TRUE ) $ ENTREZID
result.enrichGO <- enrichGO(de_entrez, OrgDb = "org.Mm.eg.db", ont = "BP")
result.enrichGO <- replaceEntrezID(result.enrichGO, organism = "mmu")
write.csv(result.enrichGO@result, file = "./Results_enrichment/enrichGO_DE_IM_vsMafbKO.csv")
result.enrichGO@result
```

```
## # A tibble: 2,479 x 9
##    ID     Description  GeneRatio BgRatio   pvalue p.adjust  qvalue geneID  Count
##    <chr>  <chr>        <chr>     <chr>      <dbl>    <dbl>   <dbl> <chr>   <int>
##  1 GO:00~ cell chemot~ 16/112    303/23~ 1.37e-12  2.40e-9 1.63e-9 Cmklr1~    16
##  2 GO:00~ leukocyte c~ 14/112    219/23~ 2.91e-12  2.40e-9 1.63e-9 Cmklr1~    14
```

```
##  3 GO:00~ myeloid leu~ 14/112    219/23~ 2.91e-12  2.40e-9 1.63e-9
   Cmklr1~    14
##  4 GO:00~ positive re~ 12/112    156/23~ 1.30e-11  8.03e-9 5.46e-9
   Cmklr1~    12
##  5 GO:00~ leukocyte m~ 16/112    360/23~ 1.84e-11  9.14e-9 6.22e-9
   Cmklr1~    16
##  6 GO:00~ regulation ~ 13/112    217/23~ 4.11e-11  1.70e-8 1.16e-8
   Cmklr1~    13
##  7 GO:00~ ERK1 and ER~ 15/112    325/23~ 4.87e-11  1.72e-8 1.17e-8
   Atp6v0~    15
##  8 GO:00~ regulation ~ 10/112    121/23~ 3.50e-10  1.09e-7 7.38e-8
   Cmklr1~    10
##  9 GO:00~ mononuclear~ 9/112     89/233~ 4.51e-10  1.24e-7 8.45e-8
   Cmklr1~     9
## 10 GO:00~ positive re~ 9/112     96/233~ 8.95e-10  2.22e-7 1.51e-7
   Cmklr1~     9
## # ... with 2,469 more rows
```

5.2.3 Volcano plot of DE genes

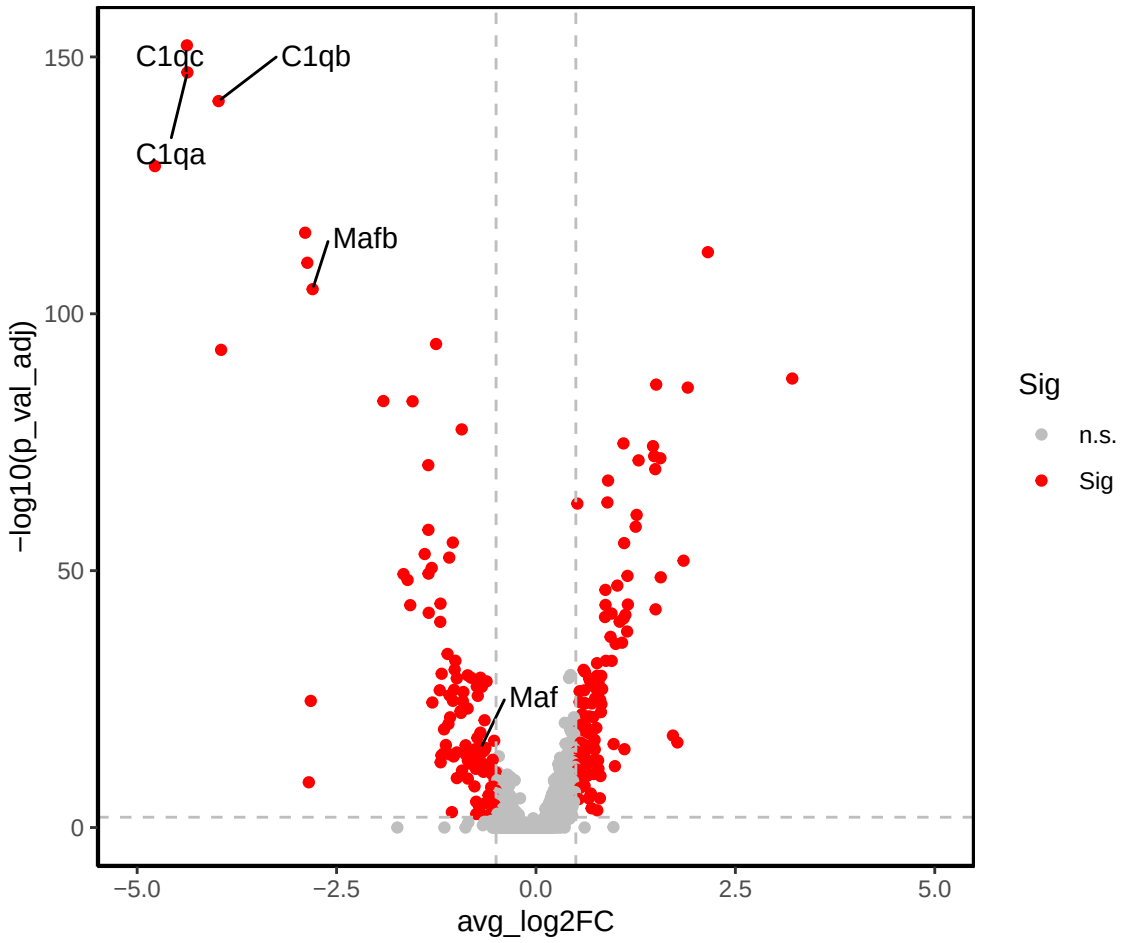
```
suppressMessages({
  library(dplyr)
  library(ggrepel)
  })

Neo_vs_IM.volcano = mutate(Neo_vs_IM,
 Sig=ifelse((abs(Neo_vs_IM$avg_log2FC)>0.5)&(Neo_vs_IM$p_val_adj < 0.05),
    "Sig", "n.s."))
# res2 = mutate(res2,
#              Sig=ifelse((abs(res2$avg_logFC)>1)&(res2$p_val_adj < 0.01),
    "Sig", "n.s."))


Neo_vs_IM.volcano$Gene <- rownames(Neo_vs_IM.volcano)
Gene.to.show.ValcanoPlot <- c("C1qa", "C1qb", "C1qc", "Maf", "Cd209", "
   Mafb")
```

```
p <- ggplot(Neo_vs_IM.volcano, aes(avg_log2FC, -log10(p_val_adj))) + geom_
   point(aes(col=Sig)) + scale_color_manual(values=c( "gray", "red"))

p + geom_text_repel(data=filter(Neo_vs_IM.volcano, Gene %in% Gene.to.show.
   ValcanoPlot), aes(label=Gene), box.padding = 1) + xlim(c(-5, 5)) +
   theme_classic() + theme(panel.border = element_rect(colour = "black",
   fill = NA, size = 1)) + geom_hline(yintercept = -log10(0.01), linetype=
   'dashed', col = 'grey') + geom_vline(xintercept = c(-0.5, 0.5),
   linetype='dashed', col = 'grey')
```




```
ggsave(filename = "../Figures/VolcanoPlot_Mafb-KO_vs_IM_DE_genes.pdf",
    height = 5, width = 6)
```

6 Session information

R sesssion:

```
sessionInfo()
```

```
## R version 4.0.3 (2020-10-10)
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: Ubuntu 20.04.3 LTS
##
## Matrix products: default
## BLAS:   /usr/lib/x86_64-linux-gnu/openblas-pthread/libblas.so.3
## LAPACK: /usr/lib/x86_64-linux-gnu/openblas-pthread/liblapack.so.3
##
## locale:
##  [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
##  [3] LC_TIME=en_GB.UTF-8        LC_COLLATE=en_US.UTF-8
##  [5] LC_MONETARY=en_GB.UTF-8    LC_MESSAGES=en_US.UTF-8
##  [7] LC_PAPER=en_GB.UTF-8       LC_NAME=C
##  [9] LC_ADDRESS=C               LC_TELEPHONE=C
```

```
## [11] LC_MEASUREMENT=en_GB.UTF-8 LC_IDENTIFICATION=C
##
## attached base packages:
##  [1] parallel  stats4    grid      stats     graphics  grDevices utils
##  [8] datasets  methods   base
##
## other attached packages:
##  [1] ggrepel_0.9.1          org.Mm.eg.db_3.12.0    AnnotationDbi_1.52.0
##  [4] IRanges_2.24.1         S4Vectors_0.28.1       Biobase_2.50.0
##  [7] BiocGenerics_0.36.1    clusterProfiler_3.18.1 ComplexHeatmap_2.6.2
## [10] dplyr_1.0.7            RColorBrewer_1.1-2     ggplot2_3.3.5
## [13] SeuratObject_4.0.4     Seurat_4.0.5
##
## loaded via a namespace (and not attached):
##   [1] shadowtext_0.0.9      circlize_0.4.13       fastmatch_1.1-3
##   [4] systemfonts_1.0.3     plyr_1.8.6            igraph_1.2.9
##   [7] lazyeval_0.2.2        splines_4.0.3         BiocParallel_1.24.1
##  [10] listenv_0.8.0         scattermore_0.7       digest_0.6.29
##  [13] yulab.utils_0.0.4     GOSemSim_2.16.1       htmltools_0.5.2
##  [16] viridis_0.6.2         GO.db_3.12.1          magick_2.7.3
##  [19] fansi_0.5.0           magrittr_2.0.1        memoise_2.0.1
##  [22] tensor_1.5            cluster_2.1.0         ROCR_1.0-11
##  [25] limma_3.46.0          graphlayouts_0.7.2    globals_0.14.0
##  [28] matrixStats_0.61.0    spatstat.sparse_2.0-0 enrichplot_1.10.2
##  [31] colorspace_2.0-2      blob_1.2.2            textshaping_0.3.6
##  [34] xfun_0.28             crayon_1.4.2          jsonlite_1.7.2
##  [37] scatterpie_0.1.7      spatstat.data_2.1-0   survival_3.2-7
##  [40] zoo_1.8-9             glue_1.5.1            polyclip_1.10-0
##  [43] gtable_0.3.0          leiden_0.3.9          GetoptLong_1.0.5
##  [46] future.apply_1.8.1    shape_1.4.6           abind_1.4-5
##  [49] scales_1.1.1          DOSE_3.16.0           DBI_1.1.1
##  [52] miniUI_0.1.1.1        Rcpp_1.0.7            viridisLite_0.4.0
##  [55] xtable_1.8-4          clue_0.3-60           reticulate_1.22
##  [58] spatstat.core_2.3-2   bit_4.0.4             htmlwidgets_1.5.4
##  [61] httr_1.4.2            fgsea_1.16.0          ellipsis_0.3.2
##  [64] ica_1.0-2             pkgconfig_2.0.3       farver_2.1.0
##  [67] uwot_0.1.11           deldir_1.0-6          utf8_1.2.2
##  [70] tidyselect_1.1.1      labeling_0.4.2        rlang_0.4.12
##  [73] reshape2_1.4.4        later_1.3.0           munsell_0.5.0
##  [76] tools_4.0.3           cachem_1.0.6          cli_3.1.0
##  [79] downloader_0.4        generics_0.1.1        RSQLite_2.2.9
##  [82] ggridges_0.5.3        evaluate_0.14         stringr_1.4.0
##  [85] fastmap_1.1.0         yaml_2.2.1            ragg_1.2.1
##  [88] goftest_1.2-3         knitr_1.36            bit64_4.0.5
##  [91] tidygraph_1.2.0       fitdistrplus_1.1-6    purrr_0.3.4
##  [94] RANN_2.6.1            ggraph_2.0.5          pbapply_1.5-0
##  [97] future_1.23.0         nlme_3.1-153          mime_0.12
## [100] DO.db_2.9             rstudioapi_0.13       compiler_4.0.3
## [103] plotly_4.10.0         png_0.1-7             spatstat.utils_2.2-0
## [106] tweenr_1.0.2          tibble_3.1.6          stringi_1.7.6
## [109] highr_0.9             lattice_0.20-41       Matrix_1.3-4
## [112] vctrs_0.3.8           pillar_1.6.4          lifecycle_1.0.1
## [115] BiocManager_1.30.16   spatstat.geom_2.3-0   lmtest_0.9-39
## [118] GlobalOptions_0.1.2   RcppAnnoy_0.0.19      data.table_1.14.2
```

```
## [121] cowplot_1.1.1          irlba_2.3.5           qvalue_2.22.0
## [124] httpuv_1.6.3           patchwork_1.1.1       R6_2.5.1
## [127] promises_1.2.0.1       KernSmooth_2.23-20    gridExtra_2.3
## [130] parallelly_1.29.0      codetools_0.2-18      MASS_7.3-53
## [133] assertthat_0.2.1       rjson_0.2.20          withr_2.4.3
## [136] sctransform_0.3.2      mgcv_1.8-33           ggfun_0.0.4
## [139] rpart_4.1-15           tidyr_1.1.4           rvcheck_0.2.1
## [142] rmarkdown_2.11         Cairo_1.5-12.2        Rtsne_0.15
## [145] ggforce_0.3.3          shiny_1.7.1
```

7 References