R Notebook

Code ▾

```
#install.packages(c("ggrepel", "dplyr", "gridExtra", "grid", "lattice", "ggpubr", "umap", "ggfortify", "magick",
"BiocManager", "readxl", "rmarkdown", "tibble", "rlist"))
#BiocManager::install(c('DESeq2', 'ComplexHeatmap', 'biomaRt', 'fgsea', 'msigdbr'))

library(rmarkdown)
library(readxl)
library(DESeq2)
library(ComplexHeatmap)
library(circlize)
library(dendsort)
library(biomaRt)
library(fgsea)
library(data.table)
library(msigdbr)
library(ggrepel)
library(dplyr)
library(gridExtra)
library(grid)
library(lattice)
library(ggpubr)
library(umap)
library(ggfortify)
library(magick)
library(tibble)
library(rlist)
library(outliers)
```

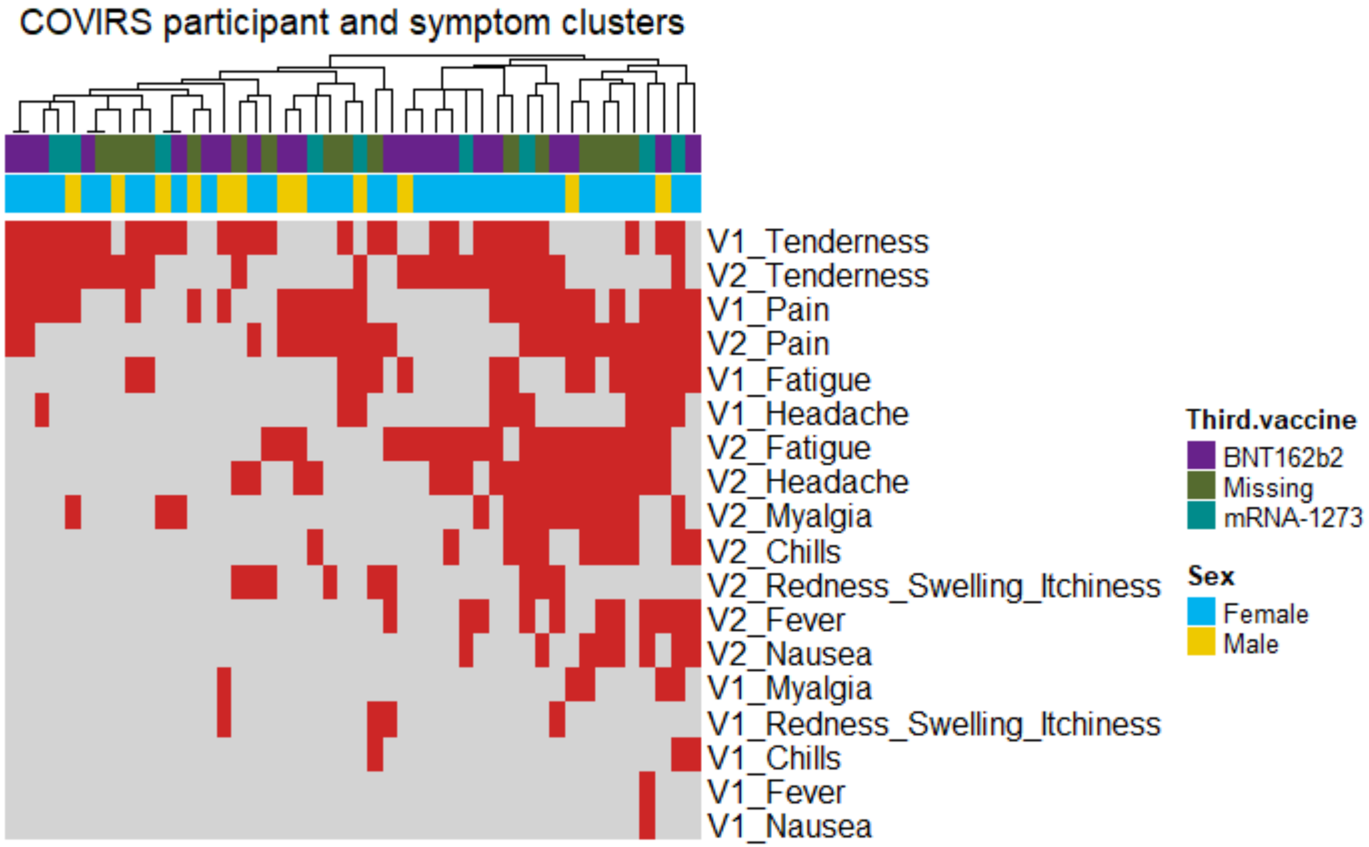
```
# ensembl=useMart("ensembl")
# ensembl = useDataset("hsapiens_gene_ensembl",mart=ensembl)
# attributes = listAttributes(ensembl)
# ensids=rownames(count_data_deseq)
# ens2gene = getBM(attributes=c('external_gene_name','ensembl_gene_id','gene_biotype','entrezgene_id', 'description'),
#                   filters = 'ensembl_gene_id',
#                   values = ensids,
#                   mart = ensembl)
# ens2gene = ens2gene[!duplicated(ens2gene$ensembl_gene_id),]
# rownames(ens2gene) = ens2gene$ensembl_gene_id
# head(ens2gene)
# saveRDS(ens2gene, "ens2gene.RDS")
```

```
count_data = readRDS('data/Raw_RNASeq_Data.RDS')
ens2gene = readRDS("data/ens2gene.RDS")
meta_data = data.frame(read.csv('data/COVIRS_metadata.csv', header = TRUE))
```

1. Defining patient groups based on symptoms

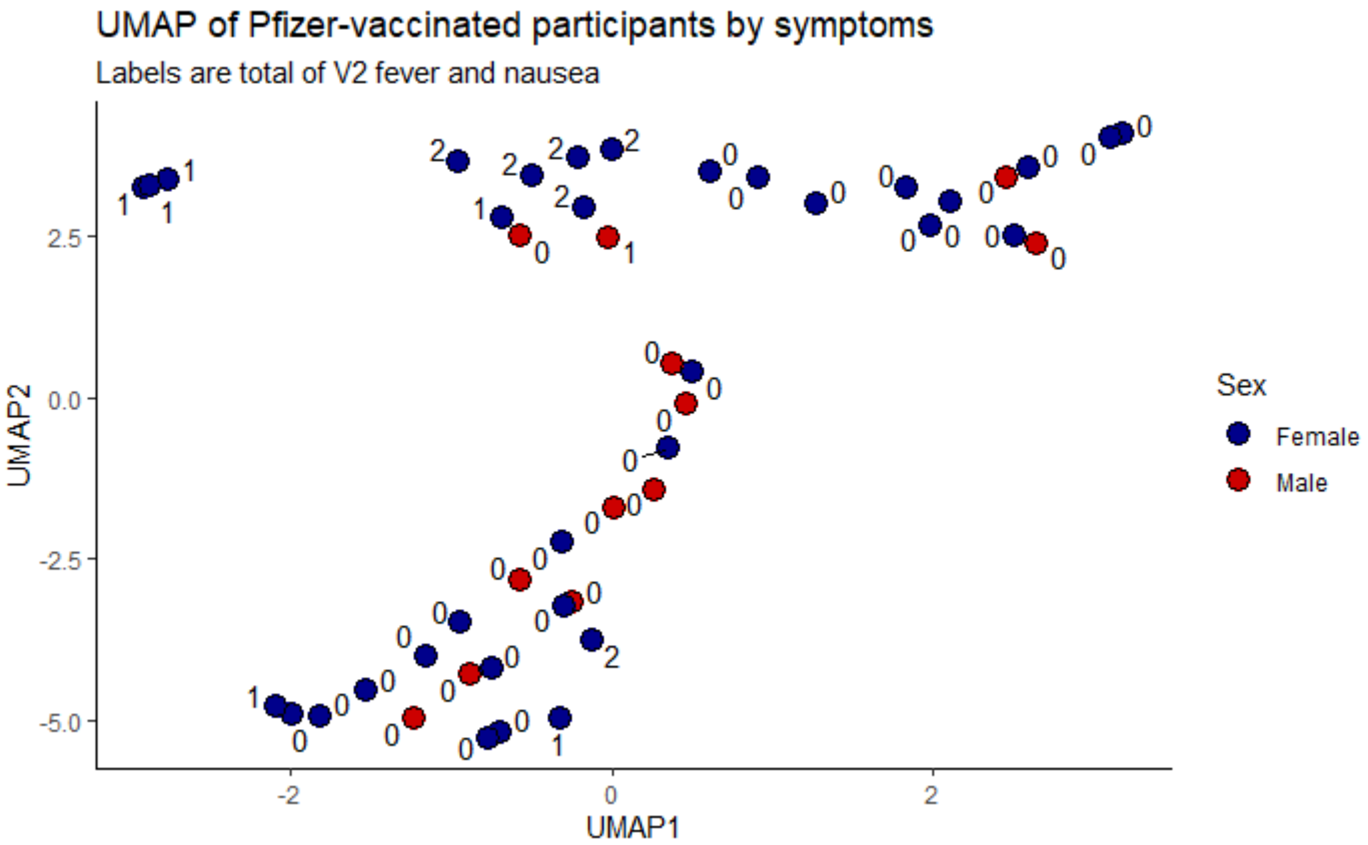
```
meta_data = meta_data[meta_data$Timepoint == 'V2A' & meta_data$First.Second.vaccine == 'BNT162b2', ]
```

```
plot = data.frame(subset(meta_data, select = -c(id, Participant.ID, Run, SAGC_ID, First.Second.vaccine, Third.vaccine, Sex, Timepoint)))
col_dend = dendsort(hclust(dist(plot)))
ha_participants = HeatmapAnnotation(df = meta_data[ ,c('Third.vaccine', 'Sex')], col = list(Third.vaccine = c('mRNA-1273' = 'darkcyan', 'BNT162b2' = 'darkorchid4', 'Missing' = 'darkolivegreen'), Sex = c('Male' = 'gold2', 'Female' = 'deepskyblue2')), show_annotation_name = FALSE)
Heatmap(as.matrix(t(plot)), show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = col_dend, show_row_dend = FALSE, top_annotation = ha_participants, show_heatmap_legend = FALSE, col = c('lightgrey', 'firebrick3'), column_title = 'COVIRS participant and symptom clusters')
```

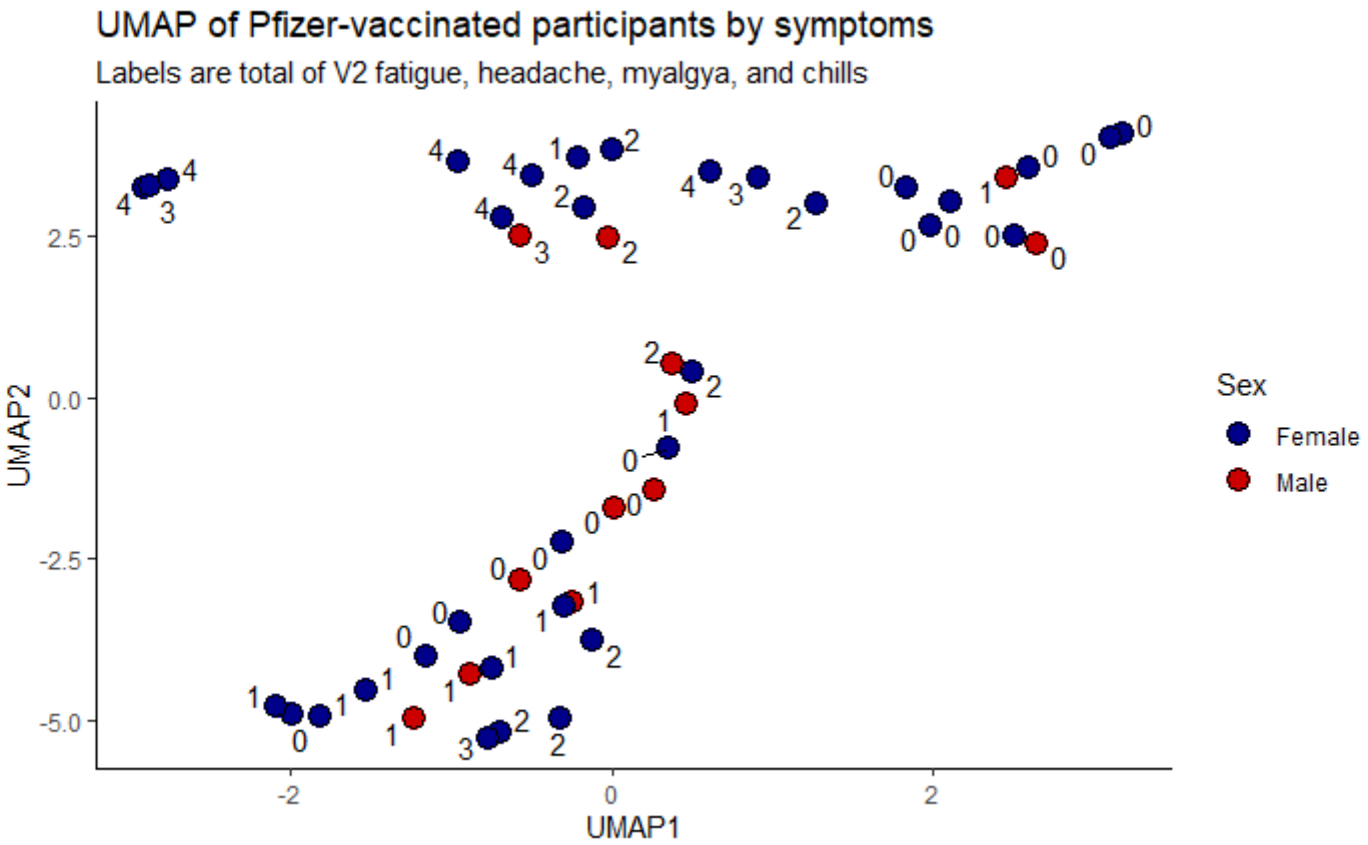


```
rm(plot); rm(ha_participants); rm(col_dend);
```

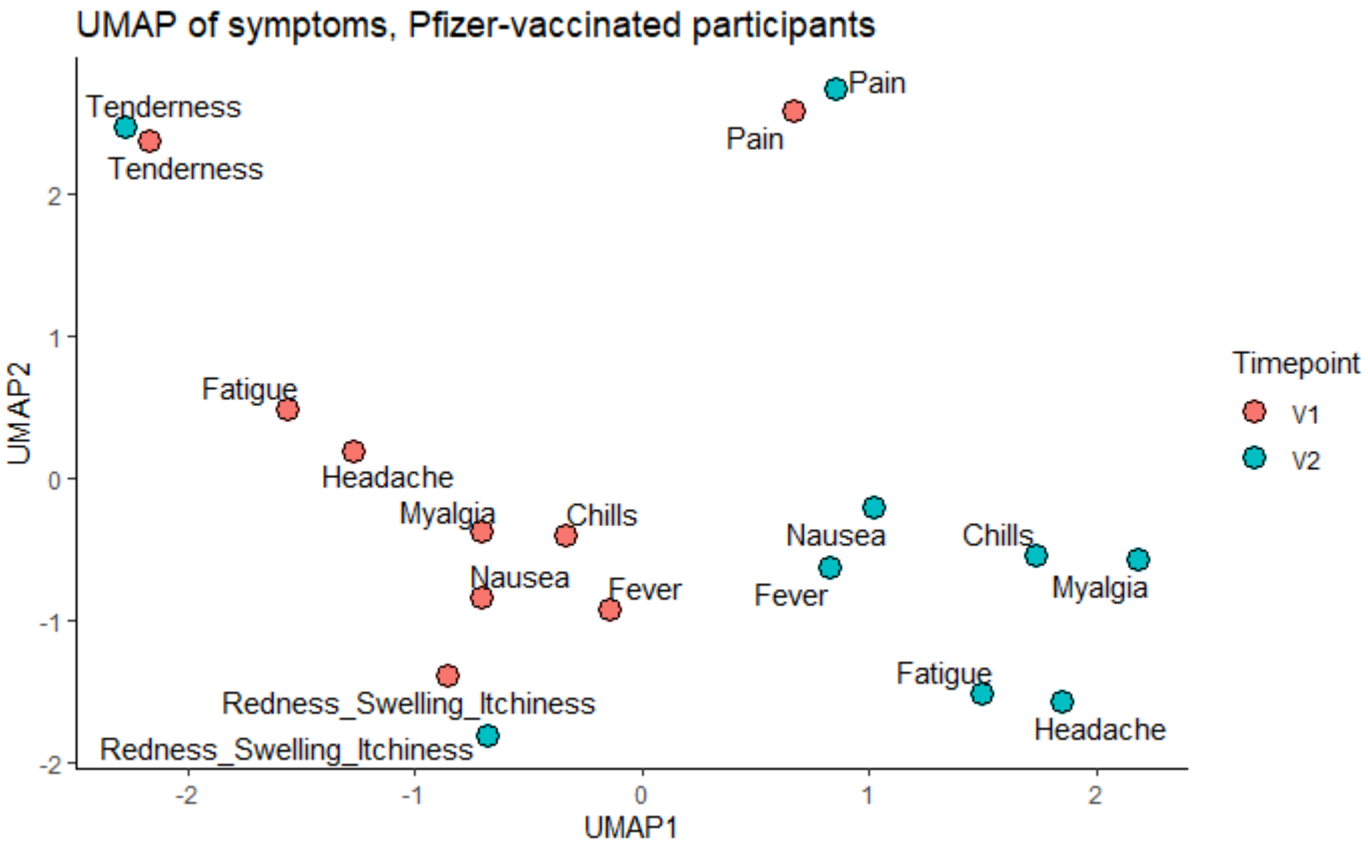
```
custom.config = umap.defaults
custom.config$random_state = 1
custom.config$n_neighbors = 3
plot = meta_data
umap = umap(as.matrix(apply(subset(plot, select = -c(id, Participant.ID, Run, SAGC_ID, Timepoint, First.Second.vaccine, Third.vaccine, Sex)), 2, as.numeric)), config = custom.config)
umap = data.frame(umap$layout)
plot$UMAP1 = umap$X1
plot$UMAP2 = umap$X2
rm(umap)
ggplot(plot, aes(x = UMAP1, y = UMAP2, fill = Sex, label = apply(subset(plot, select = c(V2_Fever, V2_Nausea)) == 1, 1, sum))) + geom_point(shape = 21, size = 4) +
  scale_fill_manual(values = c(Male = "red3", Female = "blue4")) + theme_classic() + geom_text_repel() +
  ggtitle('UMAP of Pfizer-vaccinated participants by symptoms', subtitle = 'Labels are total of V2 fever and nausea'); rm (plot)
```


UMAP of Pfizer-vaccinated participants by symptoms
Labels are total of V2 fever and nausea
UMAP2
UMAP1
2.5
0.0
-2.5
-5.0
-2
0
2
Sex
● Female
● Male
0
1
2

```r
plot = meta_data
umap = umap(as.matrix(apply(subset(plot, select = -c(id, Participant.ID, Run, SAGC_ID, Timepoint, First.Second.vaccine, Third.vaccine, Sex)), 2, as.numeric)), config = custom.config)
umap = data.frame(umap$layout)
plot$UMAP1 = umap$X1
plot$UMAP2 = umap$X2
rm(umap)
ggplot(plot, aes(x = UMAP1, y = UMAP2, fill = Sex, label = apply(subset(plot, select = c(V2_Fatigue, V2_Headache, V2_Myalgia, V2_Chills)) == 1, 1, sum))) + geom_point(shape = 21, size = 4) +
  scale_fill_manual(values = c(Male = "red3", Female = "blue4")) + theme_classic() + geom_text_repel() +
  ggtitle('UMAP of Pfizer-vaccinated participants by symptoms', subtitle = 'Labels are total of V2 fatigue, headache, myalgya, and chills'); rm (plot)
```



```
plot = data.frame(t(subset(meta_data, select = -c(id, Participant.ID, Run, SAGC_ID, Timepoint, First.Second.vacci
ne, Third.vaccine, Sex))))
umap = umap(as.matrix(apply(plot, 2, as.numeric)), config = custom.config); umap = data.frame(umap$layout)
plot$UMAP1 = umap$X1; plot$UMAP2 = umap$X2; rm(umap)
plot$Timepoint = ifelse(grepl('V1_', rownames(plot)), 'V1', 'V2')
plot$Symptom = gsub('V1_', '', rownames(plot)); plot$Symptom = gsub('V2_', '', plot$Symptom)
ggplot(plot, aes(x = UMAP1, y = UMAP2, label = Symptom, fill = Timepoint)) + geom_point(shape = 21, size = 4) + g
eom_text_repel() + theme_classic() + ggtitle('UMAP of symptoms, Pfizer-vaccinated participants')
```

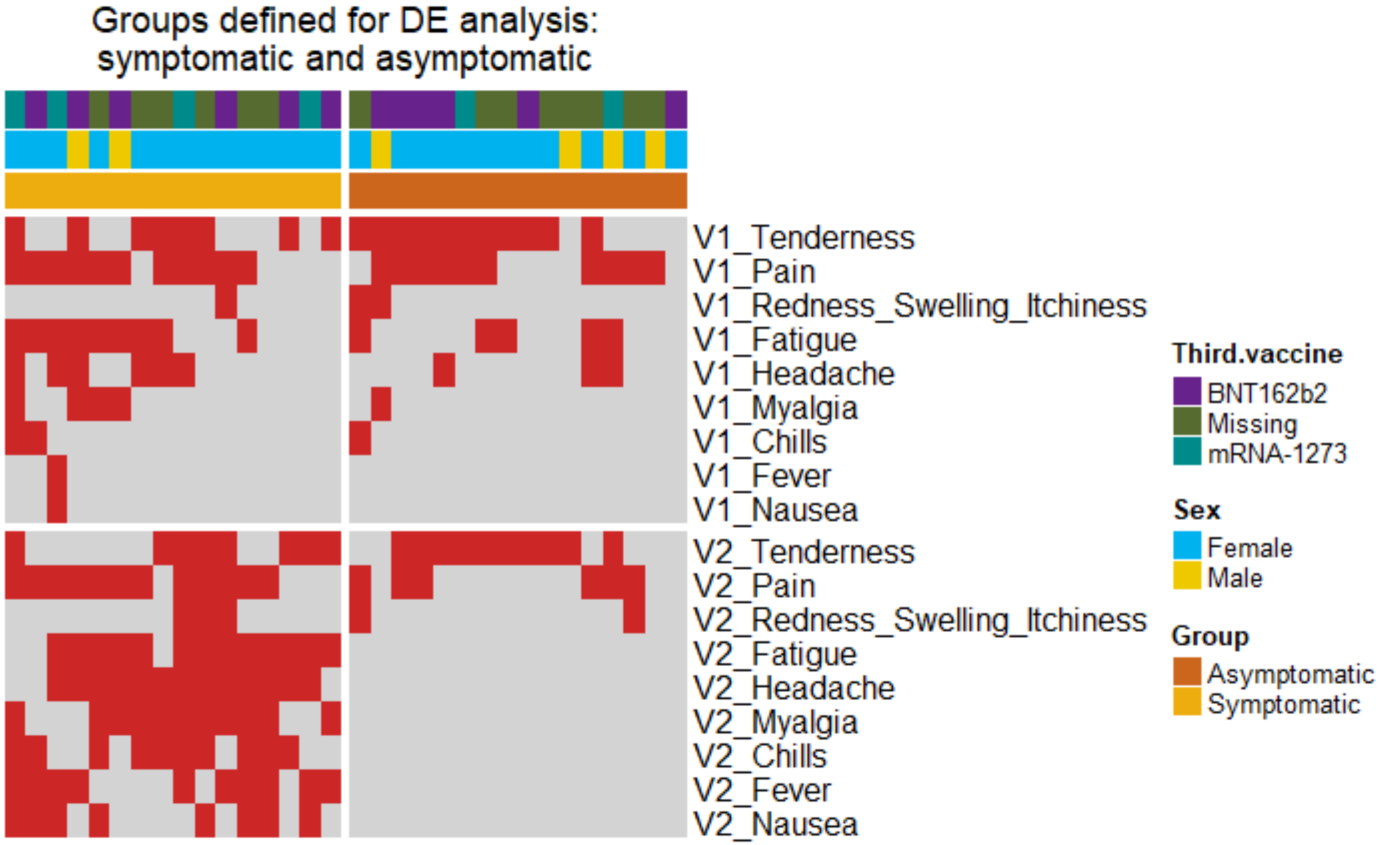


```
rm(custom.config); rm(plot)
```

```
test = meta_data[apply(subset(meta_data, select = c(V2_Fatigue, V2_Headache, V2_Myalgia, V2_Chills, V2_Fever, V2_
Nausea)) == 1, 1, sum) >= 3, ]; test$Group = "Symptomatic"
control = subset(meta_data[apply(meta_data[, 18:ncol(meta_data)] == 1, 1, sum) <= 2, ], V2_Fatigue == 0 & V2_Myal
gia == 0 & V2_Chills == 0 & V2_Headache == 0 & V2_Fever == 0 & V2_Nausea == 0); control$Group = "Asymptomatic"
```

```
meta_data = rbind(test, control); rm(test); rm(control)
count_data = count_data[, colnames(count_data) %in% meta_data$id]; meta_data = meta_data[order(meta_data$id), ]
count_data = count_data[, order(colnames(count_data))]; count_data = count_data[rownames(ens2gene), ]
count_data = data.frame(cbind(ensgene = rownames(count_data), count_data))
rownames(count_data) = NULL; rownames(meta_data) = NULL
```

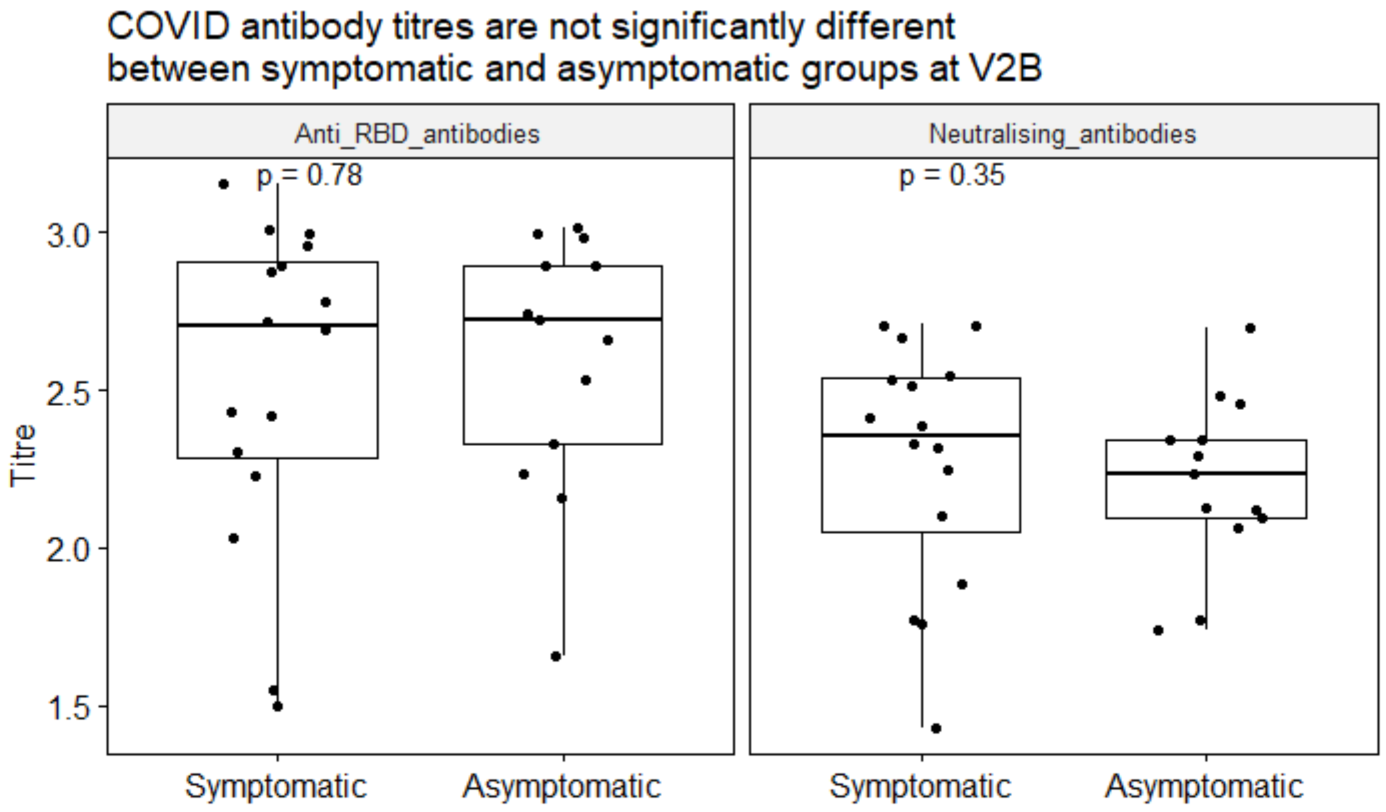
```
plot = as.matrix(meta_data[, 10:ncol(meta_data) - 1])
ha_participants = HeatmapAnnotation(df = meta_data[, c('Third.vaccine', 'Sex', 'Group')], col = list(Group = c('Asymptomatic' = 'chocolate3', 'Symptomatic' = 'darkgoldenrod2'), Third.vaccine = c('mRNA-1273' = 'darkcyan', 'BNT162b2' = 'darkorchid4', 'Missing' = 'darkolivegreen'), Sex = c('Male' = 'gold2', 'Female' = 'deepskyblue2')), show_annotation_name = FALSE)
Heatmap(t(plot), column_split = meta_data$Group, row_order = c("V1_Tenderness", "V1_Pain", "V1_Redness_Swelling_Itchiness", "V1_Fatigue", "V1_Headache", "V1_Myalgia", "V1_Chills", "V1_Fever", "V1_Nausea", "V2_Tenderness", "V2_Pain", "V2_Redness_Swelling_Itchiness", "V2_Fatigue", "V2_Headache", "V2_Myalgia", "V2_Chills", "V2_Fever", "V2_Nausea"), cluster_columns = TRUE, show_row_dend = FALSE, show_column_dend = FALSE, column_dend_side = 'top', col = c('lightgrey', 'firebrick3'), top_annotation = ha_participants, show_column_names = FALSE, show_heatmap_legend = FALSE, row_split = grepl('V2_', rownames(t(plot))), row_title = NULL, column_title = 'Groups defined for DE analysis:\nsymptomatic and asymptomatic'); rm(plot)
```



2. Antibodies and T cell responses

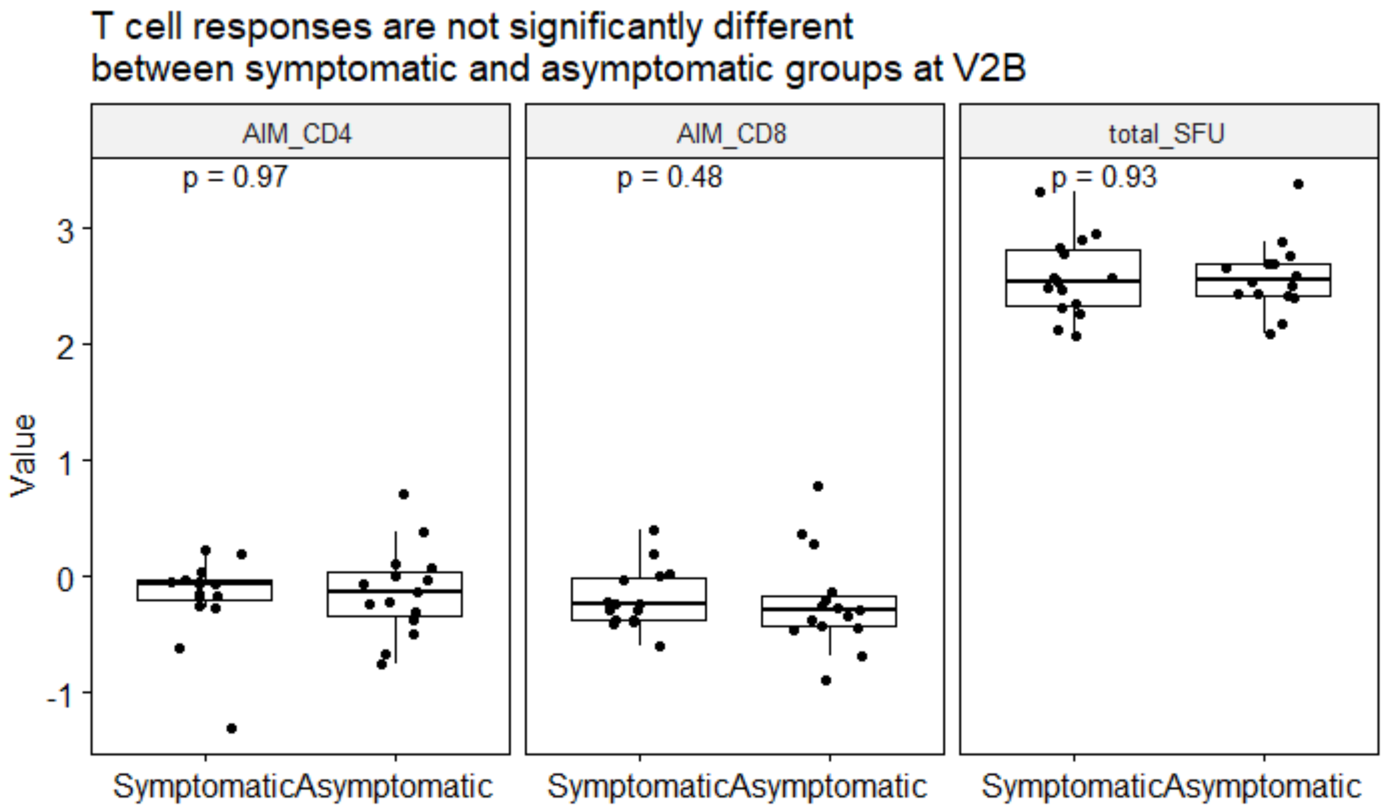
```
antibody_data = read.csv('data/antibodies.csv')
antibody_data = subset(antibody_data, Patient %in% meta_data$Participant.ID); meta_anti = meta_data[meta_data$Participant.ID %in% antibody_data$Patient, ]; antibody_data$Group = meta_anti$Group; rm(meta_anti)
```

```
ggboxplot(antibody_data, x = 'Group', y = 'Titre', add = 'jitter', facet.by = 'Antibody.type') + stat_compare_means(label = "p.format") + xlab('') + ggtitle('COVID antibody titres are not significantly different\nbetween symptomatic and asymptomatic groups at V2B')
```



```
tcell_data = read.csv('data/tcells.csv')
tcell_data = subset(tcell_data, Patient_ID %in% meta_data$Participant.ID); meta_tcell = meta_data[meta_data$Parti
cipant.ID %in% tcell_data$Patient_ID, ]; tcell_data$Group = meta_tcell$Group; rm(meta_tcell); tcell_data$Value =
log10(tcell_data$Value)
```

```
ggboxplot(tcell_data, x = 'Group', y = 'Value', add = 'jitter', facet.by = 'Type') + stat_compare_means(label = "
p.format") + xlab('') + ggtitle('T cell responses are not significantly different\nbetween symptomatic and asympt
omatic groups at V2B')
```

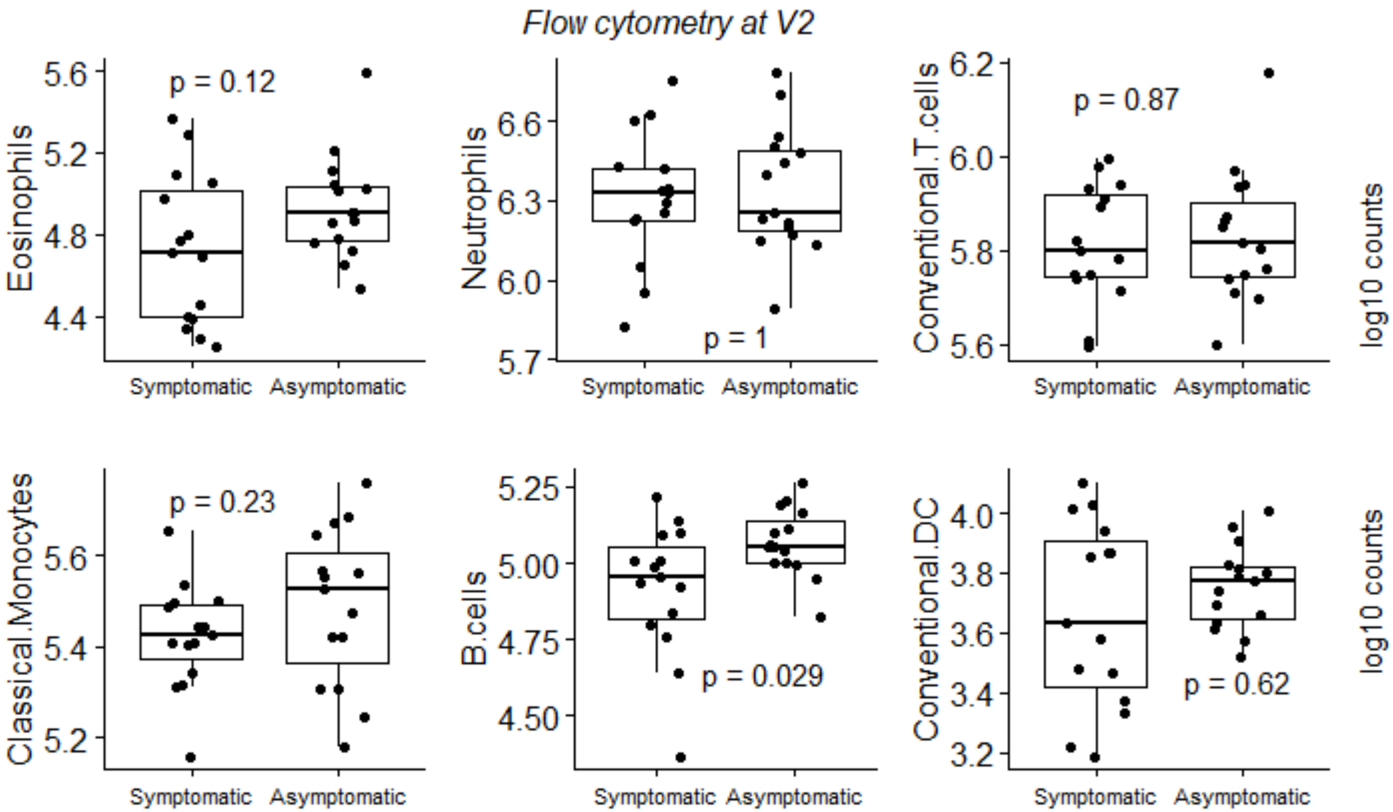


```
rm(antibody_data); rm(ha_participants); rm(tcell_data)
```

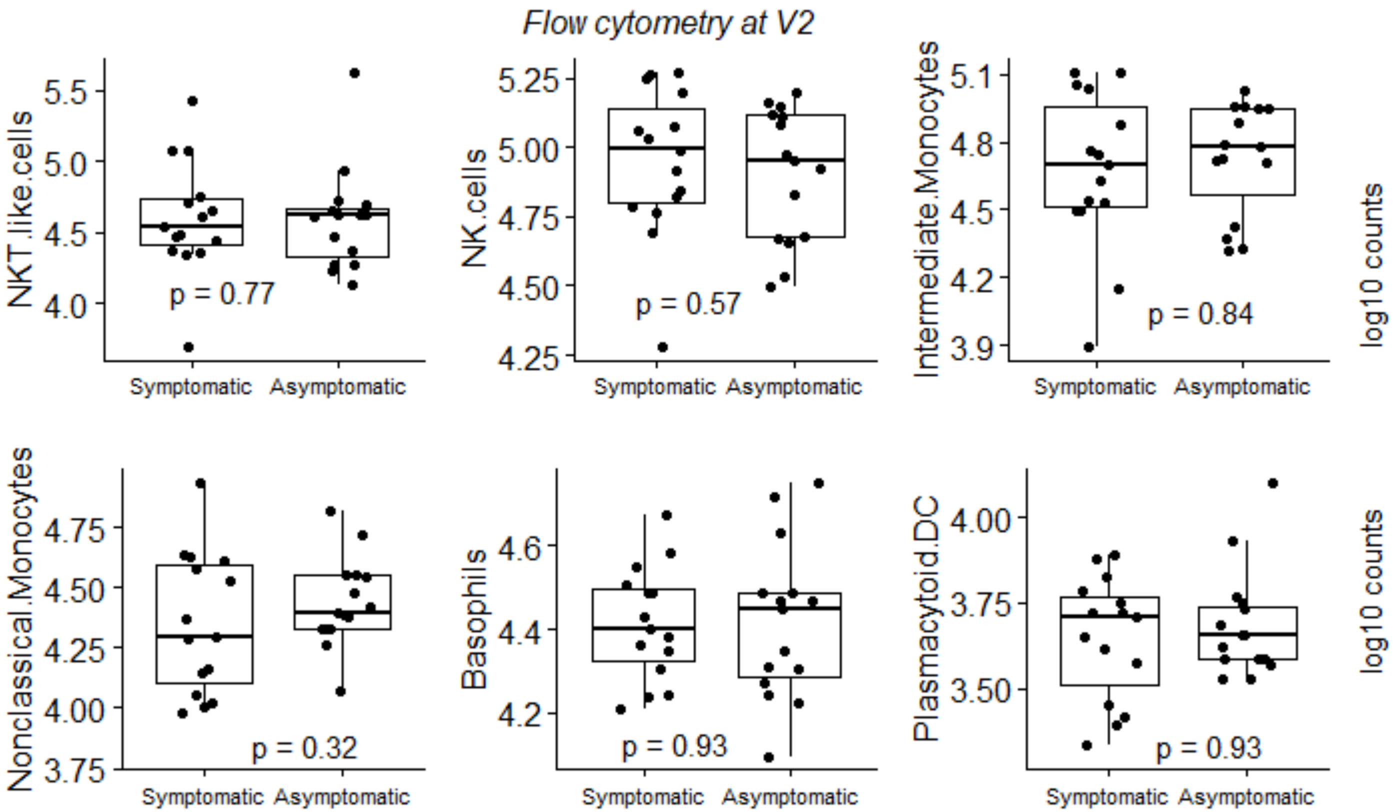
3. Flow cytometry

```
flow = read.csv('data/COVIRS_flow_data_20230202.csv'); flow = subset(flow, Patient %in% meta_data$Participant.ID
& Timepoint == 'V2A'); meta_flow = subset(meta_data, Participant.ID %in% flow$Patient); flow = flow[order(flow$Pa
tient), ]; flow$Total = rowSums(flow[, 5:ncol(flow)-1]); flow$Group = meta_flow$Group;
```

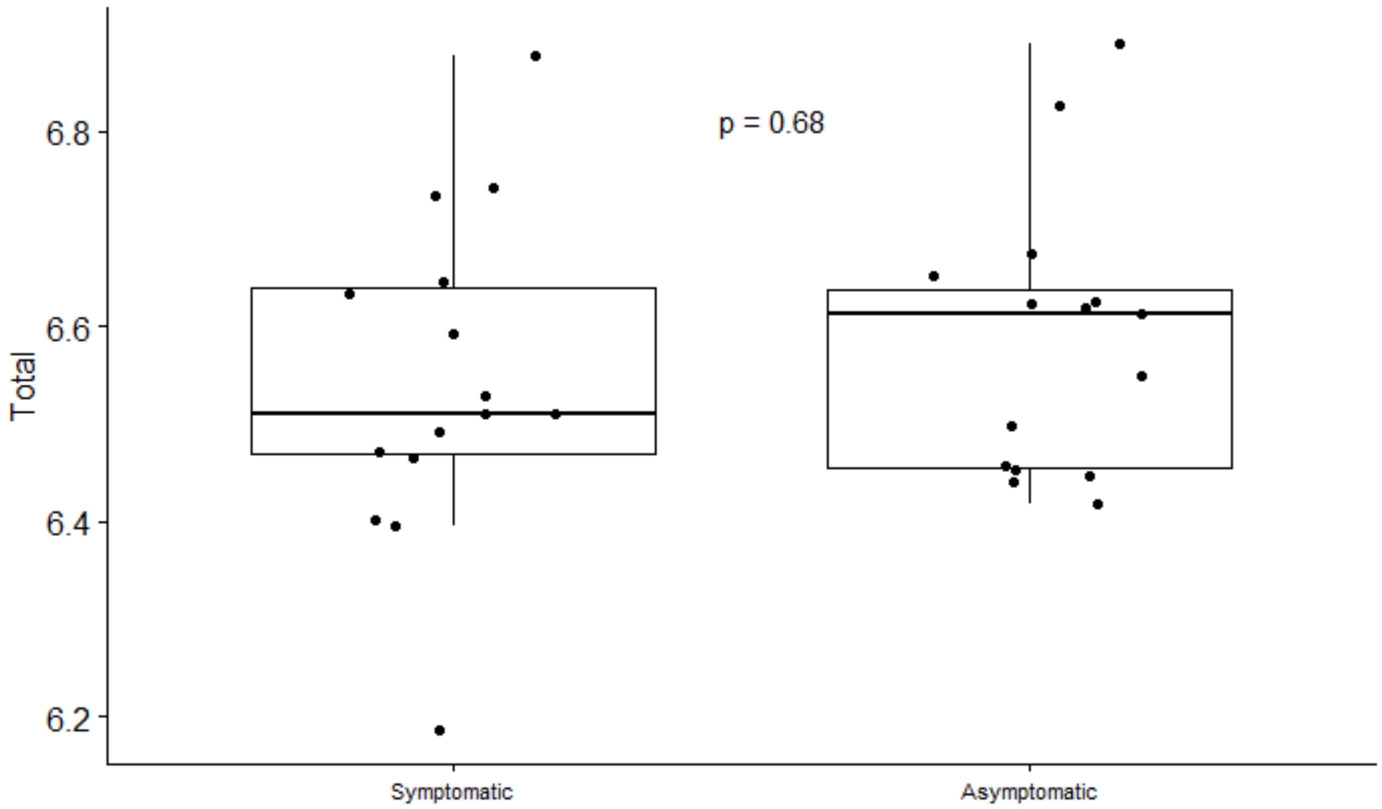
```
flow_log = log10(flow[, 5:ncol(flow) - 1]); flow_log$Group = meta_flow$Group

plot1 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[1], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1, label.y = 5.5) + xlab('') + theme(axis.text.x = element_text(size = 8))
plot2 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[2], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1.5, label.y = 5.75) + xlab('') + theme(axis.text.x = element_text(size = 8))
plot3 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[3], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1, label.y = 6.1) + xlab('') + theme(axis.text.x = element_text(size = 8))
plot6 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[6], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1, label.y = 5.7) + xlab('') + theme(axis.text.x = element_text(size = 8))
plot9 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[9], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1.5, label.y = 4.6) + xlab('') + theme(axis.text.x = element_text(size = 8))
plot11 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[11], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1.75, label.y = 3.4) + xlab('') + theme(axis.text.x = element_text(size = 8))
grid.arrange(plot1, plot2, plot3, plot6, plot9, plot11, ncol = 3, top = textGrob('Flow cytometry at V2', gp = gpar(fontface = 3)), right = textGrob('log10 counts                    log10 counts', rot = 90, gp = gpar(fontsize = 11)))
```


Flow cytometry at V2
Eosinophils
p = 0.12
5.6
5.2
4.8
4.4
Neutrophils
p = 1
6.6
6.3
6.0
5.7
Conventional.T.cells
p = 0.87
6.2
6.0
5.8
5.6
log10 counts
Classical.Monocytes
p = 0.23
5.6
5.4
5.2
B.cells
p = 0.029
5.25
5.00
4.75
4.50
Conventional.DC
p = 0.62
4.0
3.8
3.6
3.4
3.2
log10 counts
Symptomatic
Asymptomatic

```
plot4 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[4], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1, label.y = 4.0) + xlab('') + theme(axis.text.x = element_text(size = 8))
plot5 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[5], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1, label.y = 4.4) + xlab('') + theme(axis.text.x = element_text(size = 8))
plot7 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[7], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1.5, label.y = 4.0) + xlab('') + theme(axis.text.x = element_text(size = 8))
plot8 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[8], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1.5, label.y = 3.8) + xlab('') + theme(axis.text.x = element_text(size = 8))
plot10 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[10], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1, label.y = 4.1) + xlab('') + theme(axis.text.x = element_text(size = 8))
plot12 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[12], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1.5, label.y = 3.3) + xlab('') + theme(axis.text.x = element_text(size = 8))
grid.arrange(plot4, plot5, plot7, plot8, plot10, plot12, ncol = 3, top = textGrob('Flow cytometry at V2', gp = gpar(fontface = 3)), right = textGrob('log10 counts                    log10 counts', rot = 90, gp = gpar(fontsize = 11)))
```

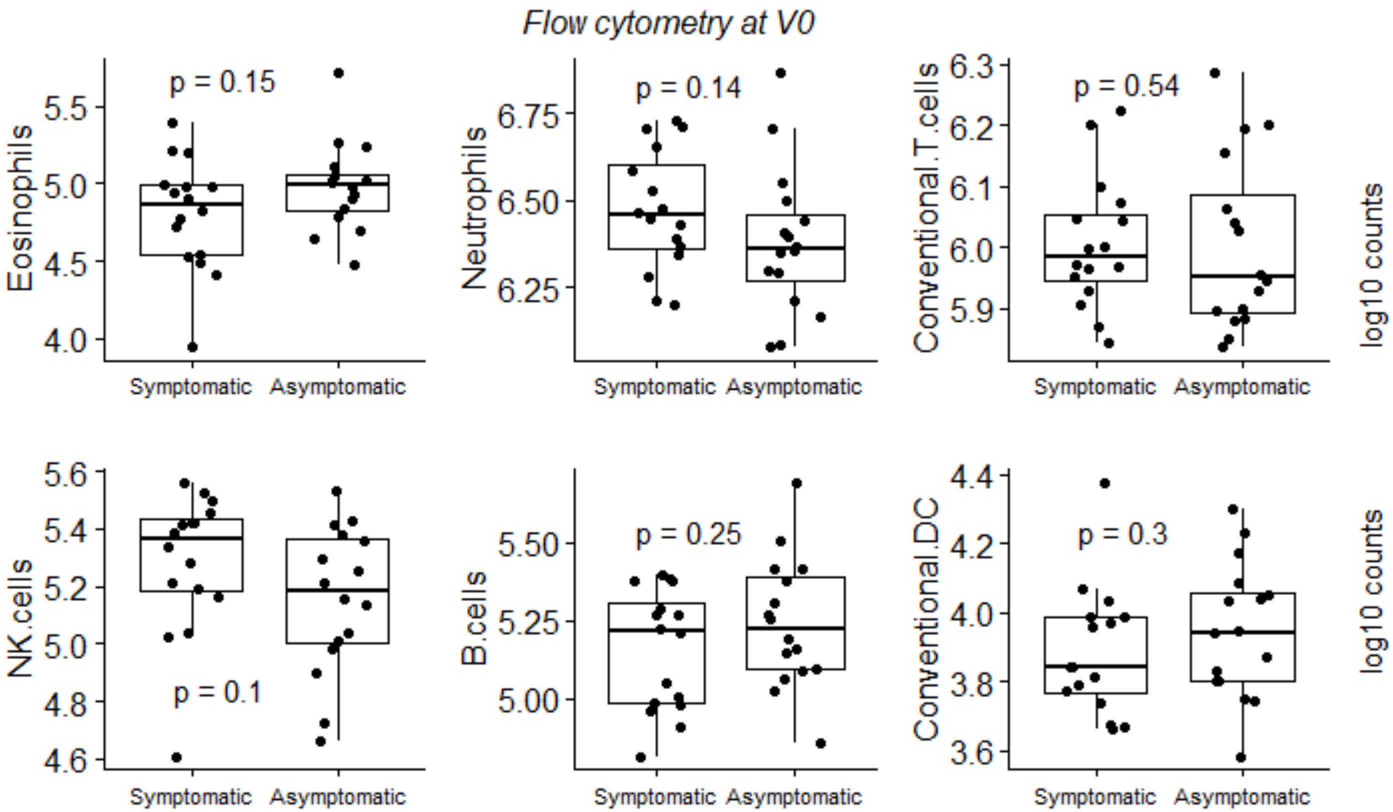


```
ggboxplot(flow_log, x = 'Group', y = 'Total', add = 'jitter') + stat_compare_means(label = "p.format", label.x =
1.5, label.y = 6.8) + xlab('') + theme(axis.text.x = element_text(size = 8))
```

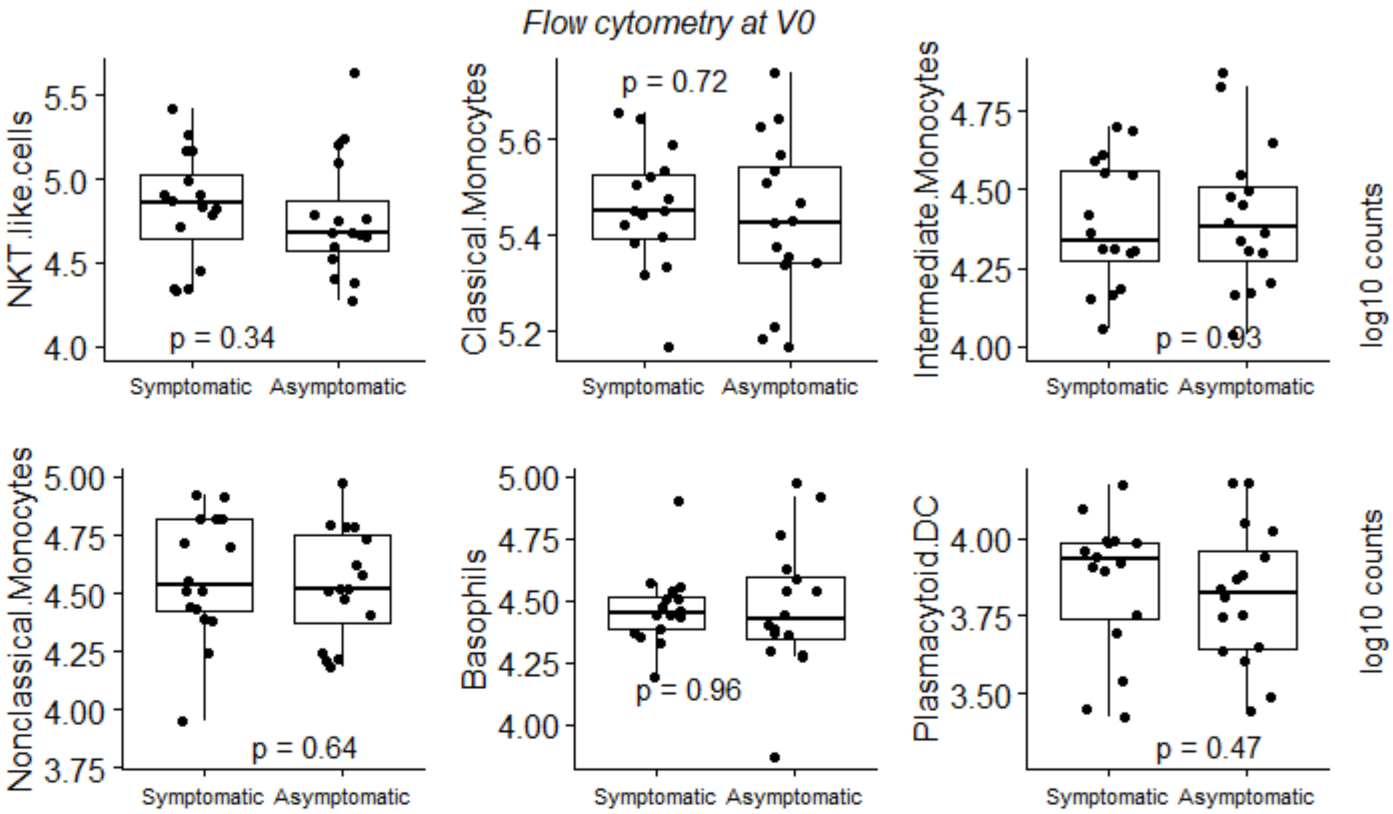


```
flow = read.csv('data/COVIRS_flow_data_20230202.csv'); flow = subset(flow, Patient %in% meta_data$Participant.ID
& Timepoint == 'V0'); meta_flow = subset(meta_data, Participant.ID %in% flow$Patient); flow = flow[order(flow$Pat
ient), ]; flow$Total = rowSums(flow[, 5:ncol(flow)-1]); flow$Group = meta_flow$Group
```

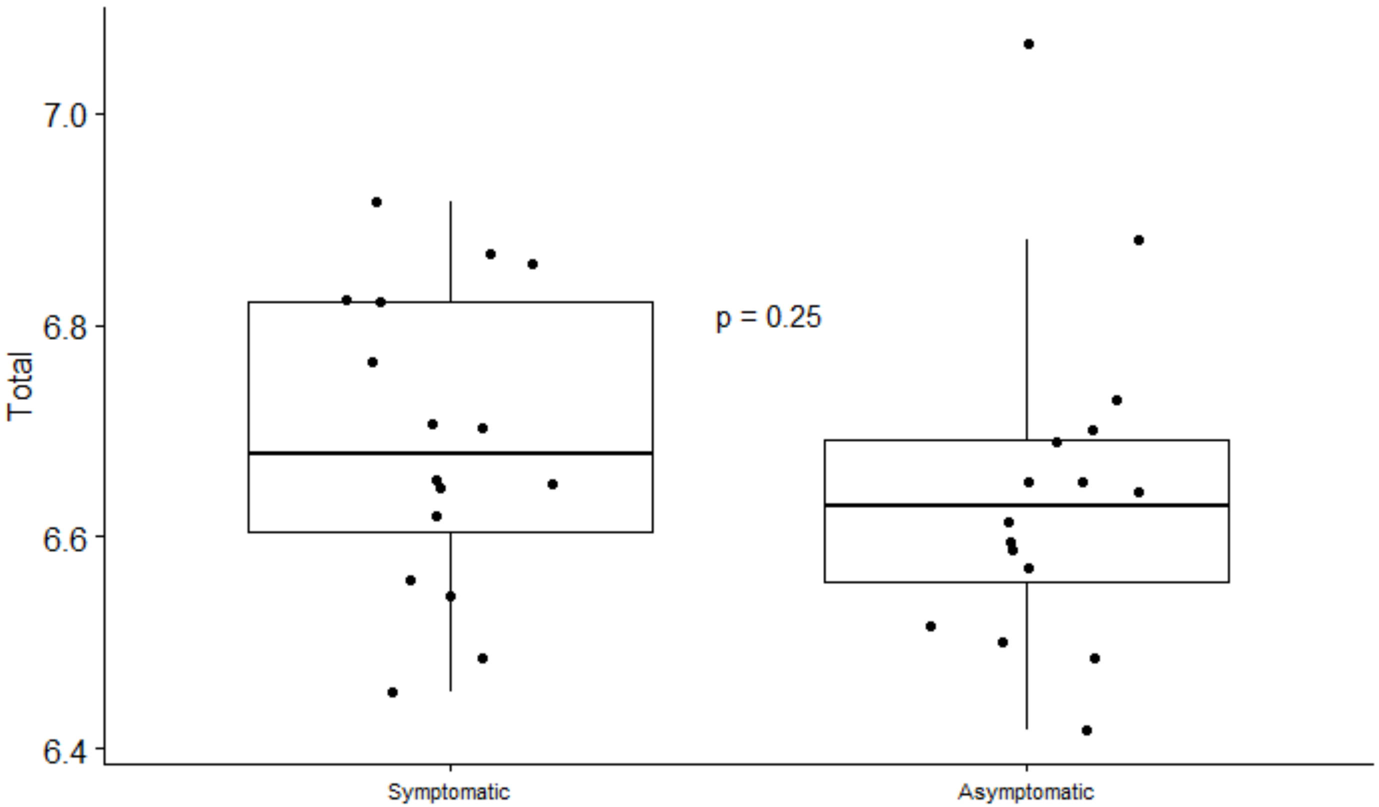
```
flow_log = log10(flow[, 5:ncol(flow) - 1]); flow_log$Group = meta_flow$Group

plot1 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[1], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1, label.y = 5.6) + xlab('') + theme(axis.text.x = element_text(size = 8))
plot2 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[2], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1, label.y = 6.8) + xlab('') + theme(axis.text.x = element_text(size = 8))
plot3 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[3], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1, label.y = 6.25) + xlab('') + theme(axis.text.x = element_text(size = 8))
plot5 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[5], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1, label.y = 4.8) + xlab('') + theme(axis.text.x = element_text(size = 8))
plot9 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[9], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1, label.y = 5.5) + xlab('') + theme(axis.text.x = element_text(size = 8))
plot11 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[11], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1, label.y = 4.2) + xlab('') + theme(axis.text.x = element_text(size = 8))
grid.arrange(plot1, plot2, plot3, plot5, plot9, plot11, ncol = 3, top = textGrob('Flow cytometry at V0', gp = gpar(fontface = 3)), right = textGrob('log10 counts                    log10 counts', rot = 90, gp = gpar(fontsize = 11)))
```


Flow cytometry at V0
Eosinophils
p = 0.15
5.5
5.0
4.5
4.0
Symptomatic
Asymptomatic
Neutrophils
p = 0.14
6.75
6.50
6.25
Conventional.T.cells
p = 0.54
6.3
6.2
6.1
6.0
5.9
log10 counts
NK.cells
p = 0.1
5.6
5.4
5.2
5.0
4.8
4.6
B.cells
p = 0.25
5.50
5.25
5.00
Conventional.DC
p = 0.3
4.4
4.2
4.0
3.8
3.6
log10 counts

```
plot4 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[4], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1, label.y = 4.0) + xlab('') + theme(axis.text.x = element_text(size = 8))
plot6 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[6], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1, label.y = 5.7) + xlab('') + theme(axis.text.x = element_text(size = 8))
plot7 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[7], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1.5, label.y = 4.0) + xlab('') + theme(axis.text.x = element_text(size = 8))
plot8 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[8], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1.5, label.y = 3.8) + xlab('') + theme(axis.text.x = element_text(size = 8))
plot10 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[10], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1, label.y = 4.1) + xlab('') + theme(axis.text.x = element_text(size = 8))
plot12 = ggboxplot(flow_log, x = 'Group', y = colnames(flow_log)[12], add = 'jitter') + stat_compare_means(label = "p.format", label.x = 1.5, label.y = 3.3) + xlab('') + theme(axis.text.x = element_text(size = 8))
grid.arrange(plot4, plot6, plot7, plot8, plot10, plot12, ncol = 3, top = textGrob('Flow cytometry at V0', gp = gpar(fontface = 3)), right = textGrob('log10 counts                    log10 counts', rot = 90, gp = gpar(fontsize = 11)))
```



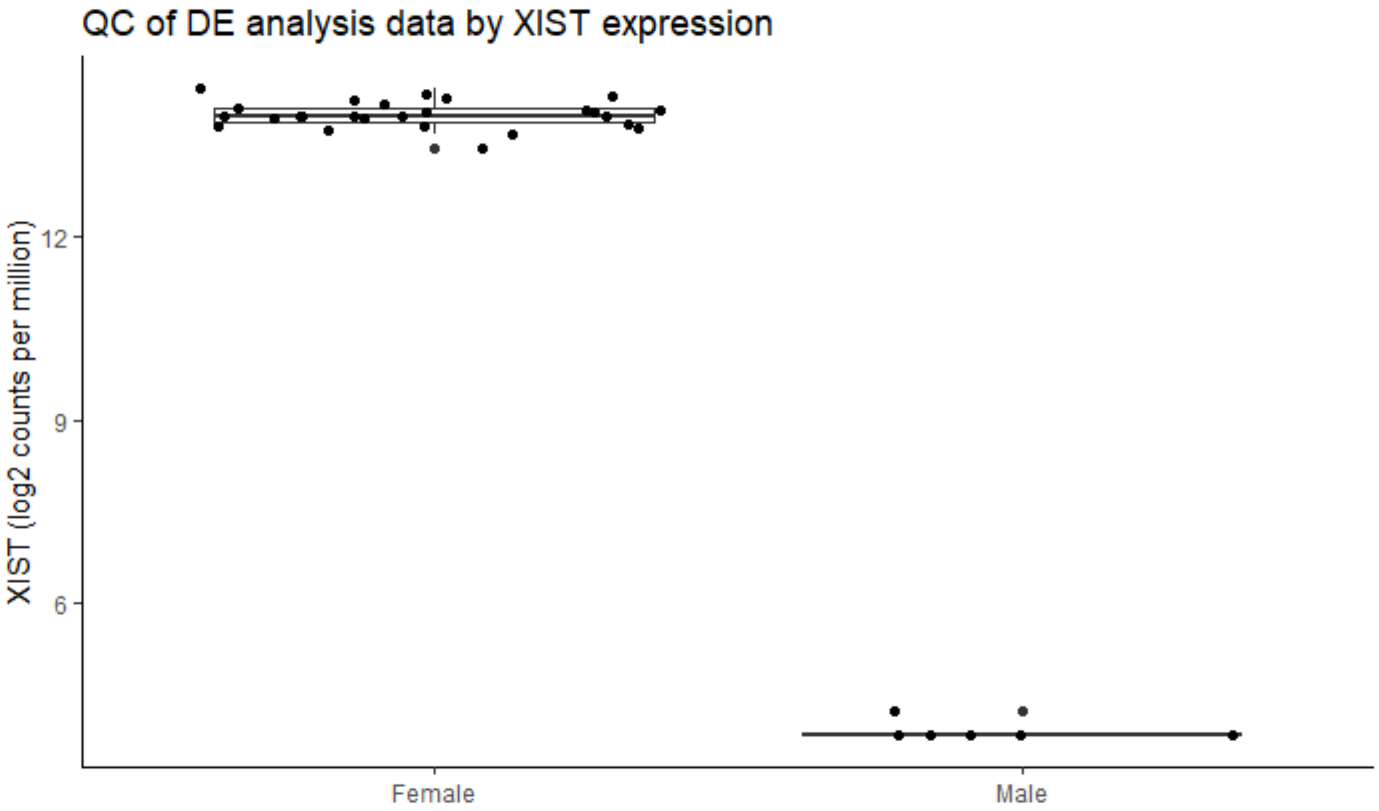
```
ggboxplot(flow_log, x = 'Group', y = 'Total', add = 'jitter') + stat_compare_means(label = "p.format", label.x =
1.5, label.y = 6.8) + xlab('') + theme(axis.text.x = element_text(size = 8))
```



```
rm(flow); rm(meta_flow); rm(flow_log); rm(plot4); rm(plot6); rm(plot7); rm(plot8); rm(plot10); rm(plot12); rm(plo
t1); rm(plot2); rm(plot3); rm(plot5); rm(plot9); rm(plot11)
```

4. V2 gene set enrichment analysis

```
dds = DESeqDataSetFromMatrix(countData = count_data,
                              colData = meta_data,
                              design = ~ Sex, tidy = TRUE)
vsd = vst(dds, blind = FALSE); vsd.df = assay(vsd)
vsd.df = data.frame(t(vsd.df))
vsd.df$Group = meta_data$Group; vsd.df$Sex = meta_data$Sex
ggplot(vsd.df, aes(x = Sex, y = vsd.df[, 'ENSG00000229807'])) + geom_boxplot() + geom_jitter() + ylab('XIST (log2 counts per million)') + theme_classic() + xlab('') + ggtitle('QC of DE analysis data by XIST expression')
```



```
dds = DESeqDataSetFromMatrix(countData = count_data,
                              colData = meta_data,
                              design = ~ Group + Sex, tidy = TRUE)
```

```
vsd = vst(dds, blind = FALSE); vsd.df = assay(vsd)
write.csv(vsd.df, file = 'results/counts_norm_v2.csv', row.names = TRUE)
```

```
keep = rowSums(counts(dds)>100) >= 4; dds = dds[keep, ]; dds = DESeq(dds)
```

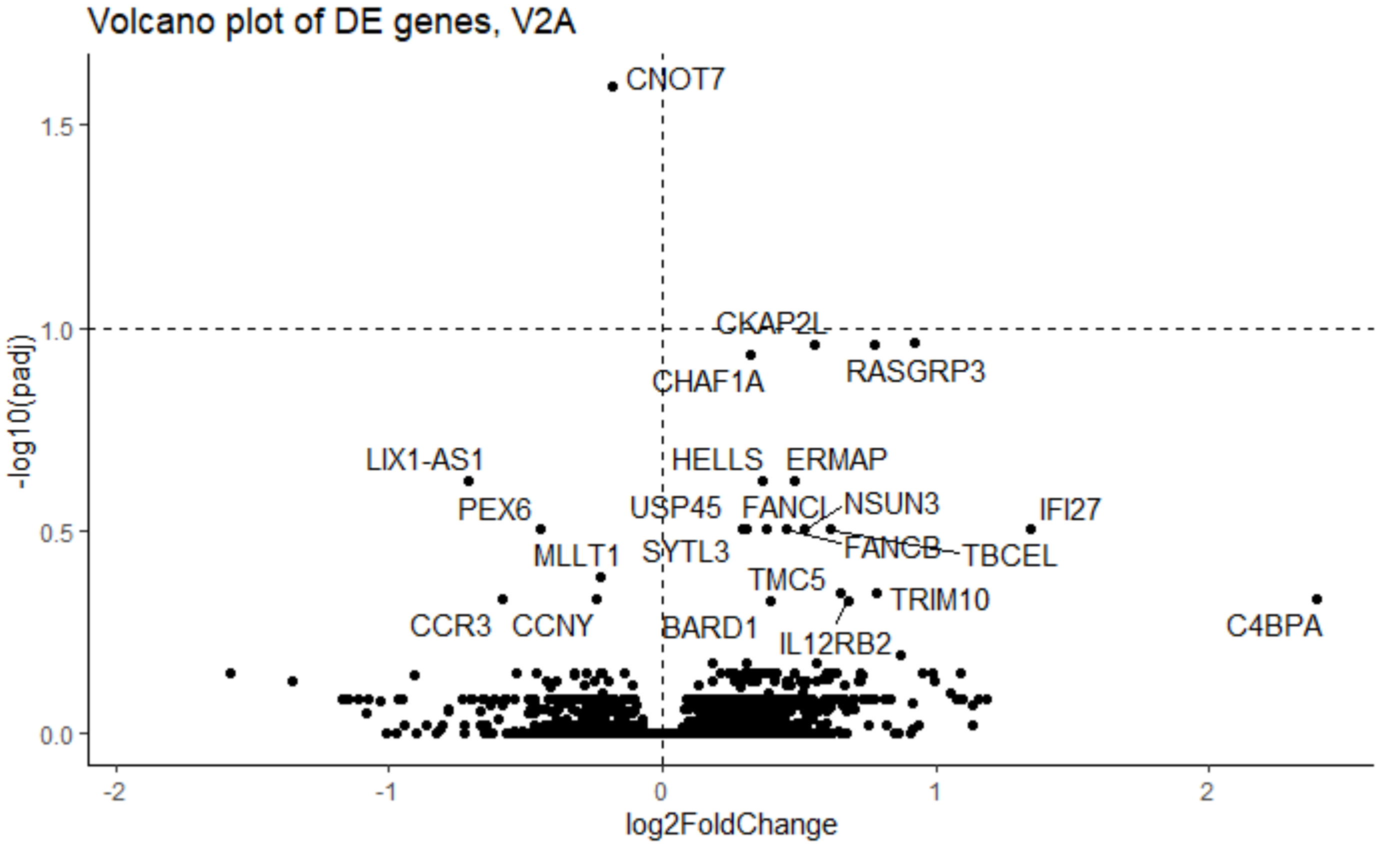
```
estimating size factors
estimating dispersions
gene-wise dispersion estimates
mean-dispersion relationship
final dispersion estimates
fitting model and testing
-- replacing outliers and refitting for 20 genes
-- DESeq argument 'minReplicatesForReplace' = 7
-- original counts are preserved in counts(dds)
estimating dispersions
fitting model and testing
```

```
res = results(dds, contrast = c("Group", "Symptomatic", "Asymptomatic")); res = res[order(res$padj, decreasing =
FALSE), ]
res; summary(res)
```

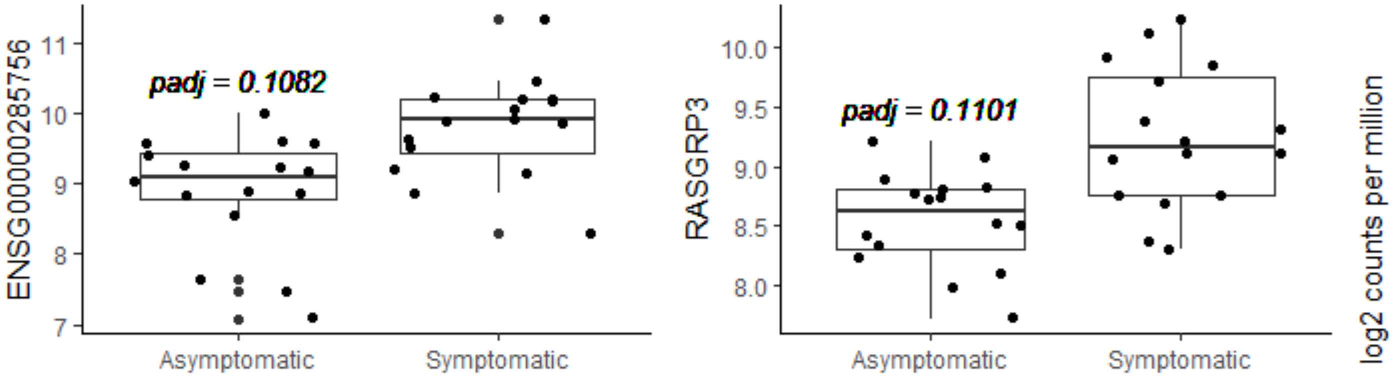
```
log2 fold change (MLE): Group Symptomatic vs Asymptomatic
Wald test p-value: Group Symptomatic vs Asymptomatic
DataFrame with 11403 rows and 6 columns
                 baseMean log2FoldChange     lfcSE         stat      pvalue      padj
                <numeric>      <numeric> <numeric>    <numeric>   <numeric> <numeric>
ENSG00000198791 1120.7692      -0.183659 0.0388160     -4.73151 2.22853e-06  0.025403
ENSG00000285756  750.6775       0.923212 0.2158807      4.27649 1.89861e-05  0.108211
ENSG00000152689  499.4901       0.774705 0.1856994      4.17182 3.02170e-05  0.110068
ENSG00000169607   73.5999       0.560044 0.1360796      4.11556 3.86239e-05  0.110068
ENSG00000167670  653.8076       0.322582 0.0796486      4.05006 5.12048e-05  0.116737
...                   ...            ...       ...          ...         ...       ...
ENSG00000089048   319.702    3.87279e-06 0.0633473  6.11359e-05    0.999951  0.999982
ENSG00000207005  1044.569   -1.88696e+00 0.6090192 -3.09836e+00          NA        NA
ENSG00000184500   222.258   -1.78817e-01 0.3468684 -5.15518e-01          NA        NA
ENSG00000078114    25.276   -1.59377e+00 0.5449411 -2.92467e+00          NA        NA
ENSG00000107566   470.830    8.18822e-03 0.2855543  2.86748e-02          NA        NA

out of 11403 with nonzero total read count
adjusted p-value < 0.1
LFC > 0 (up)       : 0, 0%
LFC < 0 (down)     : 1, 0.0088%
outliers [1]       : 4, 0.035%
low counts [2]     : 0, 0%
(mean count < 25)
[1] see 'cooksCutoff' argument of ?results
[2] see 'independentFiltering' argument of ?results
```

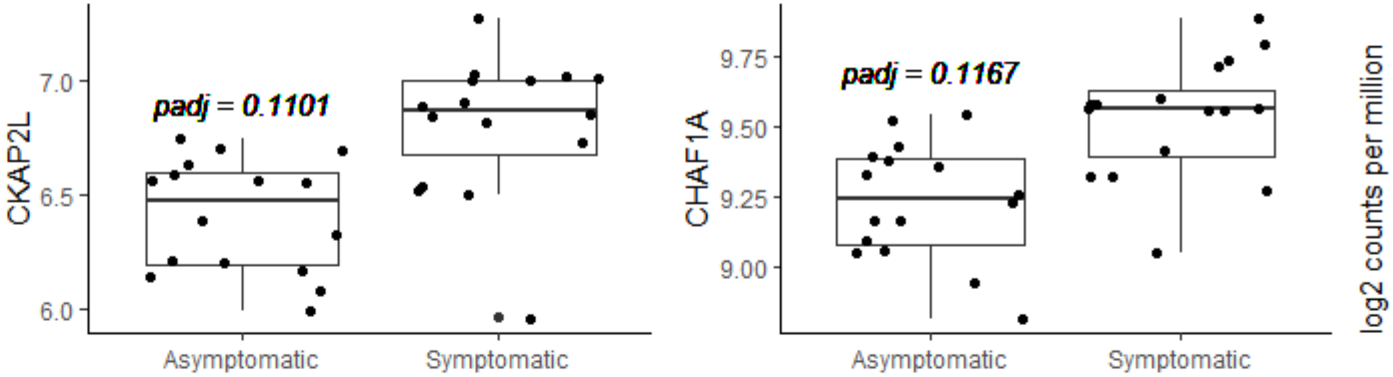
```
res.df = data.frame(res); res.df = data.frame(res.df, ens2gene[rownames(res.df), ])
write.csv(res.df, file = 'results/res_deseq_v2.csv', row.names = TRUE)
rm(keep)
```

```
plot = subset(res.df, padj < 0.8)
ggplot(res.df, aes(x = log2FoldChange, y = -log10(padj))) + geom_point() +
  geom_hline(yintercept = -log10(0.1), linetype = 2) + geom_vline(xintercept = 0, linetype = 2) +
  theme_classic() + geom_text_repel(max.overlaps = 13, data = subset(plot, -log10(padj) > 0.2), aes(x = log2FoldC
hange, y = -log10(padj), label = external_gene_name)) + ggtitle("Volcano plot of DE genes, V2A"); rm(plot)
```

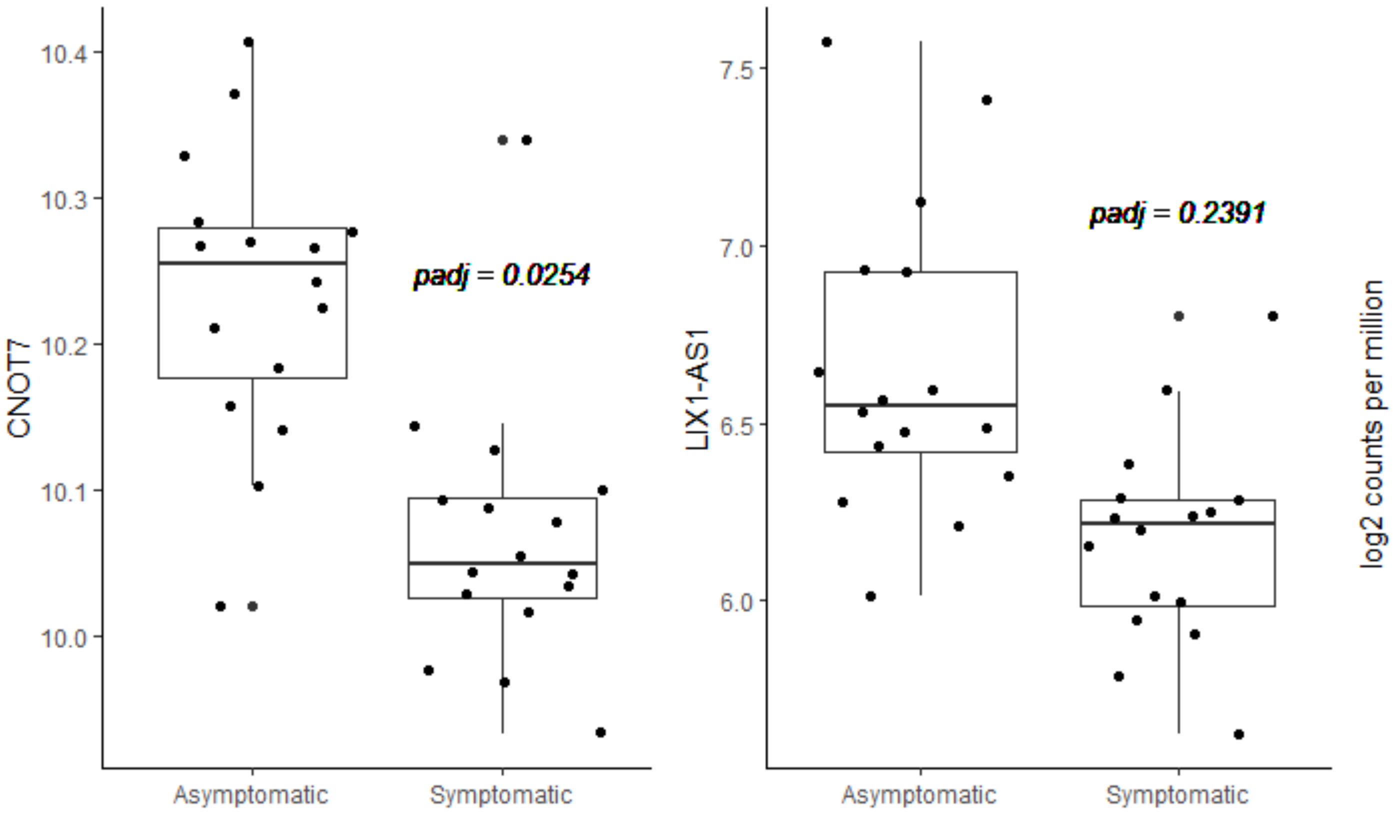


```
plot = rbind(filter(res.df, log2FoldChange > 0)[1:4, ], filter(res.df, log2FoldChange < 0)[1:2, ])
vsd.df = data.frame(t(vsd.df)); vsd.df$Group = meta_data$Group; vsd.df$Sex = meta_data$Sex

plot1 = ggplot(vsd.df, aes(x = Group, y = vsd.df[, rownames(plot)[1]])) + geom_boxplot() + geom_jitter() + ylab(res.df[rownames(plot)[1], 'ensembl_gene_id']) + geom_text(aes(x = 1, y = 10.5, fontface = 3), label = paste0('padj = ', as.character(round(res.df[rownames(plot)[1], 'padj'], 4)))) + theme_classic() + xlab('')
plot2 = ggplot(vsd.df, aes(x = Group, y = vsd.df[, rownames(plot)[2]])) + geom_boxplot() + geom_jitter() + ylab(res.df[rownames(plot)[2], 'external_gene_name']) + geom_text(aes(x = 1, y = 9.5, fontface = 3), label = paste0('padj = ', as.character(round(res.df[rownames(plot)[2], 'padj'], 4)))) + theme_classic() + xlab('')
plot3 = ggplot(vsd.df, aes(x = Group, y = vsd.df[, rownames(plot)[3]])) + geom_boxplot() + geom_jitter() + ylab(res.df[rownames(plot)[3], 'external_gene_name']) + geom_text(aes(x = 1, y = 6.9, fontface = 3), label = paste0('padj = ', as.character(round(res.df[rownames(plot)[3], 'padj'], 4)))) + theme_classic() + xlab('')
plot4 = ggplot(vsd.df, aes(x = Group, y = vsd.df[, rownames(plot)[4]])) + geom_boxplot() + geom_jitter() + ylab(res.df[rownames(plot)[4], 'external_gene_name']) + geom_text(aes(x = 1, y = 9.7, fontface = 3), label = paste0('padj = ', as.character(round(res.df[rownames(plot)[4], 'padj'], 4)))) + theme_classic() + xlab('')
plot5 = ggplot(vsd.df, aes(x = Group, y = vsd.df[, rownames(plot)[5]])) + geom_boxplot() + geom_jitter() + ylab(res.df[rownames(plot)[5], 'external_gene_name']) + geom_text(aes(x = 2, y = 10.25, fontface = 3), label = paste0('padj = ', as.character(round(res.df[rownames(plot)[5], 'padj'], 4)))) + theme_classic() + xlab('')
plot6 = ggplot(vsd.df, aes(x = Group, y = vsd.df[, rownames(plot)[6]])) + geom_boxplot() + geom_jitter() + ylab(res.df[rownames(plot)[6], 'external_gene_name']) + geom_text(aes(x = 2, y = 7.1, fontface = 3), label = paste0('padj = ', as.character(round(res.df[rownames(plot)[6], 'padj'], 4)))) + theme_classic() + xlab('')
grid.arrange(plot1, plot2, plot3, plot4, ncol = 2, right = textGrob('log2 counts per million              log2 counts per million', rot = 90, gp = gpar(fontsize = 11)))
```



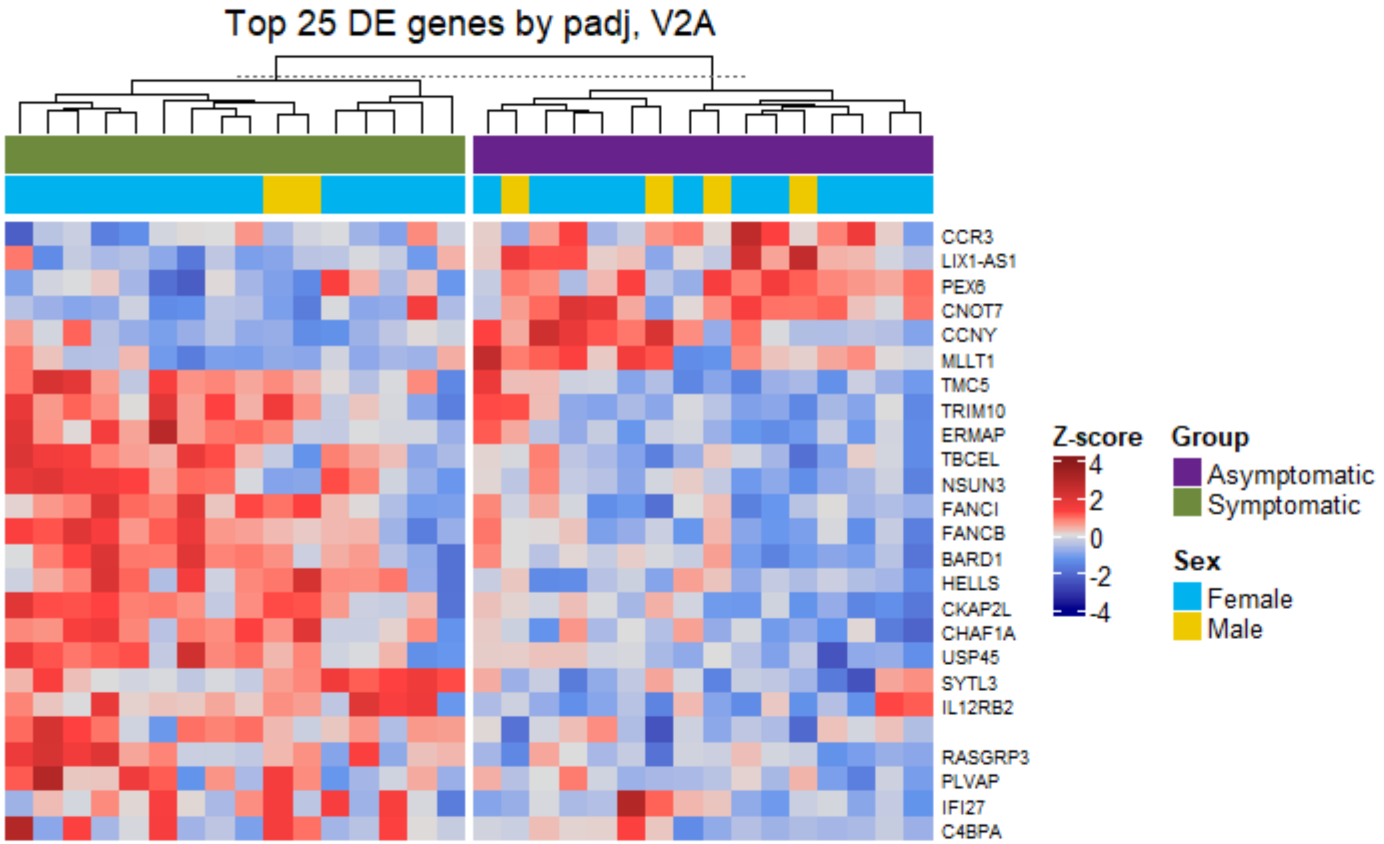


```
grid.arrange(plot5, plot6, ncol = 2, right = textGrob('log2 counts per million', rot = 90, gp = gpar(fontsize = 1
1)))
```



```
vsd.df$Sex = NULL; vsd.df$Group = NULL; vsd.df = data.frame(t(vsd.df))
rm(plot); rm(plot1); rm(plot2); rm(plot3); rm(plot4); rm(plot5); rm(plot6)
```

```
ha_top = HeatmapAnnotation(df = meta_data[, c('Group', 'Sex')], col = list(Group = c('Symptomatic' = 'darkolivegreen4', 'Asymptomatic' = 'darkorchid4'), Sex = c('Male' = 'gold2', 'Female' = 'deepskyblue2')), show_annotation_name = FALSE)
zzz = t(scale(t(vsd.df[rownames(res.df[1:25, ]), ]))); zzz = pmin(pmax(zzz, -3), 3); rownames(zzz) =  res.df[rownames(zzz), 'external_gene_name']
Heatmap(zzz, show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = TRUE, show_row_dend = FALSE, top_annotation = ha_top, heatmap_legend_param = list(title = "Z-score"), col = colorRamp2(c(-4, -1.25, 0, 1.25, 4), c("darkblue", "cornflowerblue", "gainsboro", "brown1", "firebrick4")), row_names_gp = gpar(fontsize = 7), column_title = 'Top 25 DE genes by padj, V2A', column_split = meta_data$Group)
```



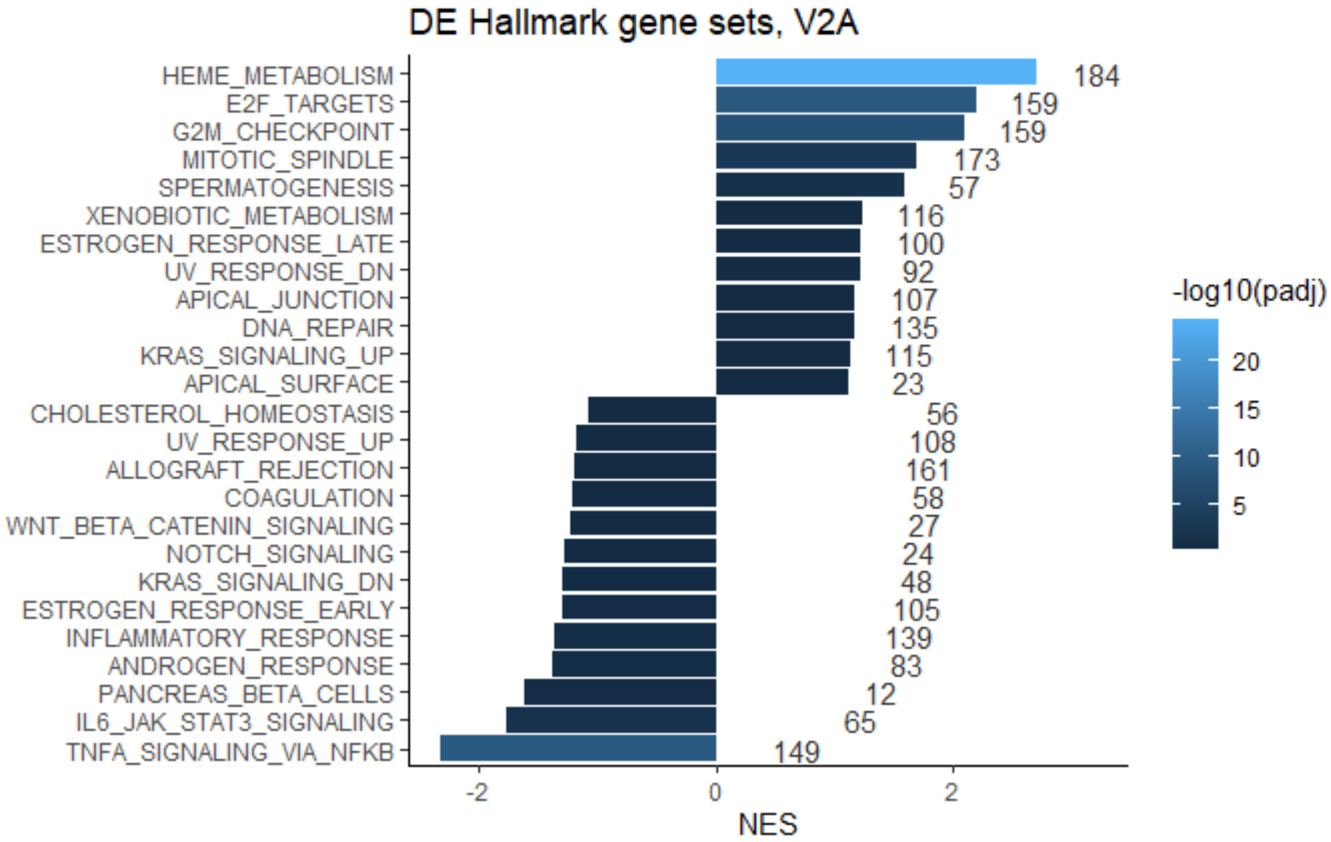
```
rm(ha_top); rm(zzz)
```

```
ranks <- res.df %>%
  as_tibble() %>%
  dplyr::select(external_gene_name, log2FoldChange) %>%
  na.omit() %>%
  distinct() %>%
  group_by(external_gene_name) %>%
  summarize(log2FoldChange=mean(log2FoldChange))
ranks <- deframe(ranks)
```

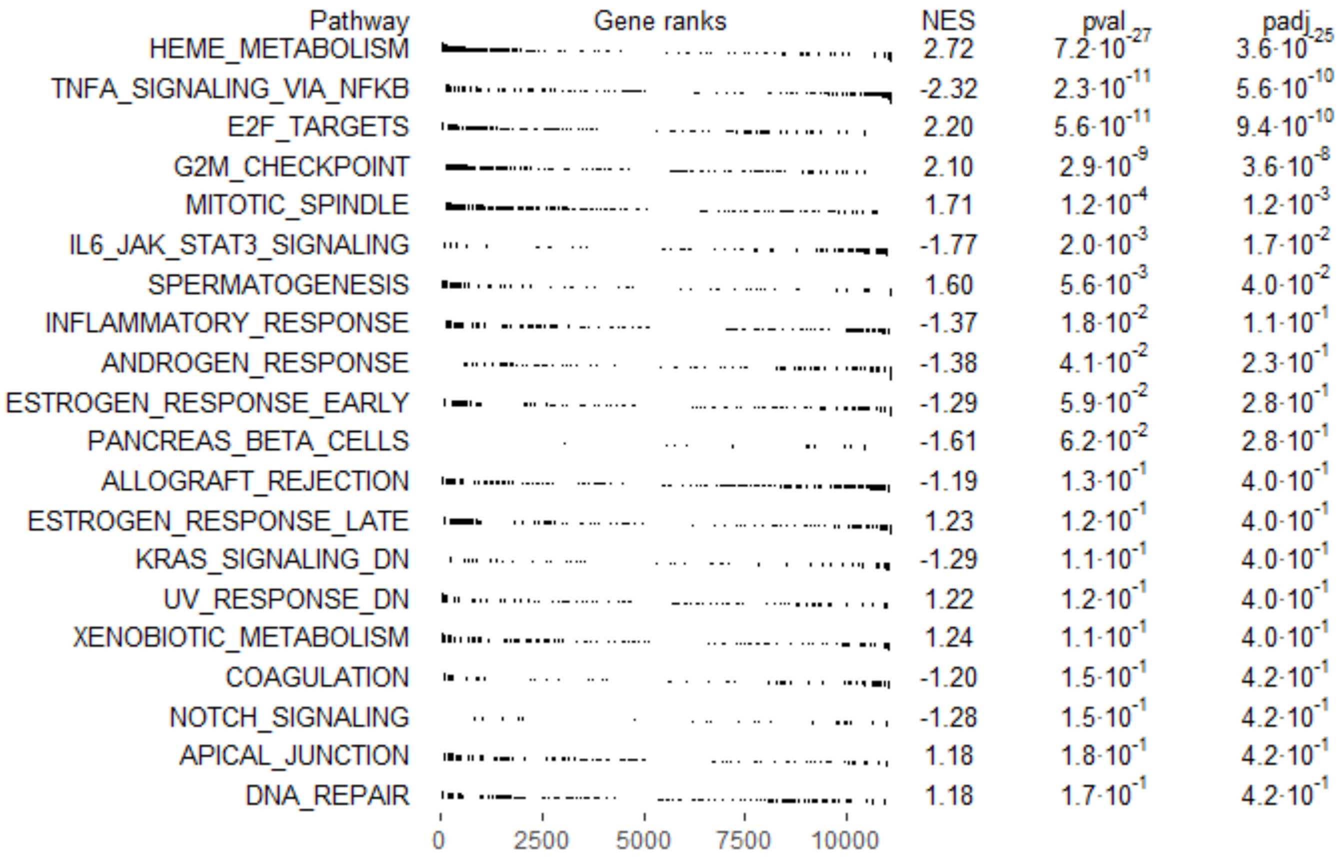
```
msig.df = as.data.frame(msigdbr(species = "Homo sapiens", category = "H"))
msigDB = by(msig.df$gene_symbol, msig.df$gs_name, function(x) as.character(x))
names(msigDB) = gsub('HALLMARK_', '', names(msigDB))
fgseaRes = fgsea(pathways = msigDB, stats = ranks, nproc = 1); fgseaRes = fgseaRes[order(fgseaRes$padj, decreasing = FALSE), ]
```

```
write.csv(subset(fgseaRes, select = -c(leadingEdge)), file = 'results/fgseaRes_H_v2.csv', row.names = FALSE)
```

```
ggplot(fgseaRes[1:25, ], aes(y = reorder(pathway, NES), x = NES, fill = -log10(padj), label = size)) +
  geom_col() + geom_text(color = 'gray20', nudge_x = ifelse(fgseaRes$NES > 0, 0.5, 3)) +
  ylab('') + theme_classic() + ggtitle('DE Hallmark gene sets, V2A')
```

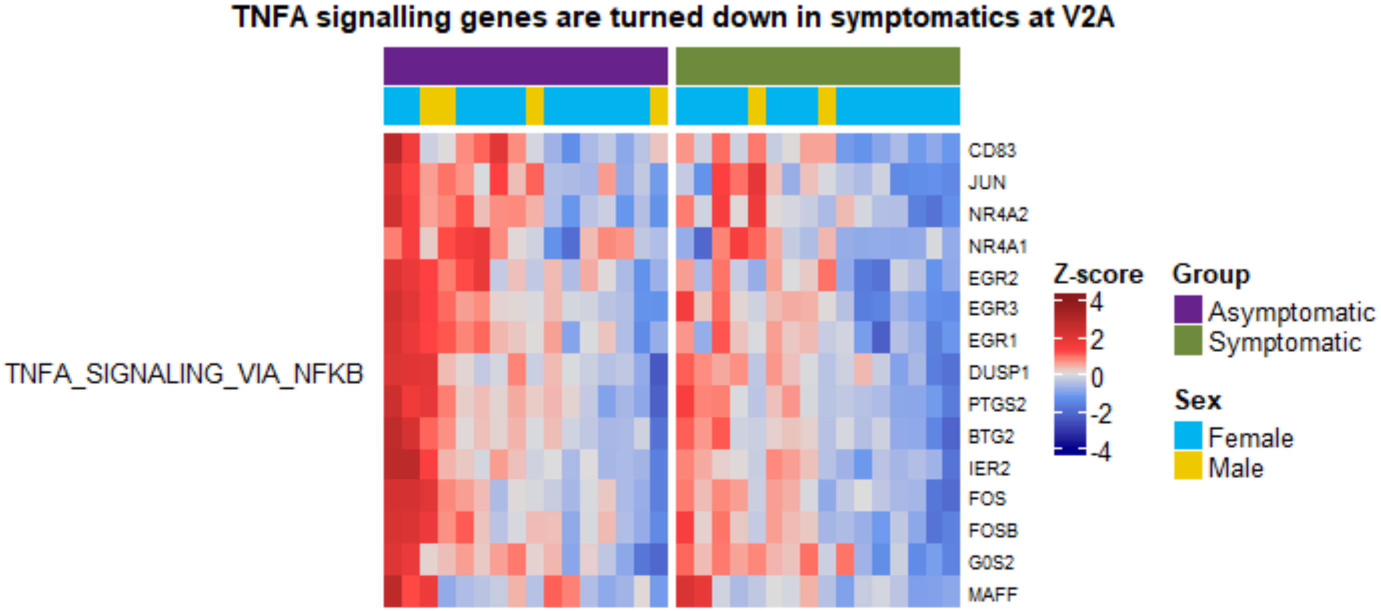


```
plotGseaTable(msigDB[fgseaRes[1:20, ][, pathway]], ranks, fgseaRes[1:20, ], gseaParam = 0.5, colwidths = c(3, 3,
0.8, 1.2, 1.2), pathwayLabelStyle = list(size = 10), headerLabelStyle = list(size = 10), valueStyle = list(size =
10))
```

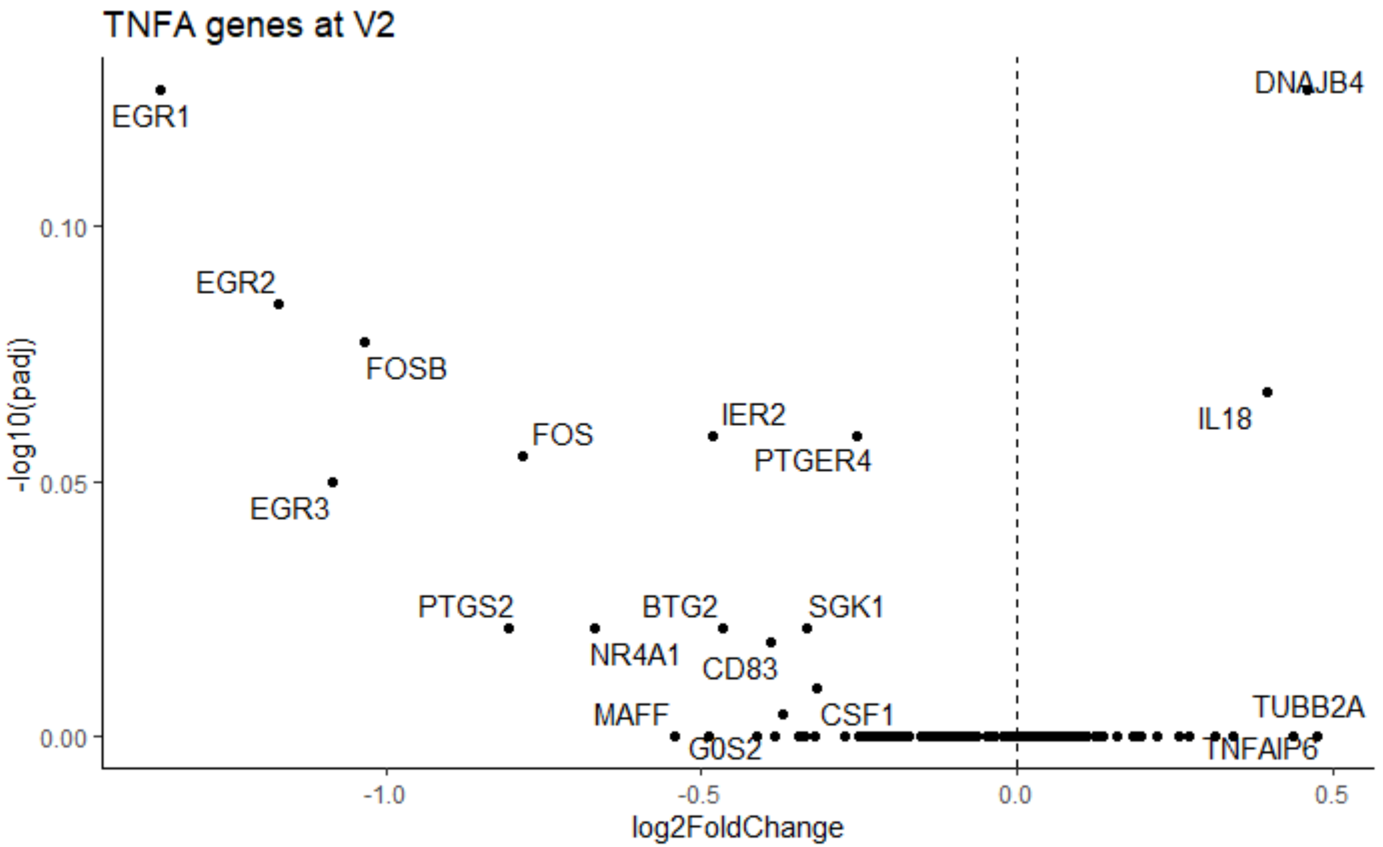


Pathway	Gene ranks	NES	pval	padj
HEME_METABOLISM		2.72	$7.2 \cdot 10^{-27}$	$3.6 \cdot 10^{-25}$
TNFA_SIGNALING_VIA_NFKB		-2.32	$2.3 \cdot 10^{-11}$	$5.6 \cdot 10^{-10}$
E2F_TARGETS		2.20	$5.6 \cdot 10^{-11}$	$9.4 \cdot 10^{-10}$
G2M_CHECKPOINT		2.10	$2.9 \cdot 10^{-9}$	$3.6 \cdot 10^{-8}$
MITOTIC_SPINDLE		1.71	$1.2 \cdot 10^{-4}$	$1.2 \cdot 10^{-3}$
IL6_JAK_STAT3_SIGNALING		-1.77	$2.0 \cdot 10^{-3}$	$1.7 \cdot 10^{-2}$
SPERMATOGENESIS		1.60	$5.6 \cdot 10^{-3}$	$4.0 \cdot 10^{-2}$
INFLAMMATORY_RESPONSE		-1.37	$1.8 \cdot 10^{-2}$	$1.1 \cdot 10^{-1}$
ANDROGEN_RESPONSE		-1.38	$4.1 \cdot 10^{-2}$	$2.3 \cdot 10^{-1}$
ESTROGEN_RESPONSE_EARLY		-1.29	$5.9 \cdot 10^{-2}$	$2.8 \cdot 10^{-1}$
PANCREAS_BETA_CELLS		-1.61	$6.2 \cdot 10^{-2}$	$2.8 \cdot 10^{-1}$
ALLOGRAFT_REJECTION		-1.19	$1.3 \cdot 10^{-1}$	$4.0 \cdot 10^{-1}$
ESTROGEN_RESPONSE_LATE		1.23	$1.2 \cdot 10^{-1}$	$4.0 \cdot 10^{-1}$
KRAS_SIGNALING_DN		-1.29	$1.1 \cdot 10^{-1}$	$4.0 \cdot 10^{-1}$
UV_RESPONSE_DN		1.22	$1.2 \cdot 10^{-1}$	$4.0 \cdot 10^{-1}$
XENOBIOTIC_METABOLISM		1.24	$1.1 \cdot 10^{-1}$	$4.0 \cdot 10^{-1}$
COAGULATION		-1.20	$1.5 \cdot 10^{-1}$	$4.2 \cdot 10^{-1}$
NOTCH_SIGNALING		-1.28	$1.5 \cdot 10^{-1}$	$4.2 \cdot 10^{-1}$
APICAL_JUNCTION		1.18	$1.8 \cdot 10^{-1}$	$4.2 \cdot 10^{-1}$
DNA_REPAIR		1.18	$1.7 \cdot 10^{-1}$	$4.2 \cdot 10^{-1}$

```
res.df = res.df[order(res.df$log2FoldChange, decreasing = FALSE), ]
ha_top = HeatmapAnnotation(df = meta_data[, c('Group', 'Sex')], col = list(Group = c('Symptomatic' = 'darkolivegreen4', 'Asymptomatic' = 'darkorchid4'), Sex = c('Male' = 'gold2', 'Female' = 'deepskyblue2')), show_annotation_name = FALSE)
plot2 = vsd.df[rownames(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[2, 'pathway']]]] & log2FoldChange < 0))[1:15], ]
zzz = t(scale(t(plot2))); zzz = pmin(pmax(zzz, -3), 3); rownames(zzz) =  res.df[rownames(zzz), 'external_gene_name']
plot2 = Heatmap(zzz, row_title = fgseaRes[[2, 'pathway']], show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = TRUE, show_row_dend = FALSE, show_column_dend = FALSE, column_split = meta_data$Group, top_annotation = ha_top, heatmap_legend_param = list(title = "Z-score"), col = colorRamp2(c(-4, -1.25, 0, 1.25, 4), c("darkblue", "cornflowerblue", "gainsboro", "brown1", "firebrick4")), row_names_gp = gpar(fontsize = 7), row_title_rot = 0, row_title_gp = gpar(fontsize = 10), column_title_gp = gpar(fontsize = 11, fontface = 2), column_title = 'TNFA signalling genes are turned down in symptomatics at V2A', heatmap_height = unit(8, "cm")); plot2
```



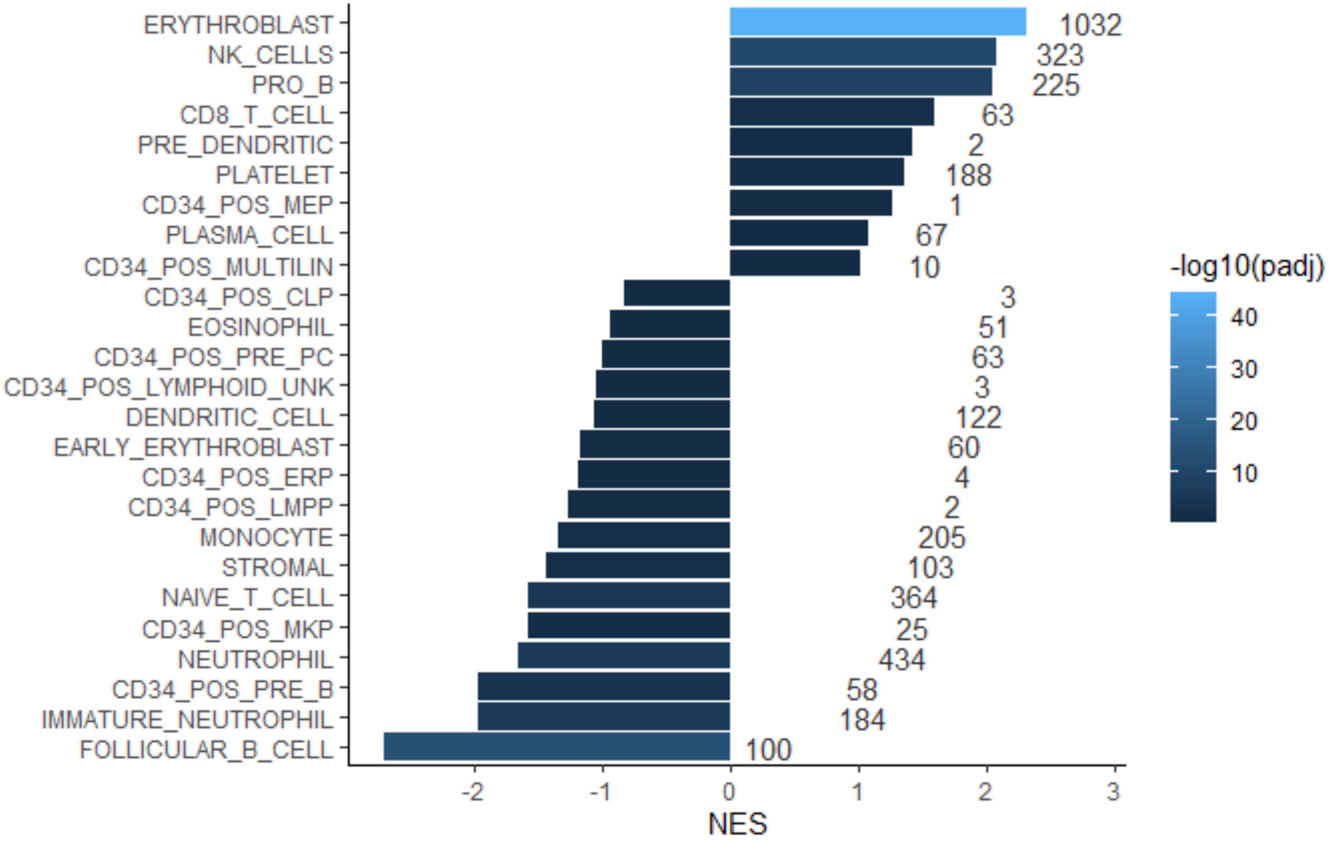
```
rm(plot2); rm(ha_top); rm(zzz)

ggplot(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[2, 'pathway']]]]), aes(x = log2FoldChange, y = -l
og10(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + geom_vline(xinter
cept = 0, linetype = 2) + ggtitle('TNFA genes at V2')
```



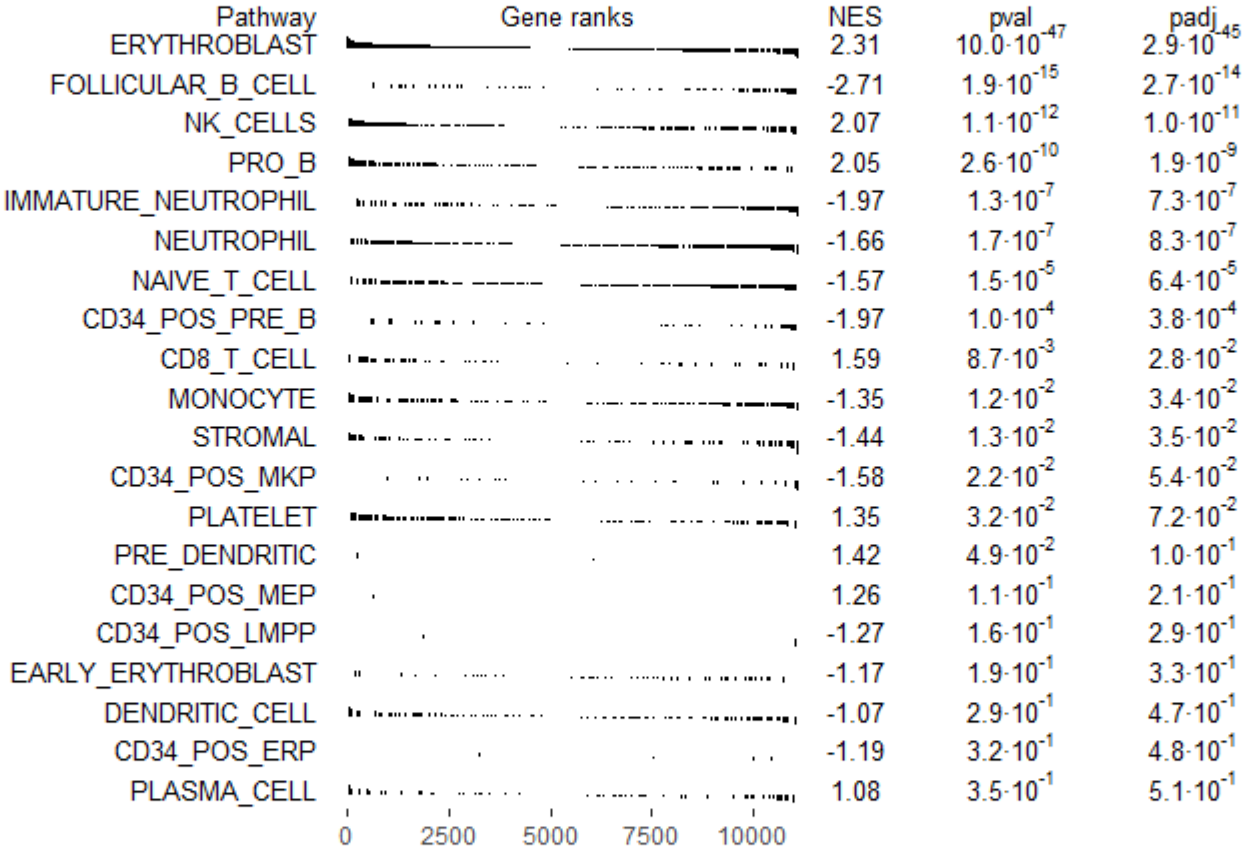
```
msig.df = as.data.frame(msigdbr(species = "Homo sapiens", category = "C8")); msig.df = msig.df[grepl('HAY_BONE_MARROW', msig.df$gs_name), ]
msigDB = by(msig.df$gene_symbol, msig.df$gs_name, function(x) as.character(x)); names(msigDB) = gsub('HAY_BONE_MARROW_', '', names(msigDB))
fgseaRes = fgsea(pathways = msigDB, stats = ranks, nproc = 1); fgseaRes = fgseaRes[order(fgseaRes$padj, decreasing = FALSE), ]
```

```
write.csv(subset(fgseaRes, select = -c(leadingEdge)), file = 'results/fgseaRes_C8_marrow_v2.csv', row.names = FAL
SE)

ggplot(fgseaRes[1:25, ], aes(y = reorder(pathway, NES), x = NES, fill = -log10(padj), label = size)) +
  geom_col() + geom_text(color = 'gray20', nudge_x = ifelse(fgseaRes$NES > 0, 0.5, 3)) +
  ylab('') + theme_classic()
```

```
Warning in x + params$x :
  longer object length is not a multiple of shorter object length
```

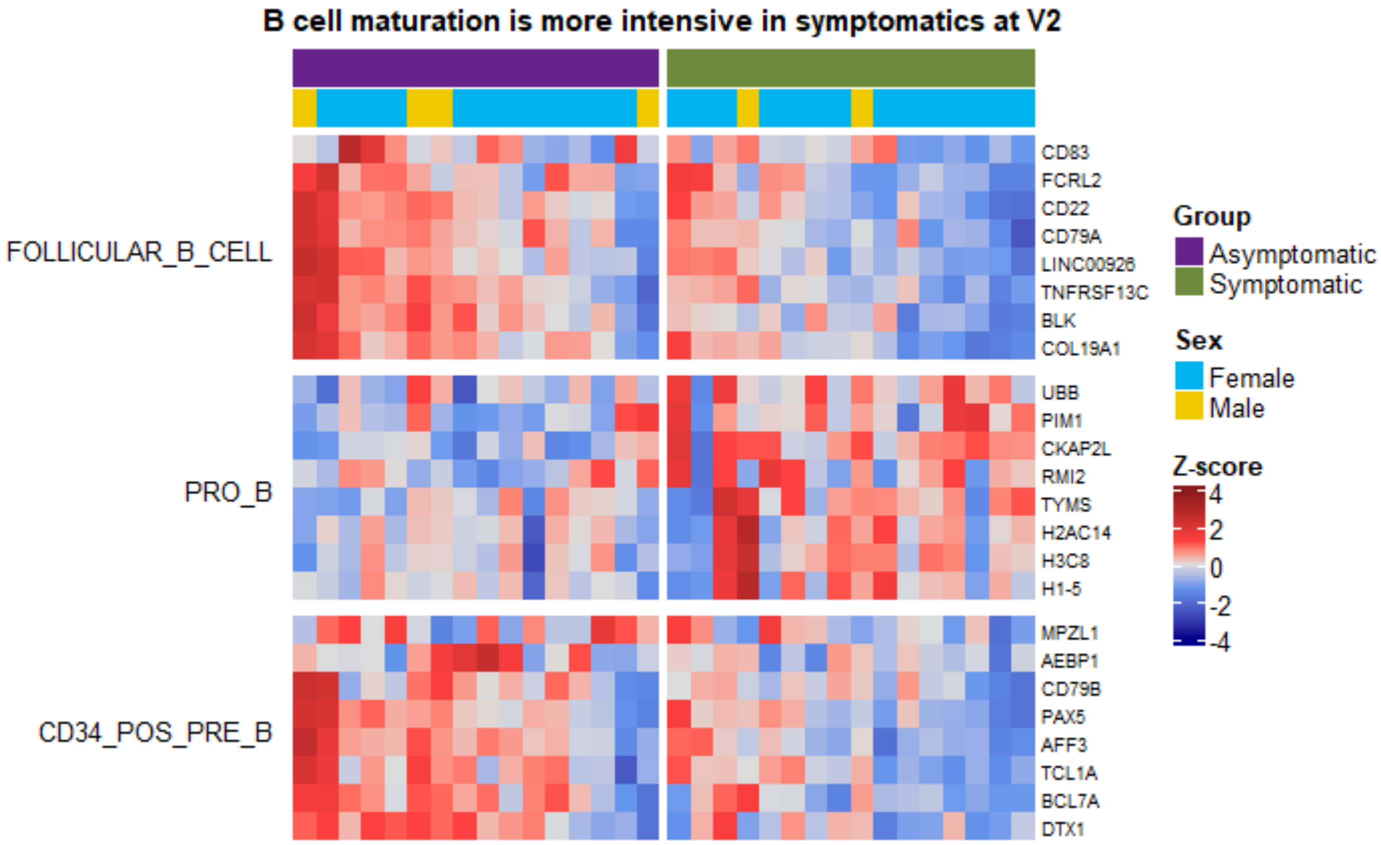


```
plotGseaTable(msigDB[fgseaRes[1:20, ][, pathway]], ranks, fgseaRes[1:20, ], gseaParam = 0.5, colwidths = c(3, 3,
0.8, 1.2, 1.2), pathwayLabelStyle = list(size = 10), headerLabelStyle = list(size = 10), valueStyle = list(size =
10))
```

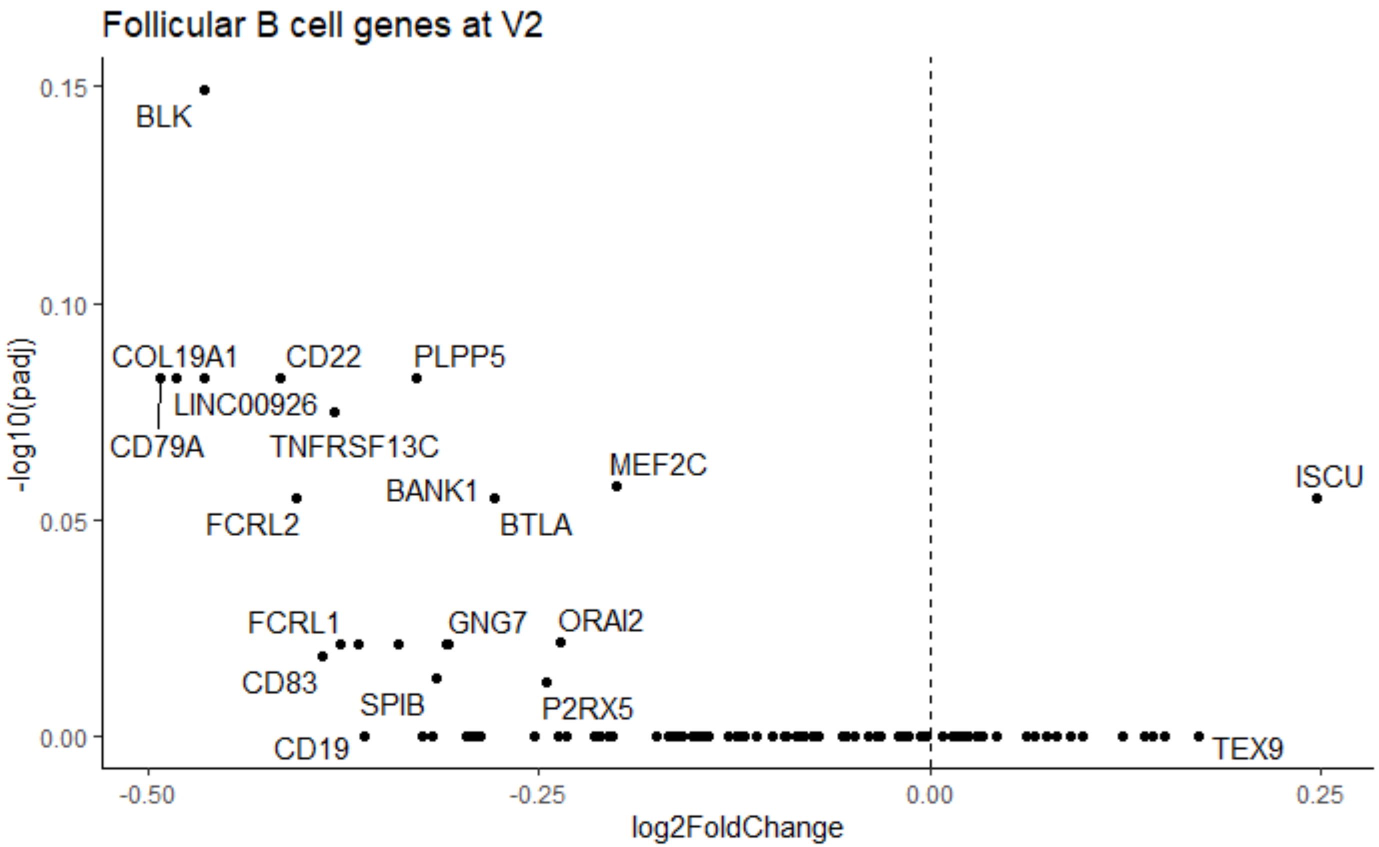


Pathway	Gene ranks	NES	pval	padj
ERYTHROBLAST		2.31	$10.0 \cdot 10^{-47}$	$2.9 \cdot 10^{-45}$
FOLLICULAR_B_CELL		-2.71	$1.9 \cdot 10^{-15}$	$2.7 \cdot 10^{-14}$
NK_CELLS		2.07	$1.1 \cdot 10^{-12}$	$1.0 \cdot 10^{-11}$
PRO_B		2.05	$2.6 \cdot 10^{-10}$	$1.9 \cdot 10^{-9}$
IMMATURE_NEUTROPHIL		-1.97	$1.3 \cdot 10^{-7}$	$7.3 \cdot 10^{-7}$
NEUTROPHIL		-1.66	$1.7 \cdot 10^{-7}$	$8.3 \cdot 10^{-7}$
NAIVE_T_CELL		-1.57	$1.5 \cdot 10^{-5}$	$6.4 \cdot 10^{-5}$
CD34_POS_PRE_B		-1.97	$1.0 \cdot 10^{-4}$	$3.8 \cdot 10^{-4}$
CD8_T_CELL		1.59	$8.7 \cdot 10^{-3}$	$2.8 \cdot 10^{-2}$
MONOCYTE		-1.35	$1.2 \cdot 10^{-2}$	$3.4 \cdot 10^{-2}$
STROMAL		-1.44	$1.3 \cdot 10^{-2}$	$3.5 \cdot 10^{-2}$
CD34_POS_MKP		-1.58	$2.2 \cdot 10^{-2}$	$5.4 \cdot 10^{-2}$
PLATELET		1.35	$3.2 \cdot 10^{-2}$	$7.2 \cdot 10^{-2}$
PRE_DENDRITIC		1.42	$4.9 \cdot 10^{-2}$	$1.0 \cdot 10^{-1}$
CD34_POS_MEP		1.26	$1.1 \cdot 10^{-1}$	$2.1 \cdot 10^{-1}$
CD34_POS_LMPP		-1.27	$1.6 \cdot 10^{-1}$	$2.9 \cdot 10^{-1}$
EARLY_ERYTHROBLAST		-1.17	$1.9 \cdot 10^{-1}$	$3.3 \cdot 10^{-1}$
DENDRITIC_CELL		-1.07	$2.9 \cdot 10^{-1}$	$4.7 \cdot 10^{-1}$
CD34_POS_ERP		-1.19	$3.2 \cdot 10^{-1}$	$4.8 \cdot 10^{-1}$
PLASMA_CELL		1.08	$3.5 \cdot 10^{-1}$	$5.1 \cdot 10^{-1}$

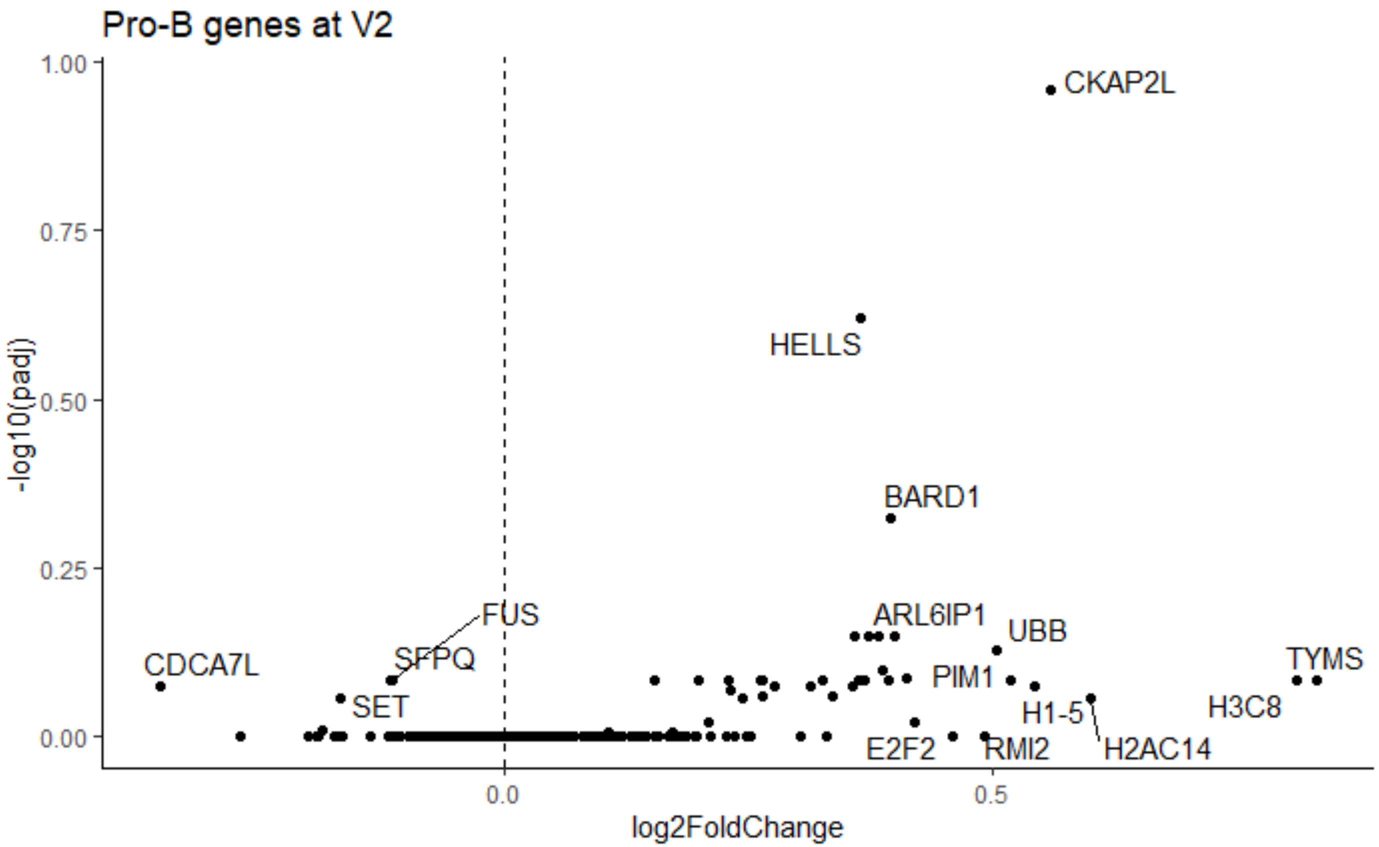
```
res.df = res.df[order(res.df$log2FoldChange, decreasing = FALSE), ]
ha_top = HeatmapAnnotation(df = meta_data[, c('Group', 'Sex')], col = list(Group = c('Symptomatic' = 'darkolivegreen4', 'Asymptomatic' = 'darkorchid4'), Sex = c('Male' = 'gold2', 'Female' = 'deepskyblue2')), show_annotation_name = FALSE)
plot1 = vsd.df[rownames(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[2, 'pathway']]]] & log2FoldChange < 0))[1:8], ]
zzz = t(scale(t(plot1))); zzz = pmin(pmax(zzz, -3), 3); rownames(zzz) =  res.df[rownames(zzz), 'external_gene_name']
plot1 = Heatmap(zzz, row_title = fgseaRes[[2, 'pathway']], show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = TRUE, show_row_dend = FALSE, show_column_dend = FALSE, column_split = meta_data$Group, top_annotation = ha_top, heatmap_legend_param = list(title = "Z-score"), col = colorRamp2(c(-4, -1.25, 0, 1.25, 4), c("darkblue", "cornflowerblue", "gainsboro", "brown1", "firebrick4")), row_names_gp = gpar(fontsize = 7), row_title_rot = 0, row_title_gp = gpar(fontsize = 10), column_title_gp = gpar(fontsize = 11, fontface = 2), column_title = 'B cell maturation is more intensive in symptomatics at V2')
res.df = res.df[order(res.df$log2FoldChange, decreasing = TRUE), ]
plot2 = vsd.df[rownames(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[4, 'pathway']]]] & log2FoldChange > 0))[1:8], ]
zzz = t(scale(t(plot2))); zzz = pmin(pmax(zzz, -3), 3); rownames(zzz) =  res.df[rownames(zzz), 'external_gene_name']
plot2 = Heatmap(zzz, show_heatmap_legend = FALSE, row_title = fgseaRes[[4, 'pathway']], show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = TRUE, show_row_dend = FALSE, show_column_dend = FALSE, column_split = meta_data$Group, heatmap_legend_param = list(title = "Z-score"), col = colorRamp2(c(-4, -1.25, 0, 1.25, 4), c("darkblue", "cornflowerblue", "gainsboro", "brown1", "firebrick4")), row_names_gp = gpar(fontsize = 7), row_title_rot = 0, row_title_gp = gpar(fontsize = 10), column_title_gp = gpar(fontsize = 11, fontface = 2))
res.df = res.df[order(res.df$log2FoldChange, decreasing = FALSE), ]
plot3 = vsd.df[rownames(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[8, 'pathway']]]] & log2FoldChange < 0))[1:8], ]
zzz = t(scale(t(plot3))); zzz = pmin(pmax(zzz, -3), 3); rownames(zzz) =  res.df[rownames(zzz), 'external_gene_name']
plot3 = Heatmap(zzz, show_heatmap_legend = FALSE, row_title = fgseaRes[[8, 'pathway']], show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = TRUE, show_row_dend = FALSE, show_column_dend = FALSE, column_split = meta_data$Group, heatmap_legend_param = list(title = "Z-score"), col = colorRamp2(c(-4, -1.25, 0, 1.25, 4), c("darkblue", "cornflowerblue", "gainsboro", "brown1", "firebrick4")), row_names_gp = gpar(fontsize = 7), row_title_rot = 0, row_title_gp = gpar(fontsize = 10), column_title_gp = gpar(fontsize = 11, fontface = 2))
draw(plot1 %v% plot2 %v% plot3); rm(plot1); rm(plot2); rm(plot3); rm(zzz); rm(ha_top)
```



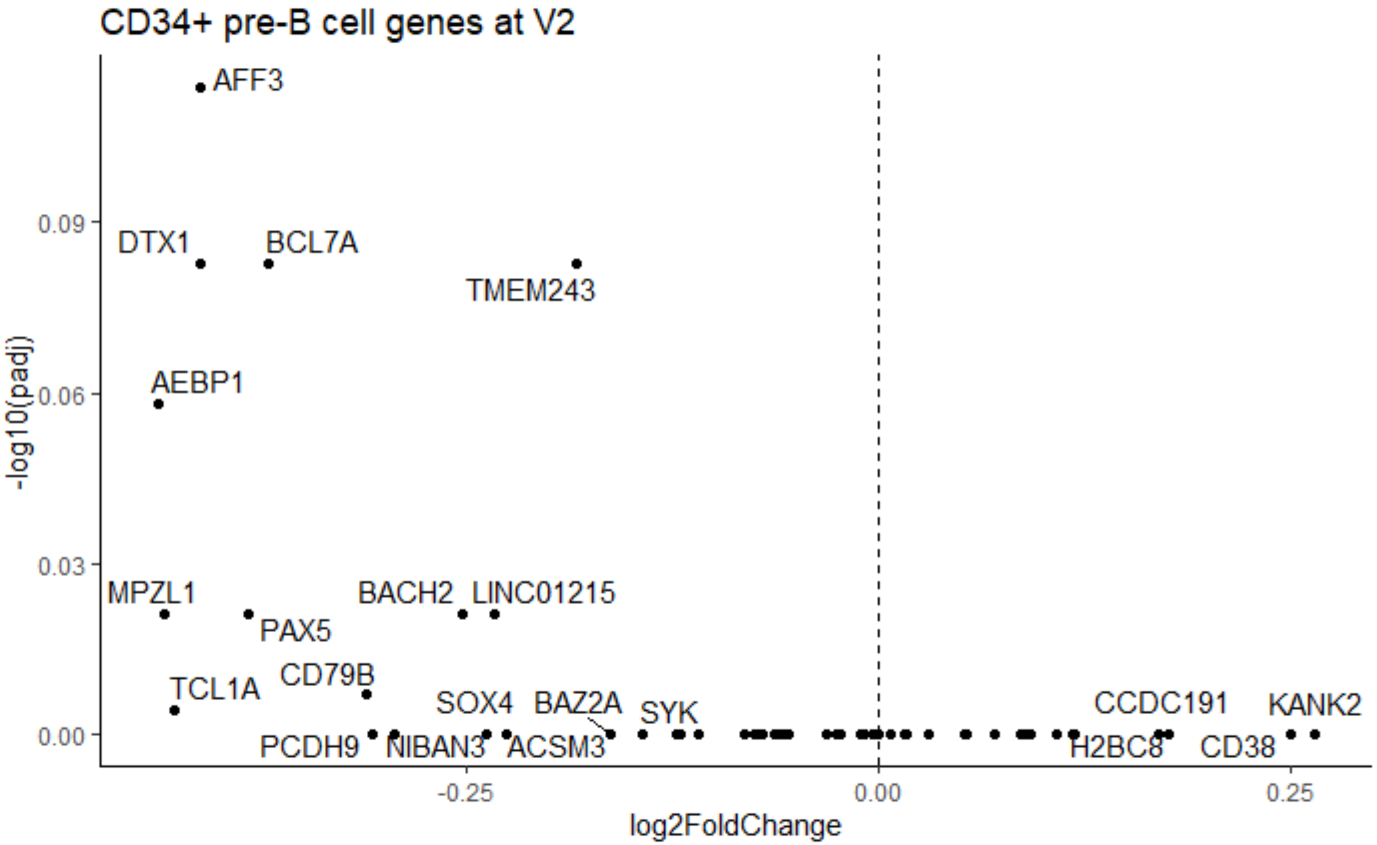
```
ggplot(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[2, 'pathway']]]]), aes(x = log2FoldChange, y = -l
og10(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + geom_vline(xinter
cept = 0, linetype = 2) + ggtitle('Follicular B cell genes at V2')
```



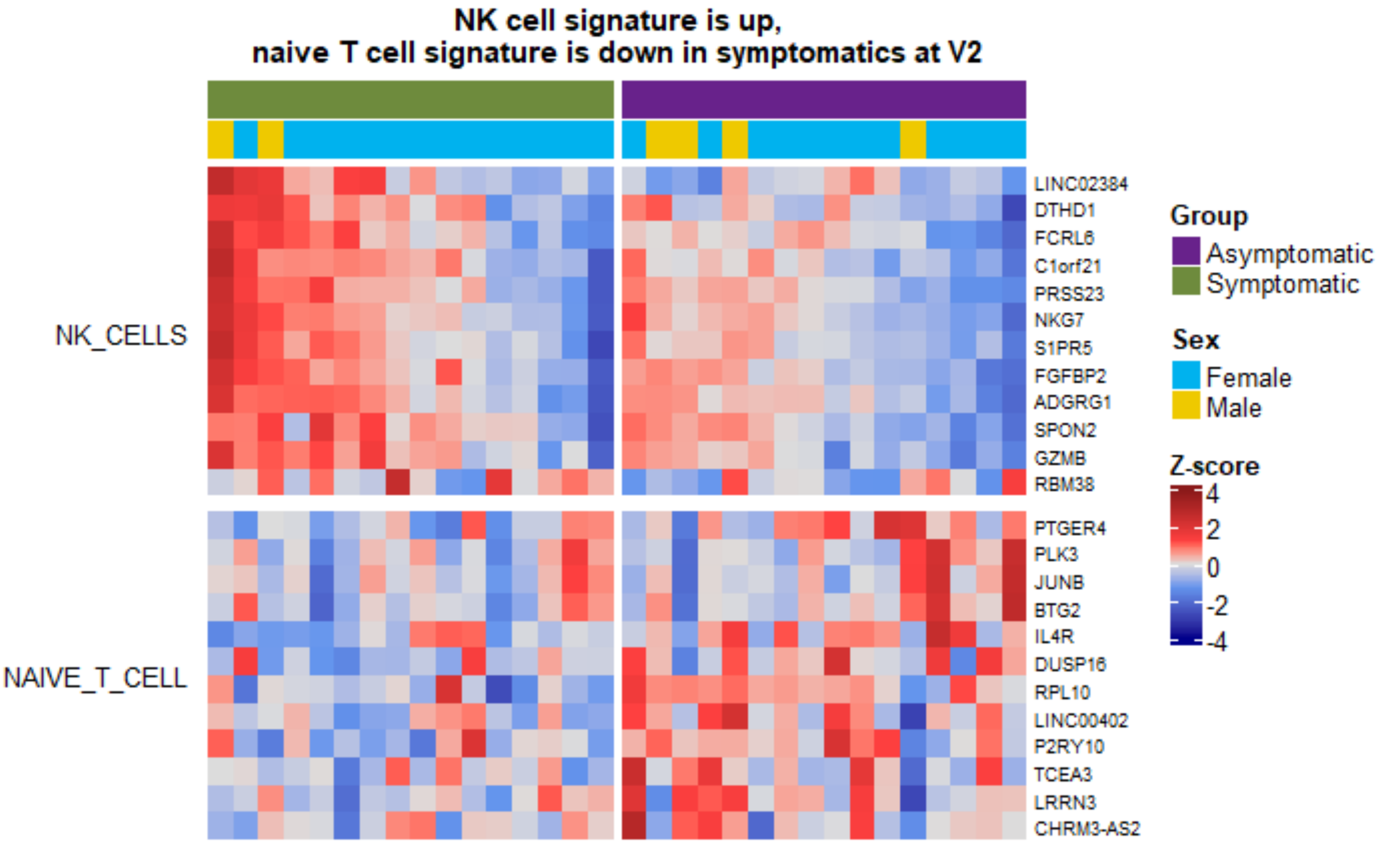
```
ggplot(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[4, 'pathway']]]]), aes(x = log2FoldChange, y = -l
og10(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + geom_vline(xinter
cept = 0, linetype = 2) + ggtitle('Pro-B genes at V2')
```



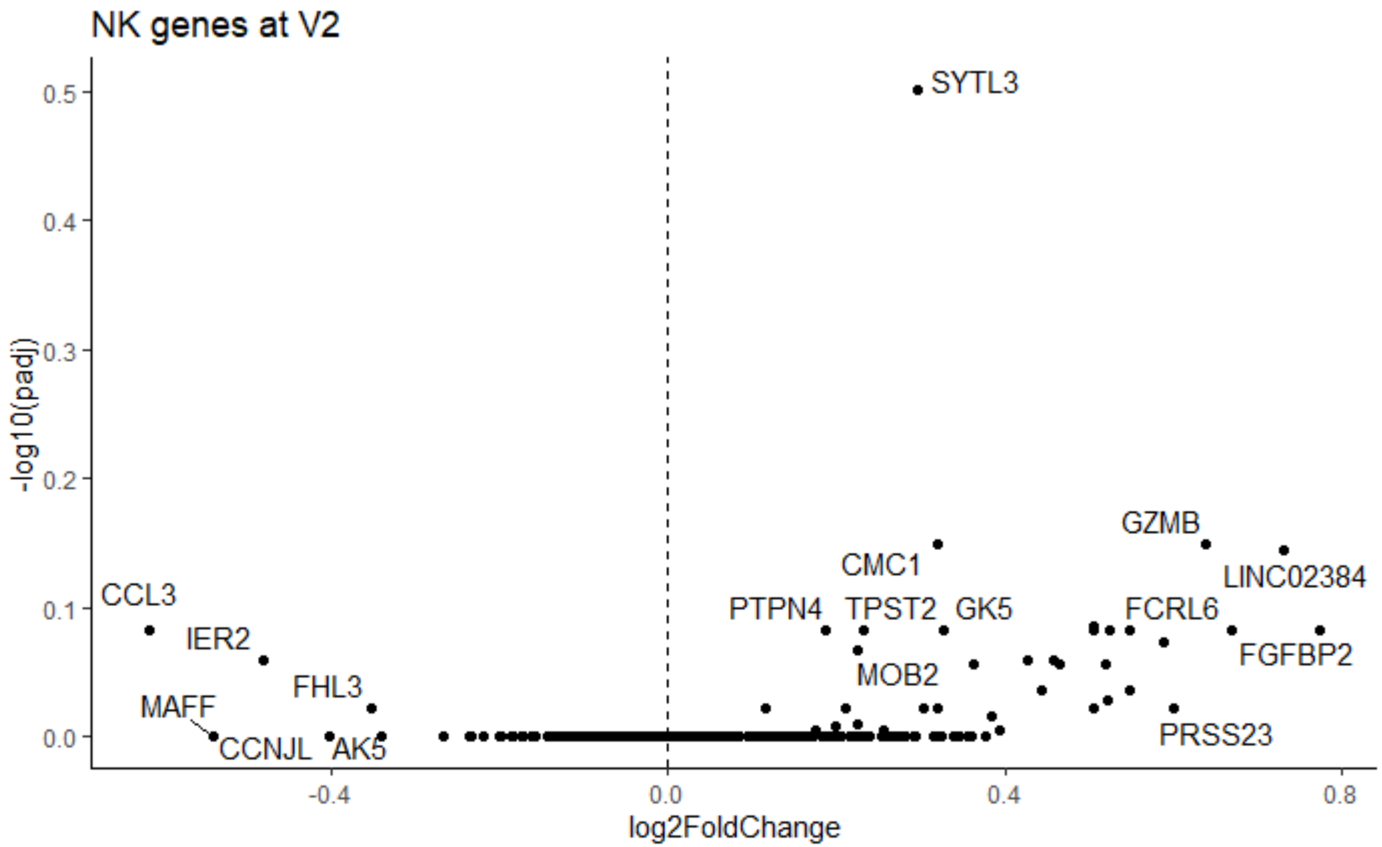
```
ggplot(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[8, 'pathway']]]]), aes(x = log2FoldChange, y = -l
og10(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + geom_vline(xinter
cept = 0, linetype = 2) + ggtitle('CD34+ pre-B cell genes at V2')
```


CD34+ pre-B cell genes at V2
AFF3
DTX1
BCL7A
TMEM243
AEBP1
MPZL1
PAX5
BACH2
LINC01215
TCL1A
CD79B
SOX4
BAZ2A
SYK
PCDH9
NIBAN3
ACSM3
CCDC191
KANK2
H2BC8
CD38
0.09
0.06
0.03
0.00
-log10(padj)
-0.25
0.00
0.25
log2FoldChange

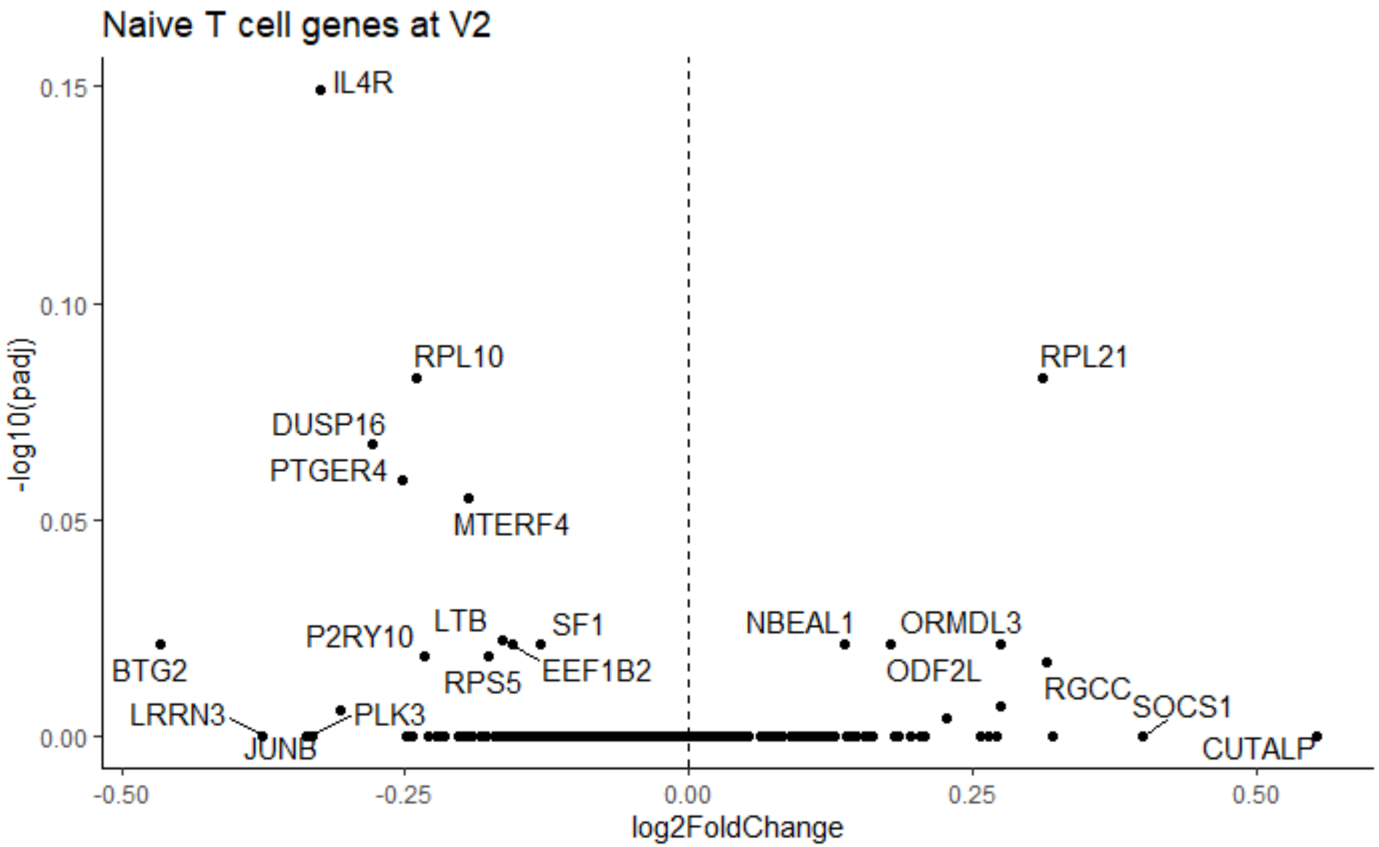
```
res.df = res.df[order(res.df$log2FoldChange, decreasing = TRUE), ]
ha_top = HeatmapAnnotation(df = meta_data[, c('Group', 'Sex')], col = list(Group = c('Symptomatic' = 'darkolivegreen4', 'Asymptomatic' = 'darkorchid4'), Sex = c('Male' = 'gold2', 'Female' = 'deepskyblue2')), show_annotation_name = FALSE)
plot1 = vsd.df[rownames(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[3, 'pathway']]]] & log2FoldChange > 0))[1:12], ]
zzz = t(scale(t(plot1))); zzz = pmin(pmax(zzz, -3), 3); rownames(zzz) =  res.df[rownames(zzz), 'external_gene_name']
plot1 = Heatmap(zzz, row_title = fgseaRes[[3, 'pathway']], show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = TRUE, show_row_dend = FALSE, show_column_dend = FALSE, column_split = meta_data$Group, top_annotation = ha_top, heatmap_legend_param = list(title = "Z-score"), col = colorRamp2(c(-4, -1.25, 0, 1.25, 4), c("darkblue", "cornflowerblue", "gainsboro", "brown1", "firebrick4")), row_names_gp = gpar(fontsize = 7), row_title_rot = 0, row_title_gp = gpar(fontsize = 10), column_title_gp = gpar(fontsize = 11, fontface = 2), column_title = 'NK cell signature is up,\nnaive T cell signature is down in symptomatics at V2')
res.df = res.df[order(res.df$log2FoldChange, decreasing = FALSE), ]
plot2 = vsd.df[rownames(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[7, 'pathway']]]] & log2FoldChange < 0))[1:12], ]
zzz = t(scale(t(plot2))); zzz = pmin(pmax(zzz, -3), 3); rownames(zzz) =  res.df[rownames(zzz), 'external_gene_name']
plot2 = Heatmap(zzz, show_heatmap_legend = FALSE, row_title = fgseaRes[[7, 'pathway']], show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = TRUE, show_row_dend = FALSE, show_column_dend = FALSE, column_split = meta_data$Group, heatmap_legend_param = list(title = "Z-score"), col = colorRamp2(c(-4, -1.25, 0, 1.25, 4), c("darkblue", "cornflowerblue", "gainsboro", "brown1", "firebrick4")), row_names_gp = gpar(fontsize = 7), row_title_rot = 0, row_title_gp = gpar(fontsize = 10), column_title_gp = gpar(fontsize = 11, fontface = 2))
res.df = res.df[order(res.df$log2FoldChange, decreasing = FALSE), ]
draw(plot1 %v% plot2); rm(plot1); rm(plot2); rm(zzz); rm(ha_top)
```



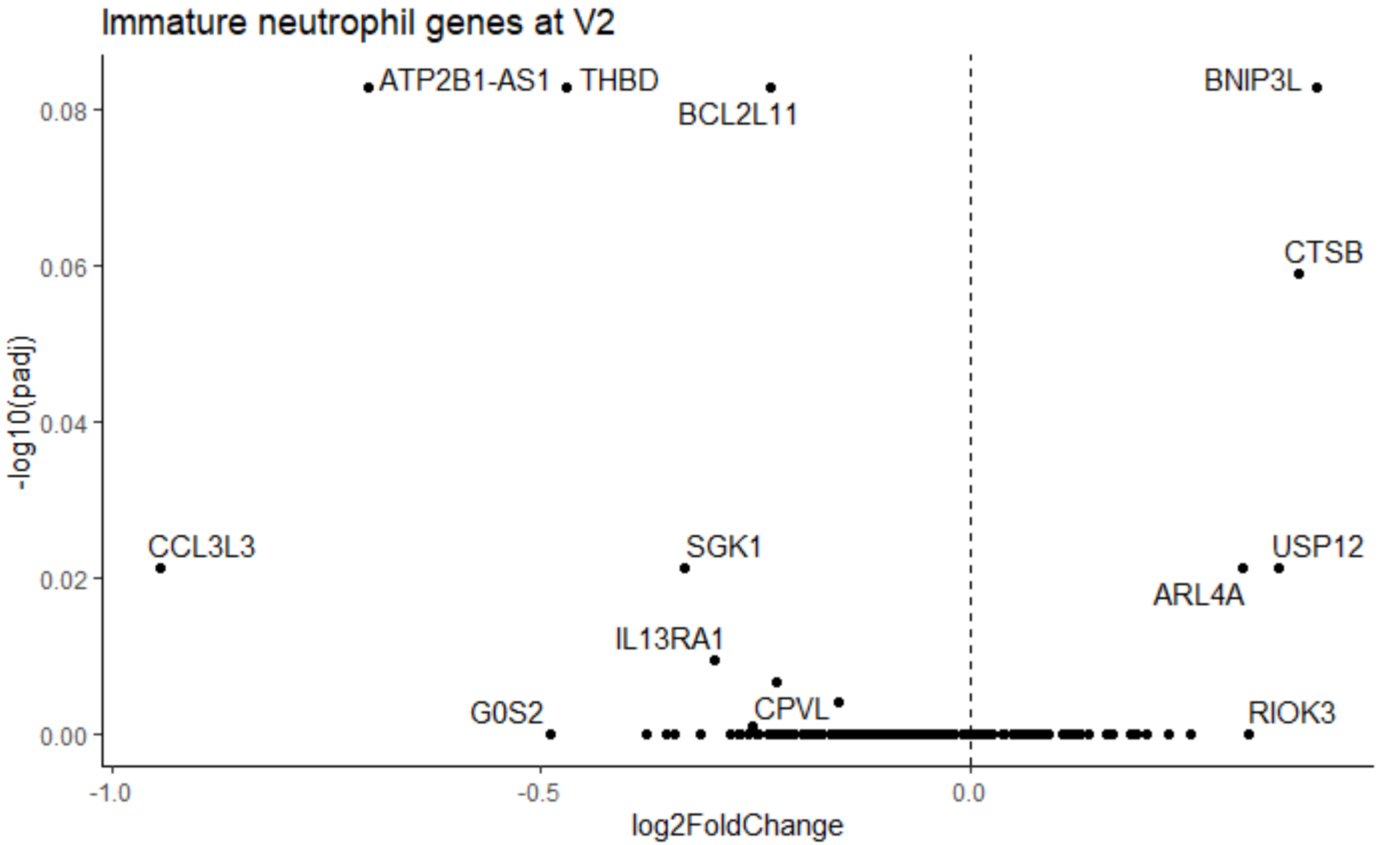
```
ggplot(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[3, 'pathway']]]]), aes(x = log2FoldChange, y = -l
og10(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + ggtitle('NK genes
at V2') + geom_vline(xintercept = 0, linetype = 2)
```



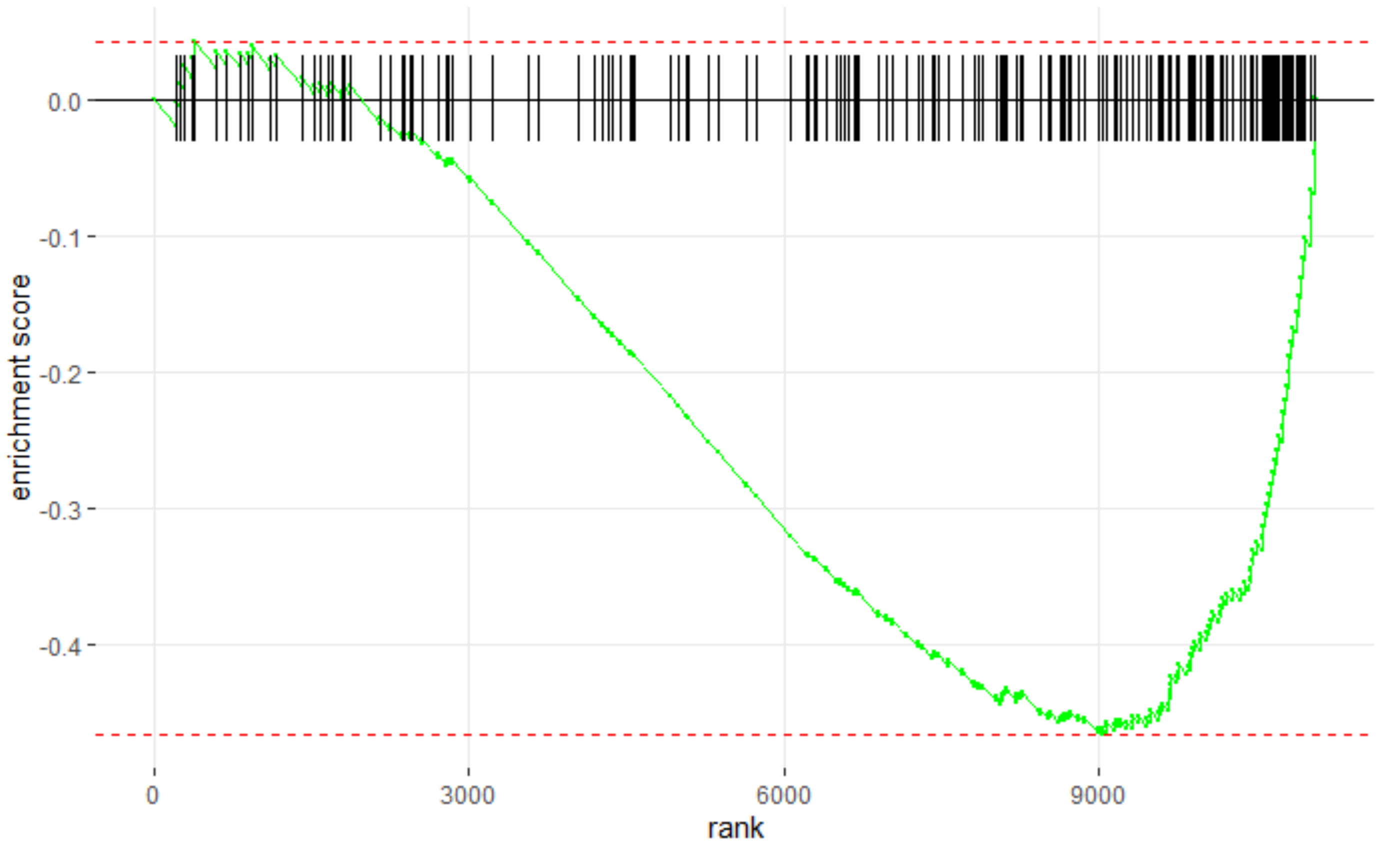
```
ggplot(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[7, 'pathway']]]]), aes(x = log2FoldChange, y = -l
og10(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + ggtitle('Naive T
cell genes at V2') + geom_vline(xintercept = 0, linetype = 2)
```



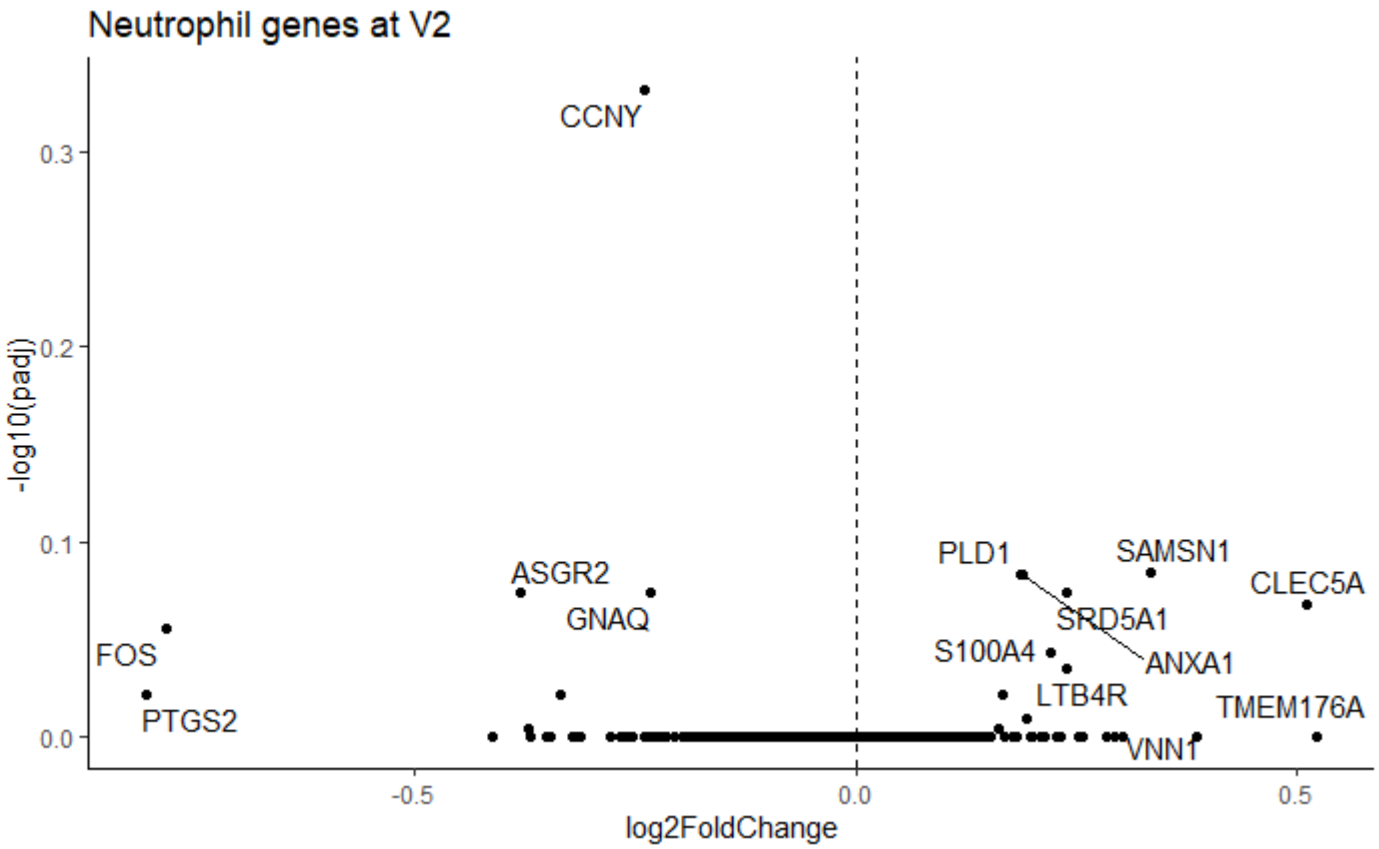
```
ggplot(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[5, 'pathway']]]]), aes(x = log2FoldChange, y = -log10(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + ggtitle('Immature neutrophil genes at V2') + geom_vline(xintercept = 0, linetype = 2)
```



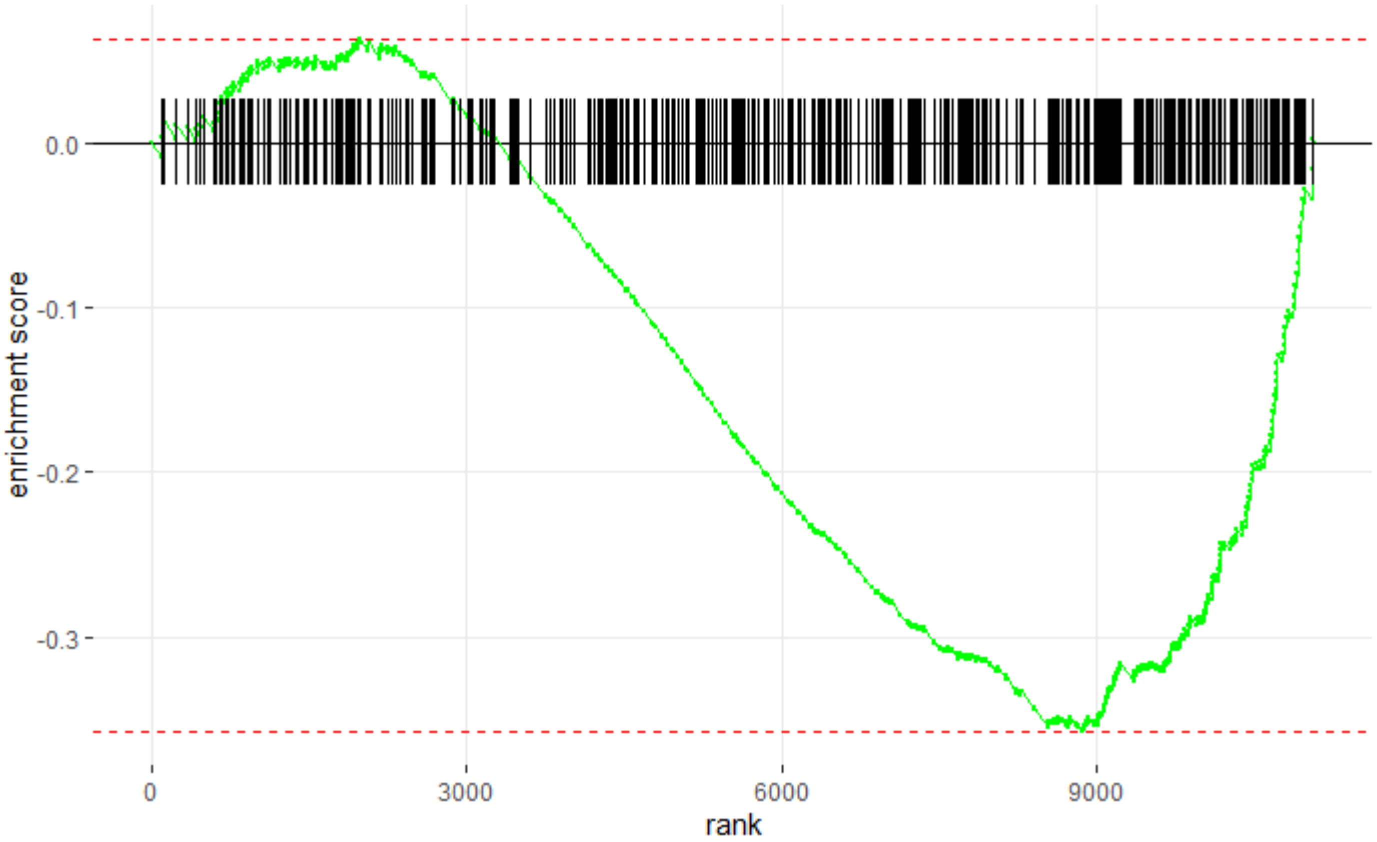
```
plotEnrichment(msigDB[[fgseaRes[[5, 'pathway']]]], ranks, gseaParam = 1, ticksSize = 0.2)
```



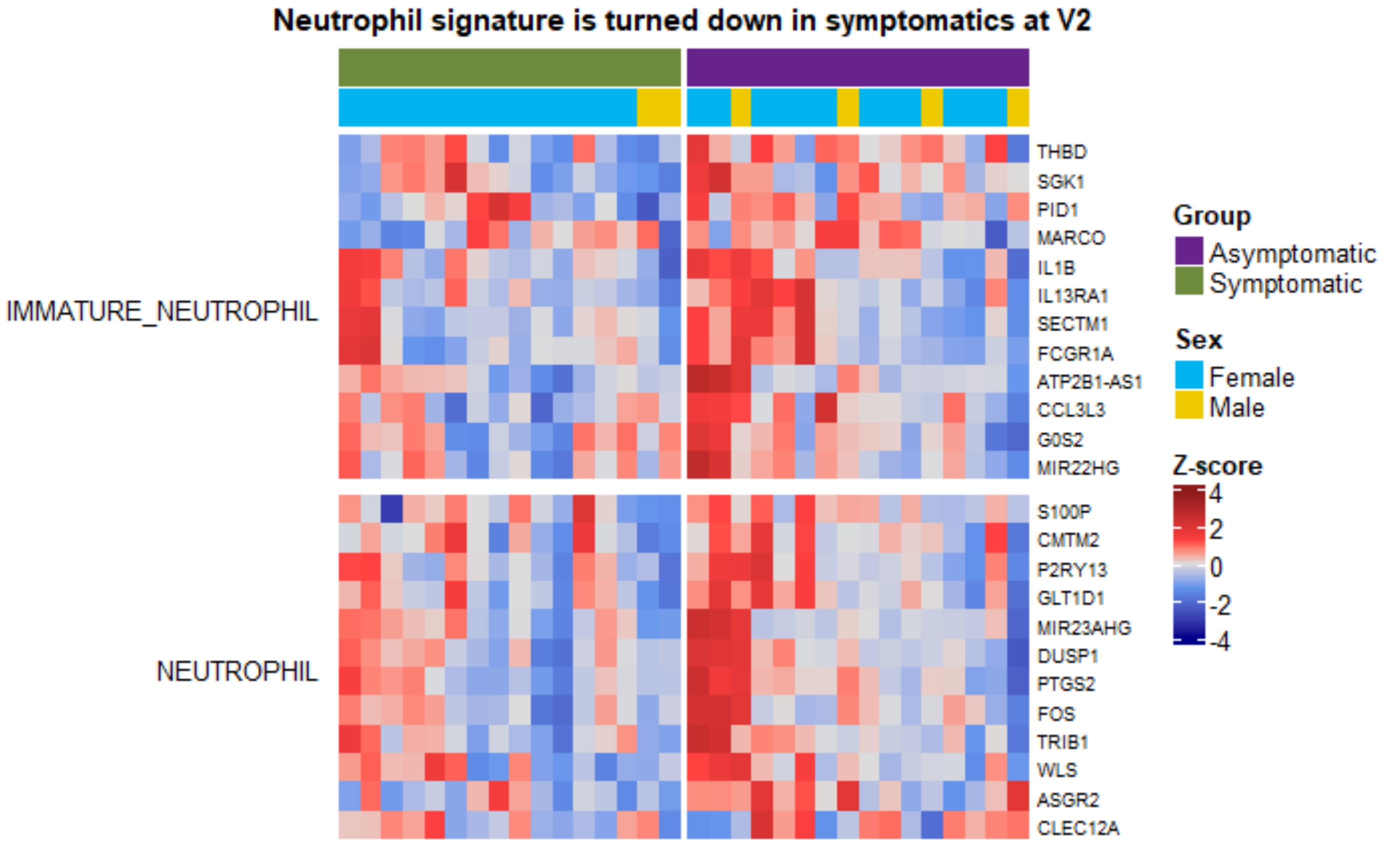
```
ggplot(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[6, 'pathway']]]]), aes(x = log2FoldChange, y = -log10(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + ggtitle('Neutrophil genes at V2') + geom_vline(xintercept = 0, linetype = 2)
```



```
plotEnrichment(msigDB[[fgseaRes[[6, 'pathway']]]], ranks, gseaParam = 1, ticksSize = 0.2)
```


0.0
-0.1
-0.2
-0.3
enrichment score
0
3000
6000
9000
rank

```r
res.df = res.df[order(res.df$log2FoldChange, decreasing = FALSE), ]
ha_top = HeatmapAnnotation(df = meta_data[, c('Group', 'Sex')], col = list(Group = c('Symptomatic' = 'darkolivegreen4', 'Asymptomatic' = 'darkorchid4'), Sex = c('Male' = 'gold2', 'Female' = 'deepskyblue2')), show_annotation_name = FALSE)
plot1 = vsd.df[rownames(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[5, 'pathway']]]] & log2FoldChange < 0))[1:12], ]
zzz = t(scale(t(plot1))); zzz = pmin(pmax(zzz, -3), 3); rownames(zzz) =  res.df[rownames(zzz), 'external_gene_name']
plot1 = Heatmap(zzz, row_title = fgseaRes[[5, 'pathway']], show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = TRUE, show_row_dend = FALSE, show_column_dend = FALSE, column_split = meta_data$Group, top_annotation = ha_top, heatmap_legend_param = list(title = "Z-score"), col = colorRamp2(c(-4, -1.25, 0, 1.25, 4), c("darkblue", "cornflowerblue", "gainsboro", "brown1", "firebrick4")), row_names_gp = gpar(fontsize = 7), row_title_rot = 0, row_title_gp = gpar(fontsize = 10), column_title_gp = gpar(fontsize = 11, fontface = 2), column_title = 'Neutrophil signature is turned down in symptomatics at V2')
plot2 = vsd.df[rownames(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[6, 'pathway']]]] & log2FoldChange < 0))[1:12], ]
zzz = t(scale(t(plot2))); zzz = pmin(pmax(zzz, -3), 3); rownames(zzz) =  res.df[rownames(zzz), 'external_gene_name']
plot2 = Heatmap(zzz, show_heatmap_legend = FALSE, row_title = fgseaRes[[6, 'pathway']], show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = TRUE, show_row_dend = FALSE, show_column_dend = FALSE, column_split = meta_data$Group, heatmap_legend_param = list(title = "Z-score"), col = colorRamp2(c(-4, -1.25, 0, 1.25, 4), c("darkblue", "cornflowerblue", "gainsboro", "brown1", "firebrick4")), row_names_gp = gpar(fontsize = 7), row_title_rot = 0, row_title_gp = gpar(fontsize = 10), column_title_gp = gpar(fontsize = 11, fontface = 2))

draw(plot1 %v% plot2); rm(plot1); rm(plot2); rm(zzz); rm(ha_top)
```



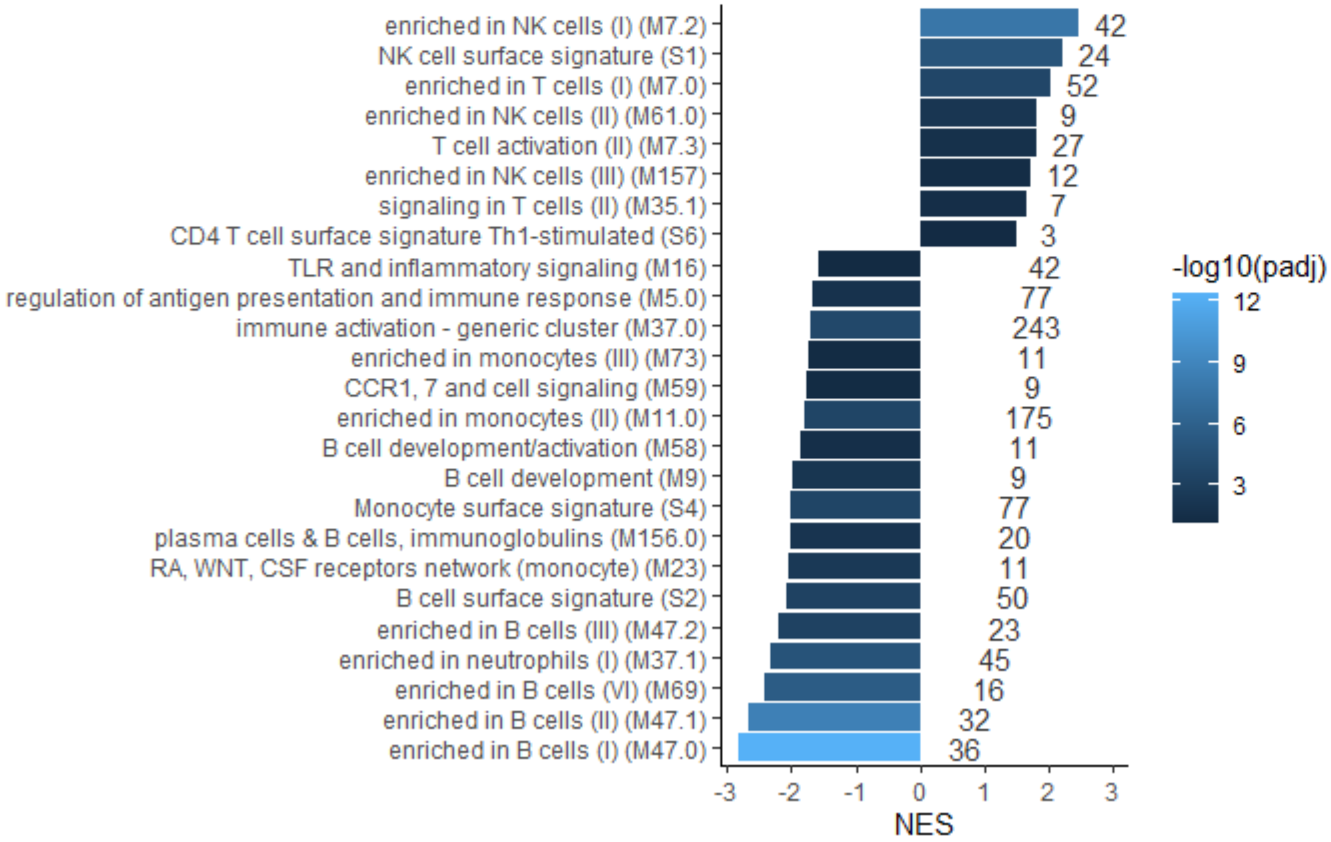
```
btm = readRDS('data/Blood_Transcript_modules.RDS'); btm_annot = read_excel('data/btm_annotation_table.xls')
```

```
New names:
• `Module size` -> `Module size...4`
• `Jaccard Index` -> `Jaccard Index...10`
• `Enrichment p value` -> `Enrichment p value...11`
• `Overlap size` -> `Overlap size...13`
• `Jaccard Index` -> `Jaccard Index...15`
• `Enrichment p value` -> `Enrichment p value...16`
• `Module size` -> `Module size...17`
• `Overlap size` -> `Overlap size...18`
• `Jaccard Index` -> `Jaccard Index...20`
• `Enrichment p value` -> `Enrichment p value...21`
• `Overlap size` -> `Overlap size...23`
```

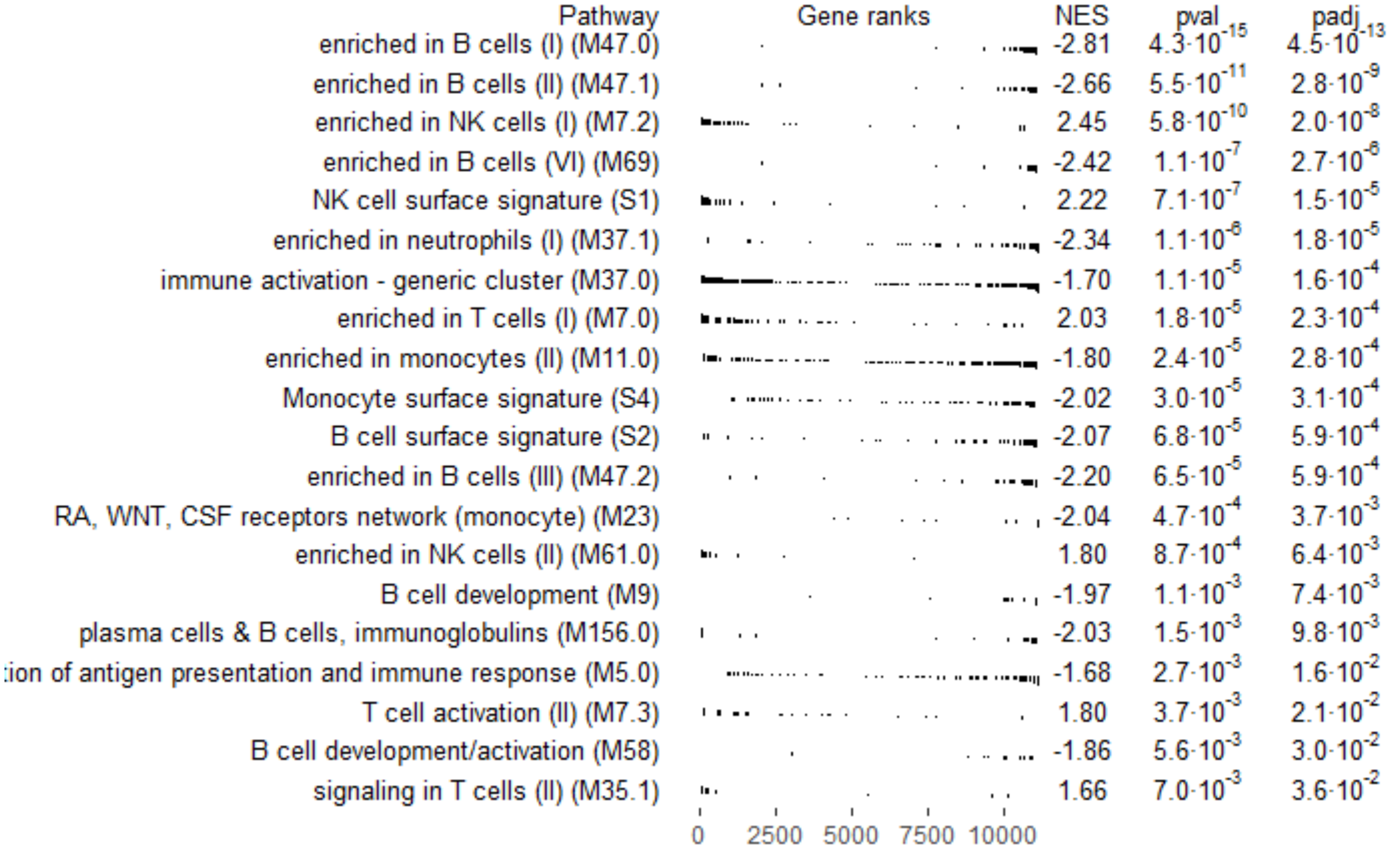
```
btm_annot = subset(btm_annot, `Module category` == 'immune'); btm = btm[c(btm_annot$ID)]; names(btm) = btm_annot$`Composite name`
fgseaRes = fgsea(pathways = btm, stats = ranks, nproc = 1); fgseaRes = fgseaRes[order(fgseaRes$padj, decreasing = FALSE), ]
```

```
write.csv(subset(fgseaRes, select = -c(leadingEdge)), file = 'results/fgseaRes_btm_immune_v2.csv', row.names = FALSE)

fgseaRes = fgseaRes[pathway %in% subset(btm_annot, `Module category` == 'immune')$`Composite name`, ]
ggplot(fgseaRes[1:25, ], aes(y = reorder(pathway, NES), x = NES, fill = -log10(padj), label = size)) + geom_col()
+ geom_text(color = 'gray20', nudge_x = ifelse(fgseaRes$NES > 0, 0.5, 3.5)) + ylab('') + theme_classic()
```

```
Warning in x + params$x :
  longer object length is not a multiple of shorter object length
```

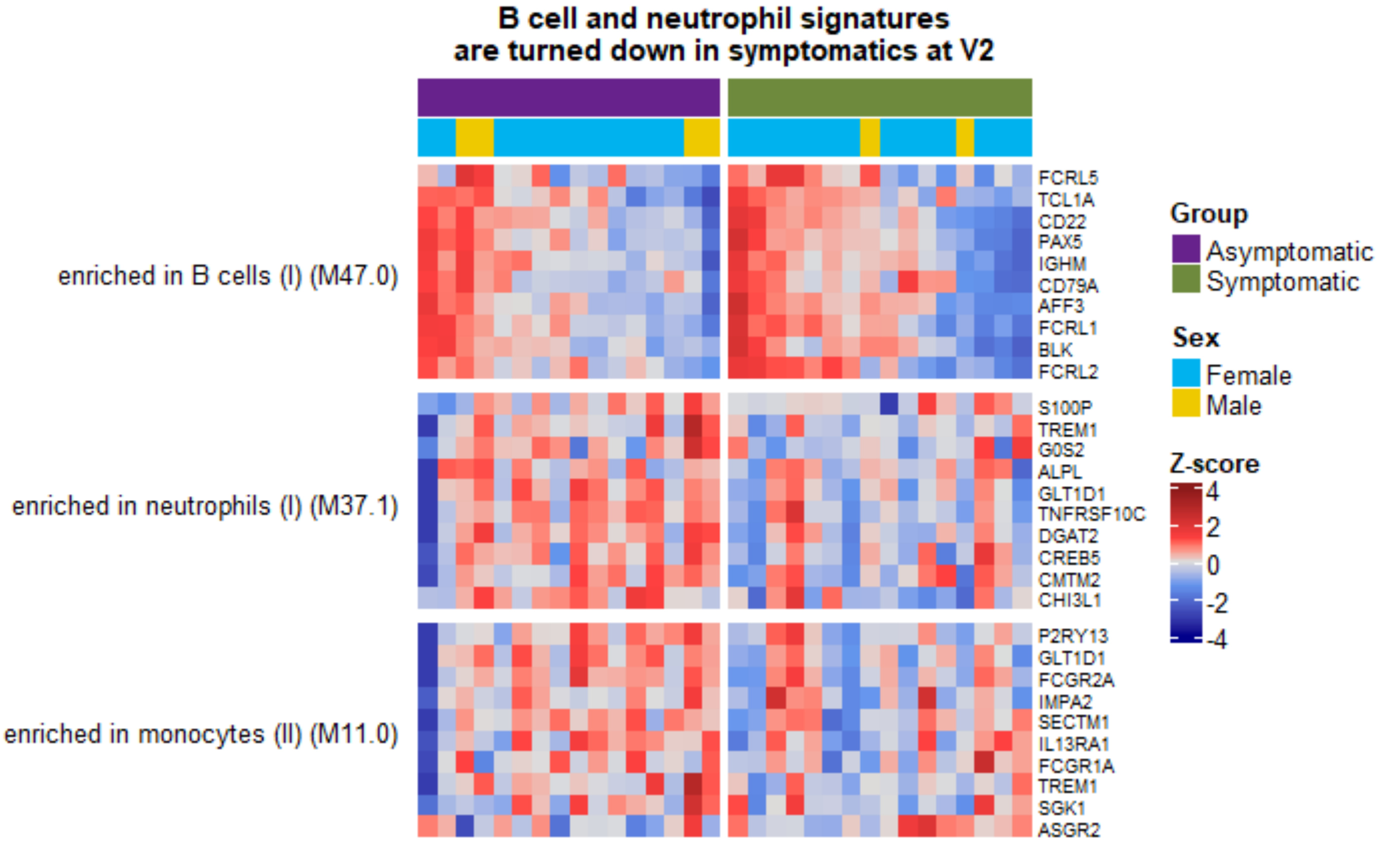


```
plotGseaTable(btm[fgseaRes[1:20, ][, pathway]], ranks, fgseaRes[1:20, ], gseaParam = 0.5, colwidths = c(6, 3, 0.8, 1.2, 1.2), pathwayLabelStyle = list(size = 10), headerLabelStyle = list(size = 10), valueStyle = list(size = 10))
```

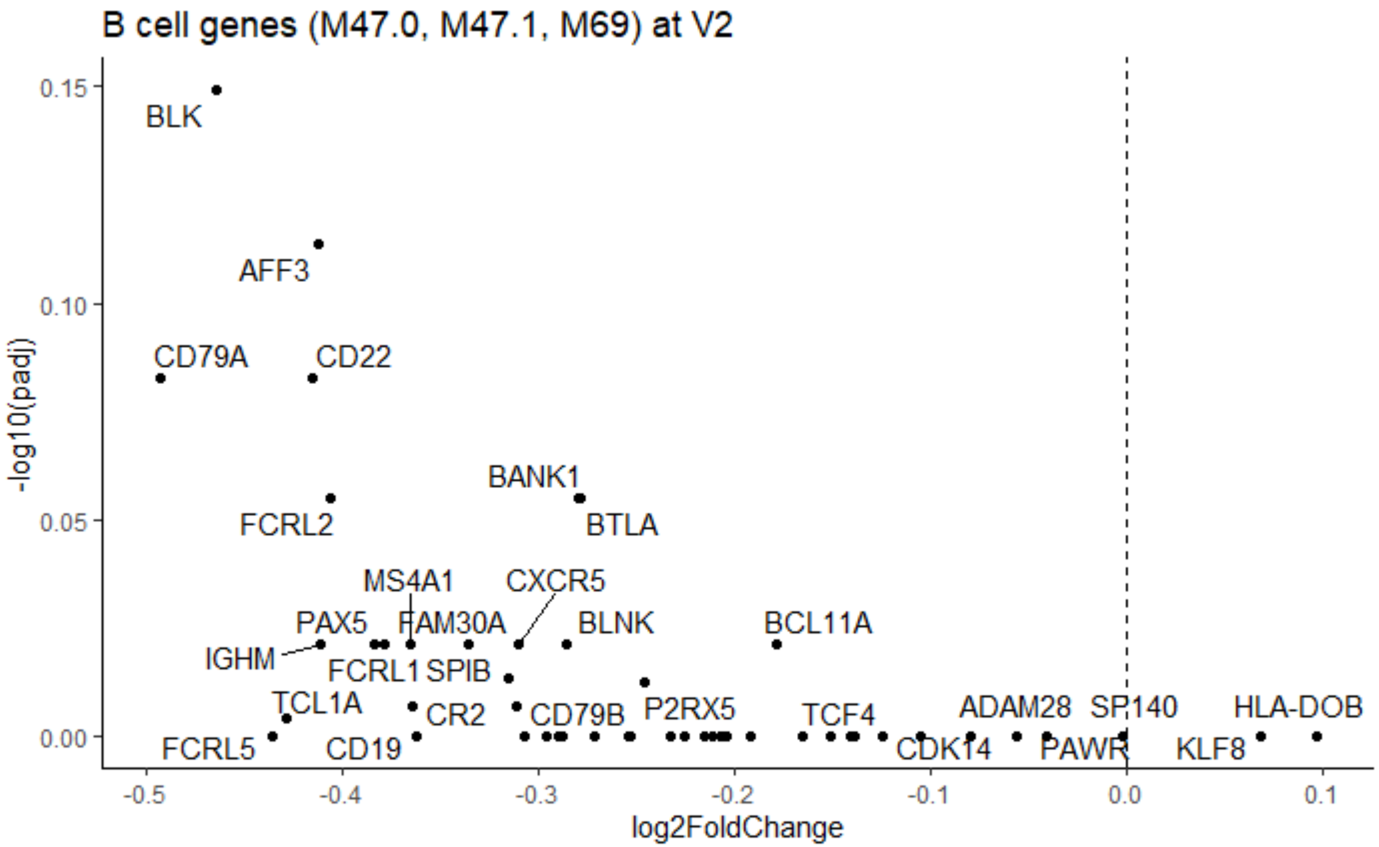


Pathway	NES	pval	padj
enriched in B cells (I) (M47.0)	-2.81	$4.3 \cdot 10^{-15}$	$4.5 \cdot 10^{-13}$
enriched in B cells (II) (M47.1)	-2.66	$5.5 \cdot 10^{-11}$	$2.8 \cdot 10^{-9}$
enriched in NK cells (I) (M7.2)	2.45	$5.8 \cdot 10^{-10}$	$2.0 \cdot 10^{-8}$
enriched in B cells (VI) (M69)	-2.42	$1.1 \cdot 10^{-7}$	$2.7 \cdot 10^{-6}$
NK cell surface signature (S1)	2.22	$7.1 \cdot 10^{-7}$	$1.5 \cdot 10^{-5}$
enriched in neutrophils (I) (M37.1)	-2.34	$1.1 \cdot 10^{-6}$	$1.8 \cdot 10^{-5}$
immune activation - generic cluster (M37.0)	-1.70	$1.1 \cdot 10^{-5}$	$1.6 \cdot 10^{-4}$
enriched in T cells (I) (M7.0)	2.03	$1.8 \cdot 10^{-5}$	$2.3 \cdot 10^{-4}$
enriched in monocytes (II) (M11.0)	-1.80	$2.4 \cdot 10^{-5}$	$2.8 \cdot 10^{-4}$
Monocyte surface signature (S4)	-2.02	$3.0 \cdot 10^{-5}$	$3.1 \cdot 10^{-4}$
B cell surface signature (S2)	-2.07	$6.8 \cdot 10^{-5}$	$5.9 \cdot 10^{-4}$
enriched in B cells (III) (M47.2)	-2.20	$6.5 \cdot 10^{-5}$	$5.9 \cdot 10^{-4}$
RA, WNT, CSF receptors network (monocyte) (M23)	-2.04	$4.7 \cdot 10^{-4}$	$3.7 \cdot 10^{-3}$
enriched in NK cells (II) (M61.0)	1.80	$8.7 \cdot 10^{-4}$	$6.4 \cdot 10^{-3}$
B cell development (M9)	-1.97	$1.1 \cdot 10^{-3}$	$7.4 \cdot 10^{-3}$
plasma cells & B cells, immunoglobulins (M156.0)	-2.03	$1.5 \cdot 10^{-3}$	$9.8 \cdot 10^{-3}$
tion of antigen presentation and immune response (M5.0)	-1.68	$2.7 \cdot 10^{-3}$	$1.6 \cdot 10^{-2}$
T cell activation (II) (M7.3)	1.80	$3.7 \cdot 10^{-3}$	$2.1 \cdot 10^{-2}$
B cell development/activation (M58)	-1.86	$5.6 \cdot 10^{-3}$	$3.0 \cdot 10^{-2}$
signaling in T cells (II) (M35.1)	1.66	$7.0 \cdot 10^{-3}$	$3.6 \cdot 10^{-2}$

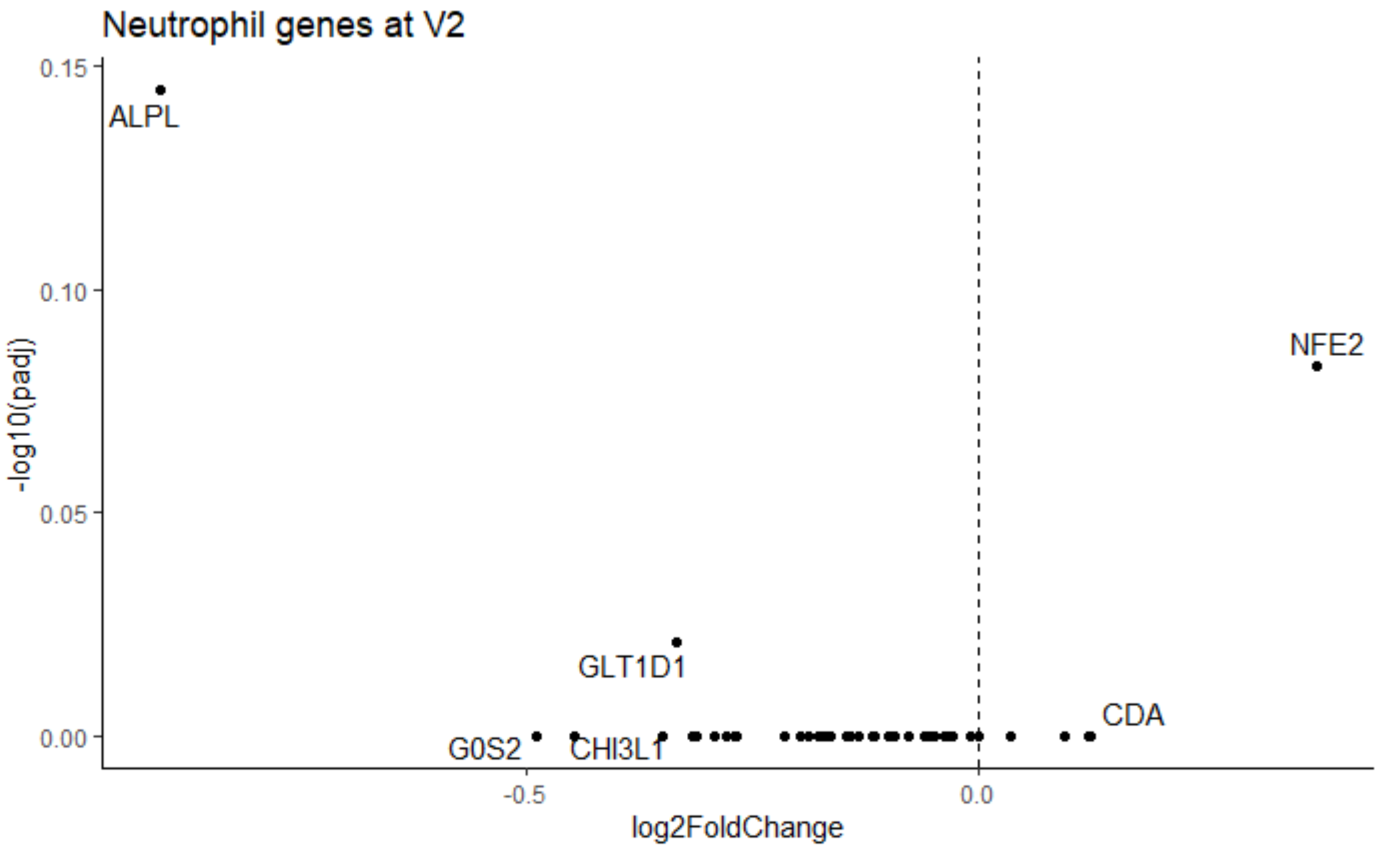
```
res.df = res.df[order(res.df$log2FoldChange, decreasing = FALSE), ]
ha_top = HeatmapAnnotation(df = meta_data[, c('Group', 'Sex')], col = list(Group = c('Symptomatic' = 'darkolivegreen4', 'Asymptomatic' = 'darkorchid4'), Sex = c('Male' = 'gold2', 'Female' = 'deepskyblue2')), show_annotation_name = FALSE)
plot1 = vsd.df[rownames(subset(res.df, external_gene_name %in% btm[[fgseaRes[[1, 'pathway']]]] & log2FoldChange < 0))[1:10], ]
zzz = t(scale(t(plot1))); zzz = pmin(pmax(zzz, -3), 3); rownames(zzz) =  res.df[rownames(zzz), 'external_gene_name']
plot1 = Heatmap(zzz, row_title = fgseaRes[[1, 'pathway']], show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = TRUE, show_row_dend = FALSE, show_column_dend = FALSE, column_split = meta_data$Group, top_annotation = ha_top, heatmap_legend_param = list(title = "Z-score"), col = colorRamp2(c(-4, -1.25, 0, 1.25, 4), c("darkblue", "cornflowerblue", "gainsboro", "brown1", "firebrick4")), row_names_gp = gpar(fontsize = 7), row_title_rot = 0, row_title_gp = gpar(fontsize = 10), column_title_gp = gpar(fontsize = 11, fontface = 2), column_title = 'B cell and neutrophil signatures\nare turned down in symptomatics at V2')
plot2 = vsd.df[rownames(subset(res.df, external_gene_name %in% btm[[fgseaRes[[6, 'pathway']]]] & log2FoldChange < 0))[1:10], ]
zzz = t(scale(t(plot2))); zzz = pmin(pmax(zzz, -3), 3); rownames(zzz) =  res.df[rownames(zzz), 'external_gene_name']
plot2 = Heatmap(zzz, show_heatmap_legend = FALSE, row_title = fgseaRes[[6, 'pathway']], show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = TRUE, show_row_dend = FALSE, show_column_dend = FALSE, column_split = meta_data$Group, heatmap_legend_param = list(title = "Z-score"), col = colorRamp2(c(-4, -1.25, 0, 1.25, 4), c("darkblue", "cornflowerblue", "gainsboro", "brown1", "firebrick4")), row_names_gp = gpar(fontsize = 7), row_title_rot = 0, row_title_gp = gpar(fontsize = 10), column_title_gp = gpar(fontsize = 11, fontface = 2))
plot3 = vsd.df[rownames(subset(res.df, external_gene_name %in% btm[[fgseaRes[[9, 'pathway']]]] & log2FoldChange < 0))[1:10], ]
zzz = t(scale(t(plot3))); zzz = pmin(pmax(zzz, -3), 3); rownames(zzz) =  res.df[rownames(zzz), 'external_gene_name']
plot3 = Heatmap(zzz, show_heatmap_legend = FALSE, row_title = fgseaRes[[9, 'pathway']], show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = TRUE, show_row_dend = FALSE, show_column_dend = FALSE, column_split = meta_data$Group, heatmap_legend_param = list(title = "Z-score"), col = colorRamp2(c(-4, -1.25, 0, 1.25, 4), c("darkblue", "cornflowerblue", "gainsboro", "brown1", "firebrick4")), row_names_gp = gpar(fontsize = 7), row_title_rot = 0, row_title_gp = gpar(fontsize = 10), column_title_gp = gpar(fontsize = 11, fontface = 2))
draw(plot1 %v% plot2 %v% plot3)
```



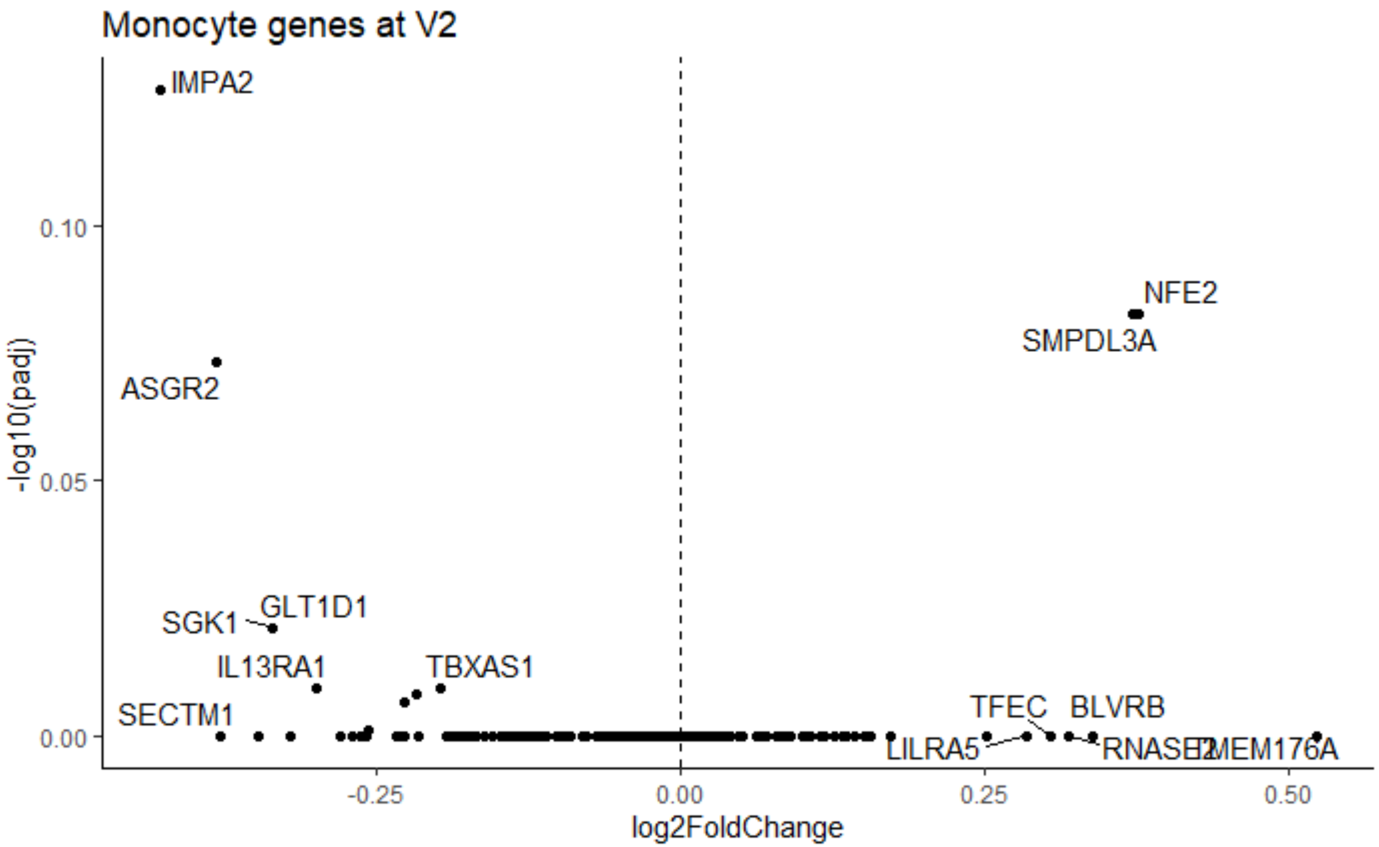
```
ggplot(subset(res.df, external_gene_name %in% btm[[fgseaRes[[1, 'pathway']]]] | external_gene_name %in% btm[[fgse
aRes[[2, 'pathway']]]] | external_gene_name %in% btm[[fgseaRes[[4, 'pathway']]]]), aes(x = log2FoldChange, y = -l
og10(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + ggtitle('B cell g
enes (M47.0, M47.1, M69) at V2') + geom_vline(xintercept = 0, linetype = 2)
```



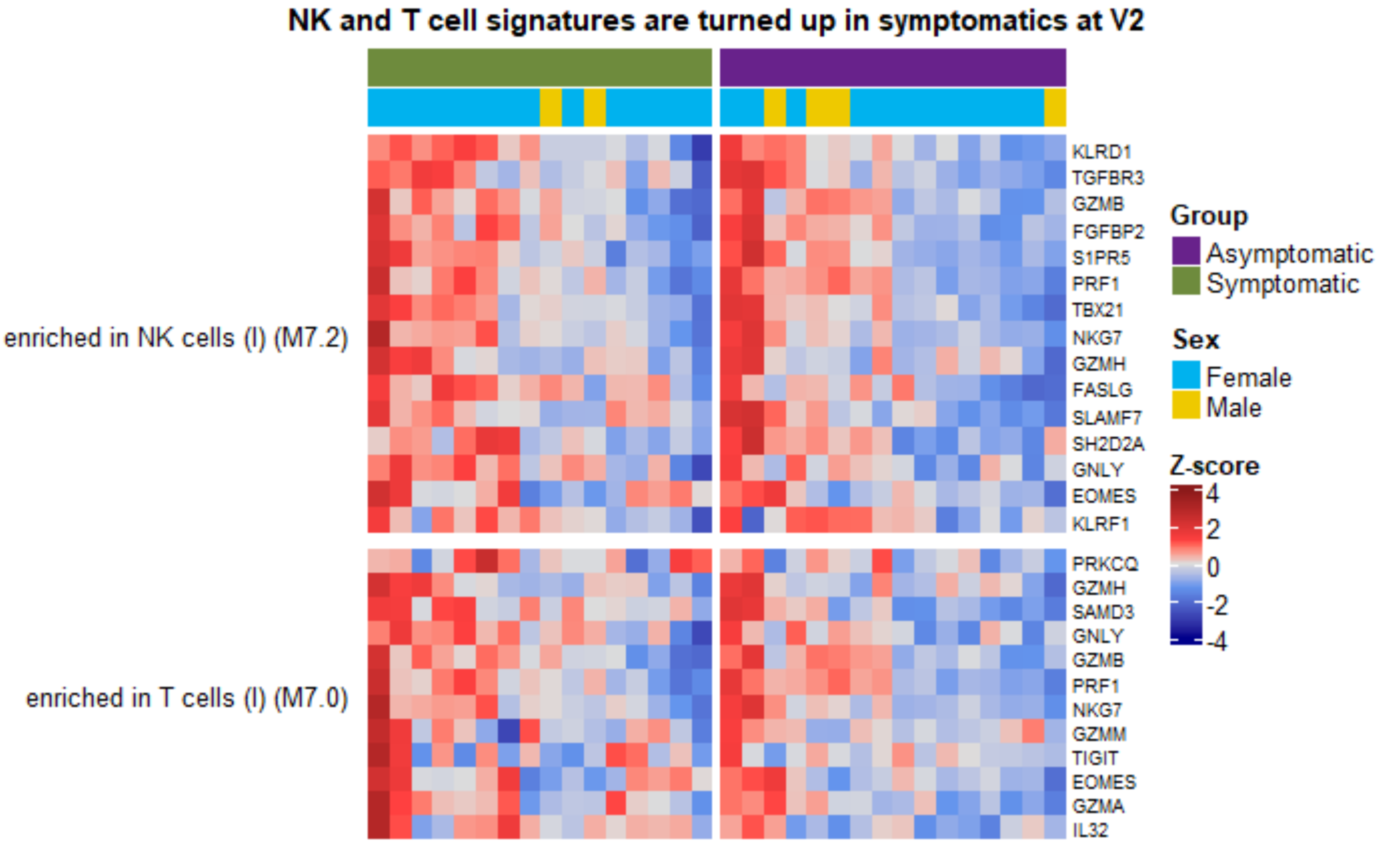
```
ggplot(subset(res.df, external_gene_name %in% btm[[fgseaRes[[6, 'pathway']]]]), aes(x = log2FoldChange, y = -log1
0(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + ggtitle('Neutrophil
genes at V2') + geom_vline(xintercept = 0, linetype = 2)
```



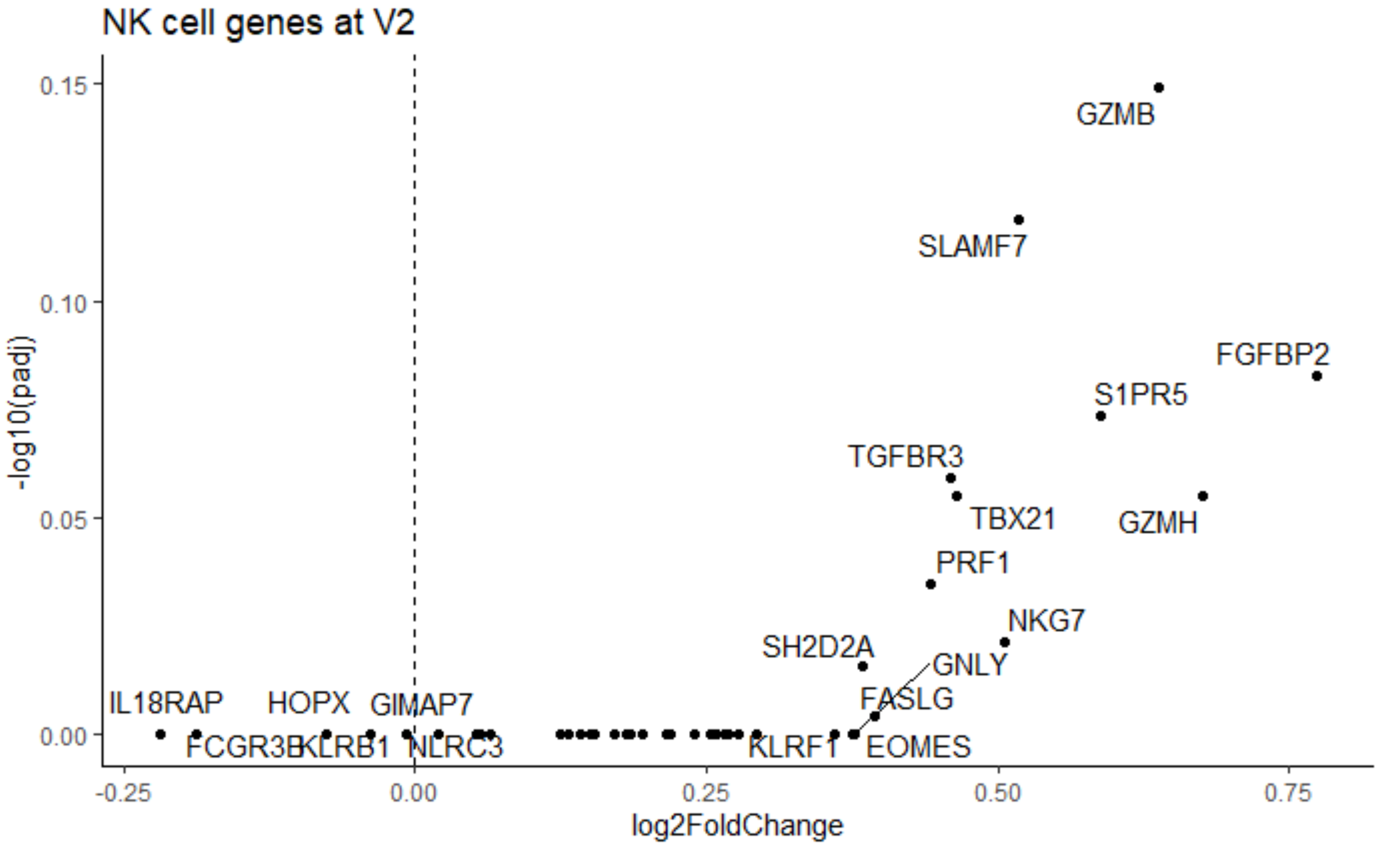
```
ggplot(subset(res.df, external_gene_name %in% btm[[fgseaRes[[9, 'pathway']]]]), aes(x = log2FoldChange, y = -log10(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + ggtitle('Monocyte genes at V2') + geom_vline(xintercept = 0, linetype = 2)
```


Monocyte genes at V2
-log10(padj)
0.10
0.05
0.00
log2FoldChange
-0.25
0.00
0.25
0.50
IMPA2
ASGR2
NFE2
SMPDL3A
SGK1
GLT1D1
IL13RA1
TBXAS1
SECTM1
TFEC
BLVRB
LILRA5
RNASE2
TMEM176A

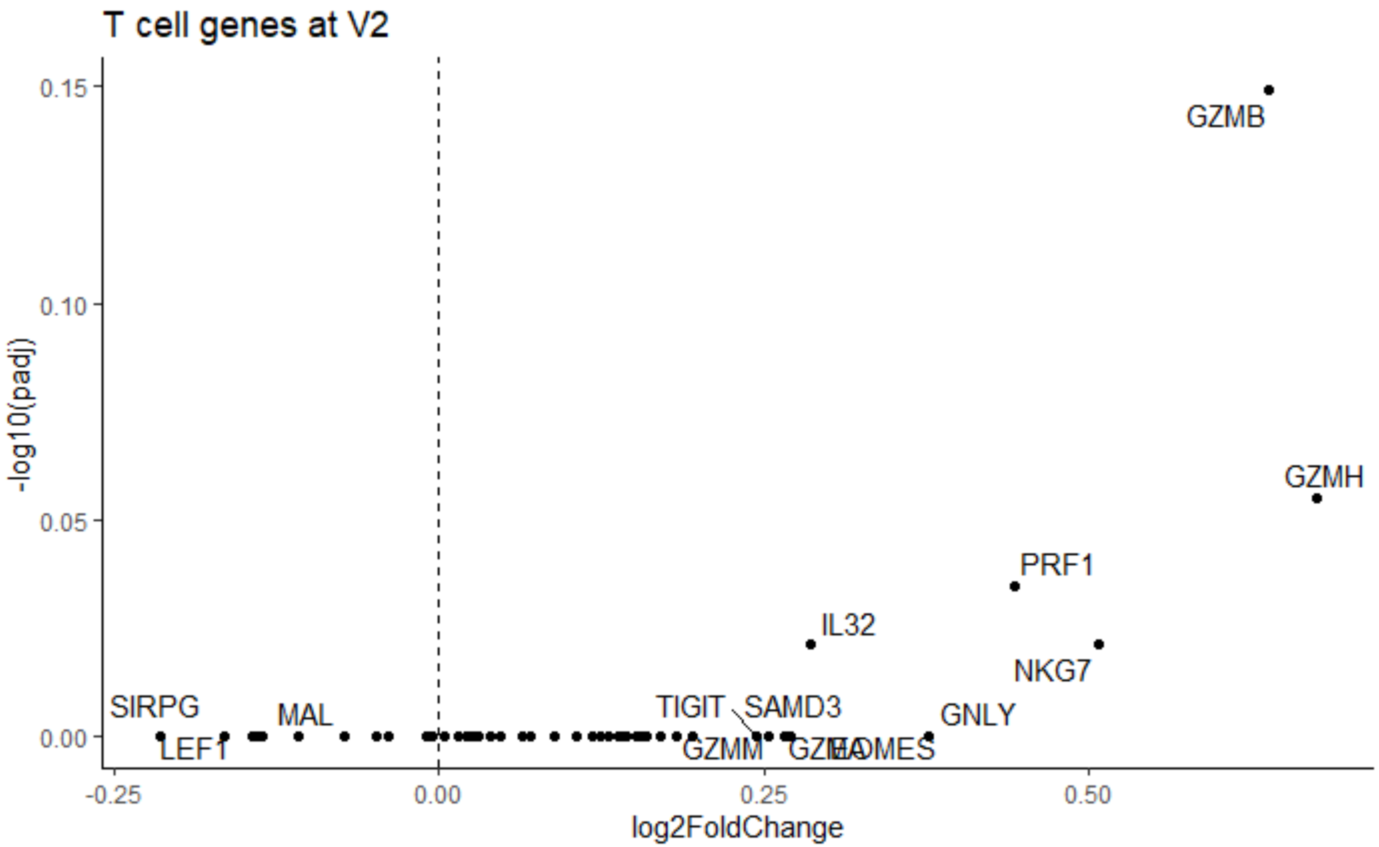
```
res.df = res.df[order(res.df$log2FoldChange, decreasing = TRUE), ]
plot4 = vsd.df[rownames(subset(res.df, external_gene_name %in% btm[[fgseaRes[[3, 'pathway']]]] & log2FoldChange > 0))[1:15], ]
zzz = t(scale(t(plot4))); zzz = pmin(pmax(zzz, -3), 3); rownames(zzz) =  res.df[rownames(zzz), 'external_gene_name']
plot4 = Heatmap(zzz, row_title = fgseaRes[[3, 'pathway']], show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = TRUE, show_row_dend = FALSE, show_column_dend = FALSE, column_split = meta_data$Group, top_annotation = ha_top, heatmap_legend_param = list(title = "Z-score"), col = colorRamp2(c(-4, -1.25, 0, 1.25, 4), c("darkblue", "cornflowerblue", "gainsboro", "brown1", "firebrick4")), row_names_gp = gpar(fontsize = 7), row_title_rot = 0, row_title_gp = gpar(fontsize = 10), column_title_gp = gpar(fontsize = 11, fontface = 2), heatmap_height = unit(7, 'cm'), column_title = 'NK and T cell signatures are turned up in symptomatics at V2')
plot5 = vsd.df[rownames(subset(res.df, external_gene_name %in% btm[[fgseaRes[[8, 'pathway']]]] & log2FoldChange > 0))[1:12], ]
zzz = t(scale(t(plot5))); zzz = pmin(pmax(zzz, -3), 3); rownames(zzz) =  res.df[rownames(zzz), 'external_gene_name']
plot5 = Heatmap(zzz, row_title = fgseaRes[[8, 'pathway']], show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = TRUE, show_row_dend = FALSE, show_column_dend = TRUE, column_split = meta_data$Group, heatmap_legend_param = list(title = "Z-score"), col = colorRamp2(c(-4, -1.25, 0, 1.25, 4), c("darkblue", "cornflowerblue", "gainsboro", "brown1", "firebrick4")), row_names_gp = gpar(fontsize = 7), row_title_rot = 0, row_title_gp = gpar(fontsize = 10), column_title_gp = gpar(fontsize = 11, fontface = 2), show_heatmap_legend = FALSE)
draw(plot4 %v% plot5)
```



```
ggplot(subset(res.df, external_gene_name %in% btm[[fgseaRes[[3, 'pathway']]]]), aes(x = log2FoldChange, y = -log10(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + ggtitle('NK cell genes at V2') + geom_vline(xintercept = 0, linetype = 2)
```



```
ggplot(subset(res.df, external_gene_name %in% btm[[fgseaRes[[8, 'pathway']]]]), aes(x = log2FoldChange, y = -log1
0(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + ggtitle('T cell gene
s at V2') + geom_vline(xintercept = 0, linetype = 2)
```



```
rm(plot1); rm(plot2); rm(plot3); rm(plot4); rm(plot5); rm(zzz); rm(ha_top)
```

5. V0 GSEA

```
keep = data.frame(cbind(Participant.ID = meta_data$Participant.ID, Group = meta_data$Group))
```

```
count_data = readRDS('data/Raw_RNASeq_Data.RDS')
ens2gene = readRDS("data/ens2gene.RDS")
meta_data = data.frame(read.csv('data/COVIRS_metadata.csv', header = TRUE))
meta_data = meta_data[meta_data$Timepoint == 'V0' & meta_data$First.Second.vaccine == 'BNT162b2', ]
meta_data = meta_data[meta_data$Participant.ID %in% keep$Participant.ID, ]; meta_data$Group = keep$Group

ensgene = rownames(count_data); count_data = count_data[, colnames(count_data) %in% meta_data$id]; count_data = data.frame(cbind(ensgene = ensgene, count_data)); rownames(count_data) = NULL; rm(ensgene); rm(keep)
```

```
dds = DESeqDataSetFromMatrix(countData = count_data,
                              colData = meta_data,
                              design= ~ Group + Sex, tidy = TRUE)
vsd = vst(dds, blind = FALSE); vsd.df = assay(vsd); write.csv(vsd.df, file = 'results/counts_norm_v0.csv', row.names = TRUE)
```

```
keep = rowSums(counts(dds) > 100) >= 4; dds = dds[keep, ]; dds = DESeq(dds)
```

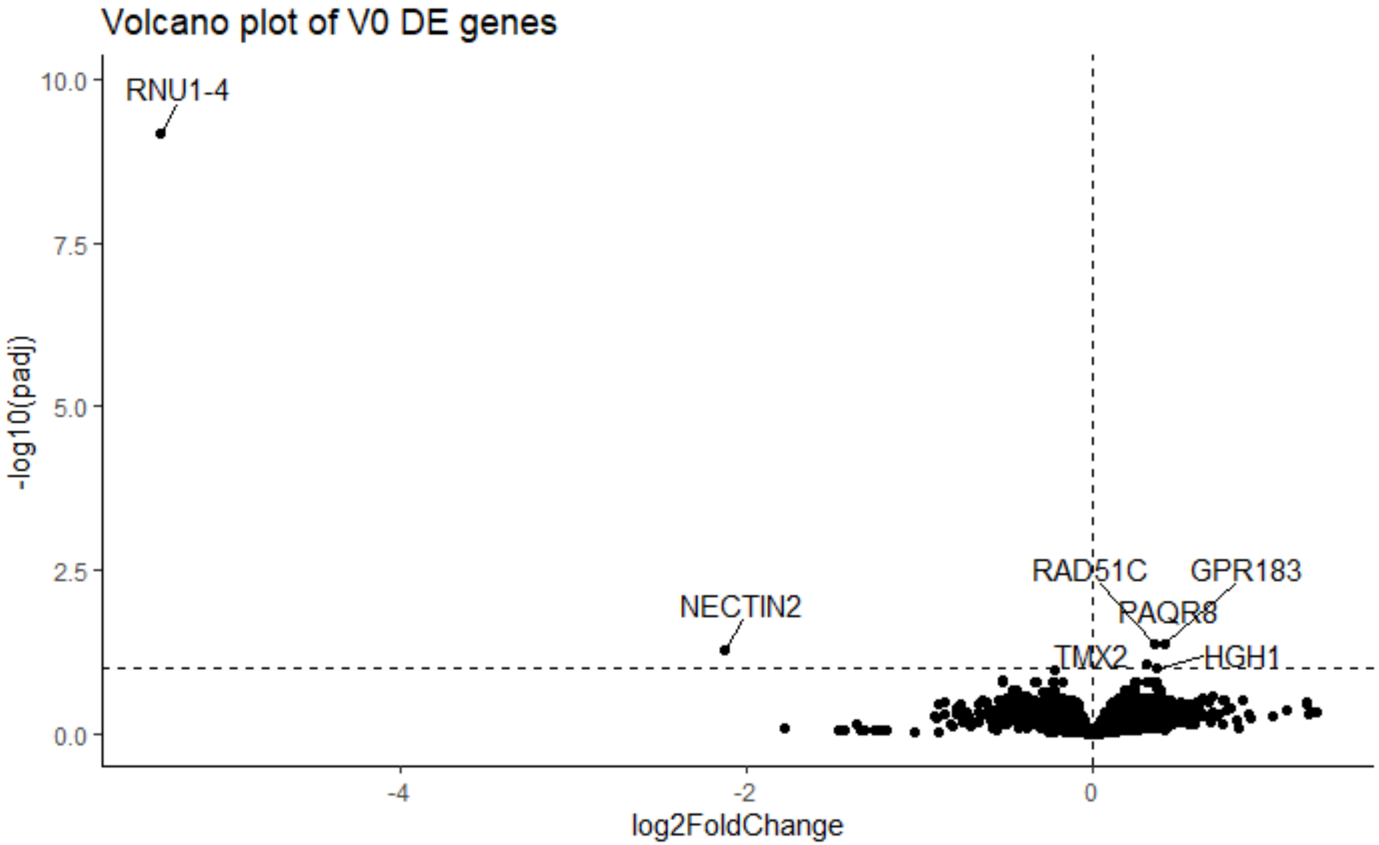
```
estimating size factors
estimating dispersions
gene-wise dispersion estimates
mean-dispersion relationship
final dispersion estimates
fitting model and testing
-- replacing outliers and refitting for 19 genes
-- DESeq argument 'minReplicatesForReplace' = 7
-- original counts are preserved in counts(dds)
estimating dispersions
fitting model and testing
```

```
res = results(dds, contrast = c("Group", "Symptomatic", "Asymptomatic")); res = res[order(res$padj, decreasing = FALSE), ]
res; summary(res); res.df = data.frame(res); res.df = data.frame(res.df, ens2gene[rownames(res.df), ]); write.csv(res.df, file = 'results/res_deseq_v0.csv', row.names = TRUE)
```

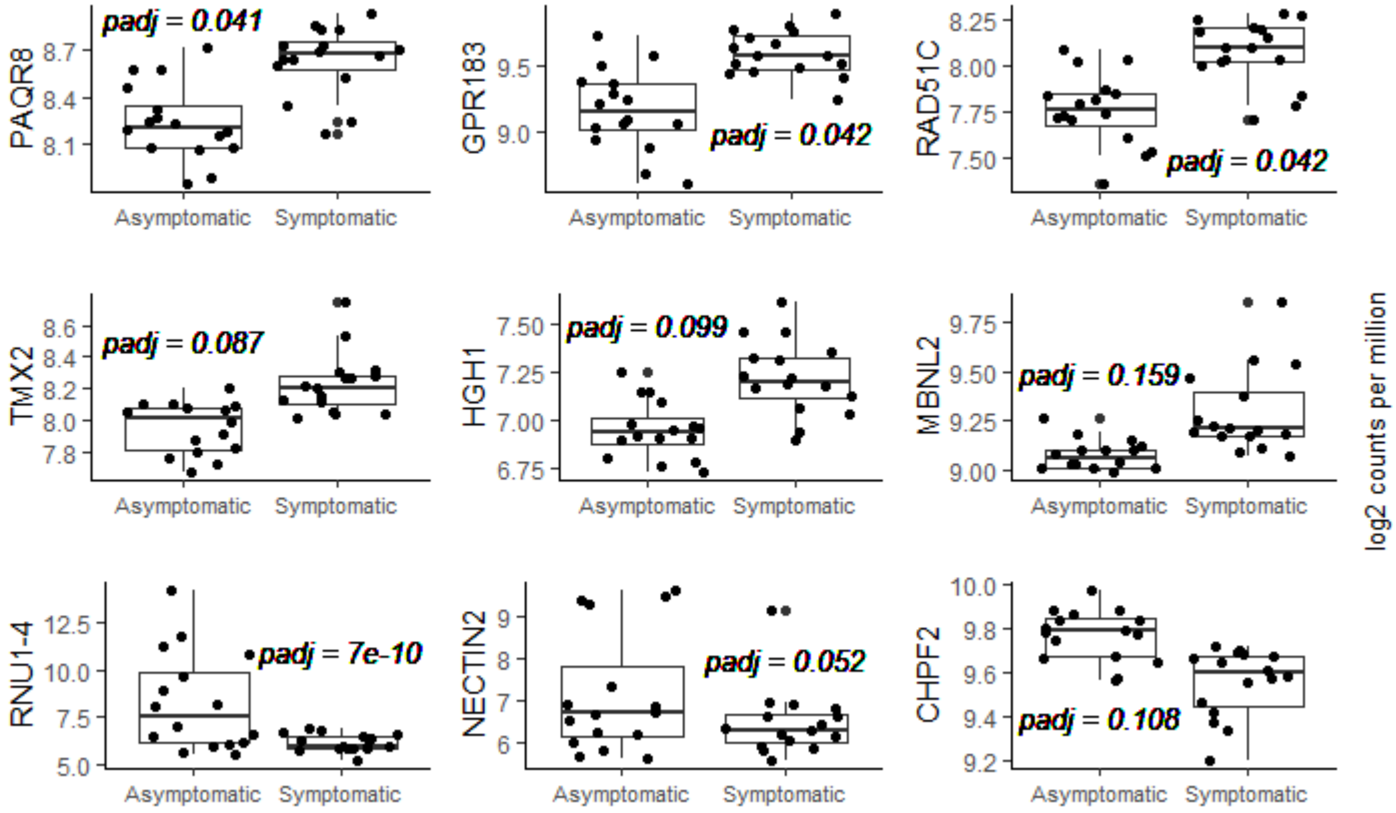
```
log2 fold change (MLE): Group Symptomatic vs Asymptomatic
Wald test p-value: Group Symptomatic vs Asymptomatic
DataFrame with 11229 rows and 6 columns
                 baseMean log2FoldChange     lfcSE       stat      pvalue        padj
                <numeric>      <numeric> <numeric>  <numeric>   <numeric>   <numeric>
ENSG00000207389   905.246      -5.386487 0.7176174   -7.50607 6.09286e-14 6.83863e-10
ENSG00000170915   313.277       0.422574 0.0942592    4.48310 7.35649e-06 4.12846e-02
ENSG00000169508   642.807       0.421562 0.0966748    4.36061 1.29698e-05 4.24441e-02
ENSG00000108384   204.533       0.363181 0.0839366    4.32684 1.51262e-05 4.24441e-02
ENSG00000130202   125.110      -2.122512 0.5016004   -4.23148 2.32157e-05 5.21147e-02
...                   ...            ...       ...        ...         ...         ...
ENSG00000123838   41.6755     -3.2523776  0.624424 -5.2086061          NA          NA
ENSG00000239899  251.3011      1.1974009  0.575360  2.0811347          NA          NA
ENSG00000233913  206.2764     -0.0132661  0.346732 -0.0382605          NA          NA
ENSG00000254612   75.9356     -0.1343210  0.199588 -0.6729914          NA          NA
ENSG00000078114   23.3453     -0.7735165  0.495343 -1.5615775          NA          NA

out of 11229 with nonzero total read count
adjusted p-value < 0.1
LFC > 0 (up)       : 5, 0.045%
LFC < 0 (down)     : 2, 0.018%
outliers [1]       : 5, 0.045%
low counts [2]     : 0, 0%
(mean count < 20)
[1] see 'cooksCutoff' argument of ?results
[2] see 'independentFiltering' argument of ?results
```

```
rm(keep)
```

```
ggplot(res.df, aes(x = log2FoldChange, y = -log10(padj))) + geom_point() +
  geom_hline(yintercept = -log10(0.1), linetype = 2) + geom_vline(xintercept = 0, linetype = 2) +
  theme_classic() + geom_text_repel(data = subset(res.df, -log10(padj) > 1), aes(x = log2FoldChange, y = -log10(padj), label = external_gene_name), max.overlaps = 9, position = position_nudge_repel(x = 0.1, y = 0.7)) + ggtitle("Volcano plot of V0 DE genes")
```

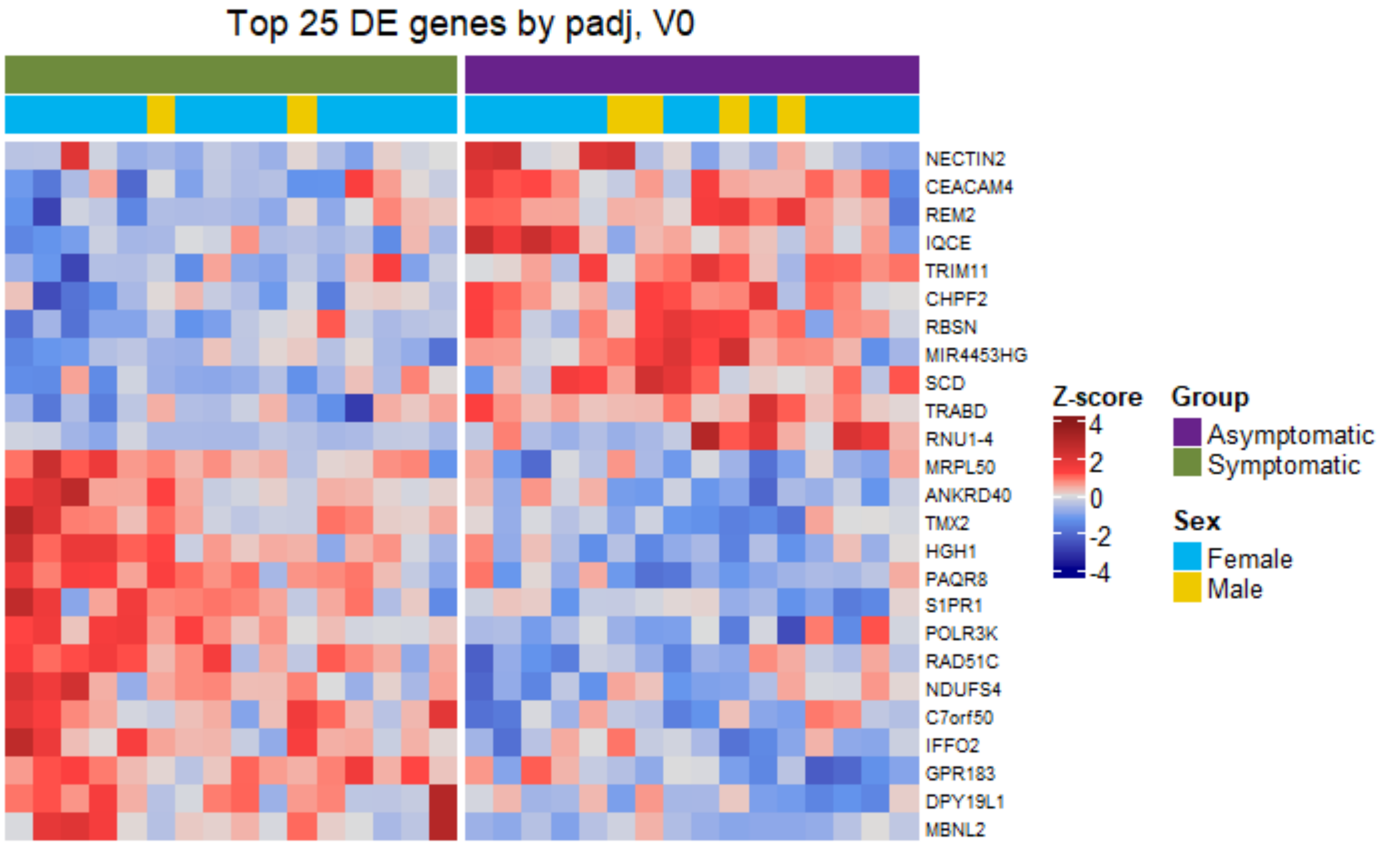

Volcano plot of V0 DE genes
-log10(padj)
10.0
7.5
5.0
2.5
0.0
log2FoldChange
-4
-2
0
RNU1-4
NECTIN2
RAD51C
GPR183
PAQR8
TMX2
HGH1

```
plot = rbind(filter(res.df, log2FoldChange > 0)[1:6, ], filter(res.df, log2FoldChange < 0)[1:3, ])
vsd.df = data.frame(t(vsd.df)); vsd.df$Group = meta_data$Group; vsd.df$Sex = meta_data$Sex

plot1 = ggplot(vsd.df, aes(x = Group, y = vsd.df[, rownames(plot)[1]])) + geom_boxplot() + geom_jitter() + ylab(r
es.df[rownames(plot)[1], 'external_gene_name']) + geom_text(aes(x = 1, y = 8.9, fontface = 3), label = paste0('pa
dj = ', as.character(round(res.df[rownames(plot)[1], 'padj'], 3)))) + theme_classic() + xlab('') + theme(axis.tex
t.x = element_text(size = 8))
plot2 = ggplot(vsd.df, aes(x = Group, y = vsd.df[, rownames(plot)[2]])) + geom_boxplot() + geom_jitter() + ylab(r
es.df[rownames(plot)[2], 'external_gene_name']) + geom_text(aes(x = 2, y = 9.0, fontface = 3), label = paste0('pa
dj = ', as.character(round(res.df[rownames(plot)[2], 'padj'], 3)))) + theme_classic() + xlab('') + theme(axis.tex
t.x = element_text(size = 8))
plot3 = ggplot(vsd.df, aes(x = Group, y = vsd.df[, rownames(plot)[3]])) + geom_boxplot() + geom_jitter() + ylab(r
es.df[rownames(plot)[3], 'external_gene_name']) + geom_text(aes(x = 2, y = 7.5, fontface = 3), label = paste0('pa
dj = ', as.character(round(res.df[rownames(plot)[3], 'padj'], 3)))) + theme_classic() + xlab('') + theme(axis.tex
t.x = element_text(size = 8))
plot4 = ggplot(vsd.df, aes(x = Group, y = vsd.df[, rownames(plot)[4]])) + geom_boxplot() + geom_jitter() + ylab(r
es.df[rownames(plot)[4], 'external_gene_name']) + geom_text(aes(x = 1, y = 8.5, fontface = 3), label = paste0('pa
dj = ', as.character(round(res.df[rownames(plot)[4], 'padj'], 3)))) + theme_classic() + xlab('') + theme(axis.tex
t.x = element_text(size = 8))
plot5 = ggplot(vsd.df, aes(x = Group, y = vsd.df[, rownames(plot)[5]])) + geom_boxplot() + geom_jitter() + ylab(r
es.df[rownames(plot)[5], 'external_gene_name']) + geom_text(aes(x = 1, y = 7.5, fontface = 3), label = paste0('pa
dj = ', as.character(round(res.df[rownames(plot)[5], 'padj'], 3)))) + theme_classic() + xlab('') + theme(axis.tex
t.x = element_text(size = 8))
plot6 = ggplot(vsd.df, aes(x = Group, y = vsd.df[, rownames(plot)[6]])) + geom_boxplot() + geom_jitter() + ylab(r
es.df[rownames(plot)[6], 'external_gene_name']) + geom_text(aes(x = 1, y = 9.5, fontface = 3), label = paste0('pa
dj = ', as.character(round(res.df[rownames(plot)[6], 'padj'], 3)))) + theme_classic() + xlab('') + theme(axis.tex
t.x = element_text(size = 8))
plot7 = ggplot(vsd.df, aes(x = Group, y = vsd.df[, rownames(plot)[7]])) + geom_boxplot() + geom_jitter() + ylab(r
es.df[rownames(plot)[7], 'external_gene_name']) + geom_text(aes(x = 2, y = 11, fontface = 3), label = paste0('pad
j = ', as.character(round(res.df[rownames(plot)[7], 'padj'], 10)))) + theme_classic() + xlab('') + theme(axis.tex
t.x = element_text(size = 8))
plot8 = ggplot(vsd.df, aes(x = Group, y = vsd.df[, rownames(plot)[8]])) + geom_boxplot() + geom_jitter() + ylab(r
es.df[rownames(plot)[8], 'external_gene_name']) + geom_text(aes(x = 2, y = 8, fontface = 3), label = paste0('padj
= ', as.character(round(res.df[rownames(plot)[8], 'padj'], 3)))) + theme_classic() + xlab('') + theme(axis.text.x
= element_text(size = 8))
plot9 = ggplot(vsd.df, aes(x = Group, y = vsd.df[, rownames(plot)[9]])) + geom_boxplot() + geom_jitter() + ylab(r
es.df[rownames(plot)[9], 'external_gene_name']) + geom_text(aes(x = 1, y = 9.4, fontface = 3), label = paste0('pa
dj = ', as.character(round(res.df[rownames(plot)[9], 'padj'], 3)))) + theme_classic() + xlab('') + theme(axis.text.x = element_text(size = 8))
grid.arrange(plot1, plot2, plot3, plot4, plot5, plot6, plot7, plot8, plot9, ncol = 3, right = textGrob('log2 counts per million', rot = 90, gp = gpar(fontsize = 10)))
```



```
vsd.df$Sex = NULL; vsd.df$Group = NULL; vsd.df = data.frame(t(vsd.df))
rm(plot); rm(plot1); rm(plot2); rm(plot3); rm(plot4); rm(plot5); rm(plot6)
```

```
ha_top = HeatmapAnnotation(df = meta_data[, c('Group', 'Sex')], col = list(Group = c('Symptomatic' = 'darkolivegreen4', 'Asymptomatic' = 'darkorchid4'), Sex = c('Male' = 'gold2', 'Female' = 'deepskyblue2')), show_annotation_name = FALSE)
zzz = t(scale(t(vsd.df[rownames(res.df[1:25, ]), ]))); zzz = pmin(pmax(zzz, -3), 3); rownames(zzz) =  res.df[rownames(zzz), 'external_gene_name']
Heatmap(zzz, show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = TRUE, show_row_dend = FALSE, show_column_dend = FALSE, top_annotation = ha_top, heatmap_legend_param = list(title = "Z-score"), col = colorRamp2(c(-4, -1.25, 0, 1.25, 4), c("darkblue", "cornflowerblue", "gainsboro", "brown1", "firebrick4")), row_names_gp = gpar(fontsize = 7), column_title = 'Top 25 DE genes by padj, V0', column_split = meta_data$Group)
```

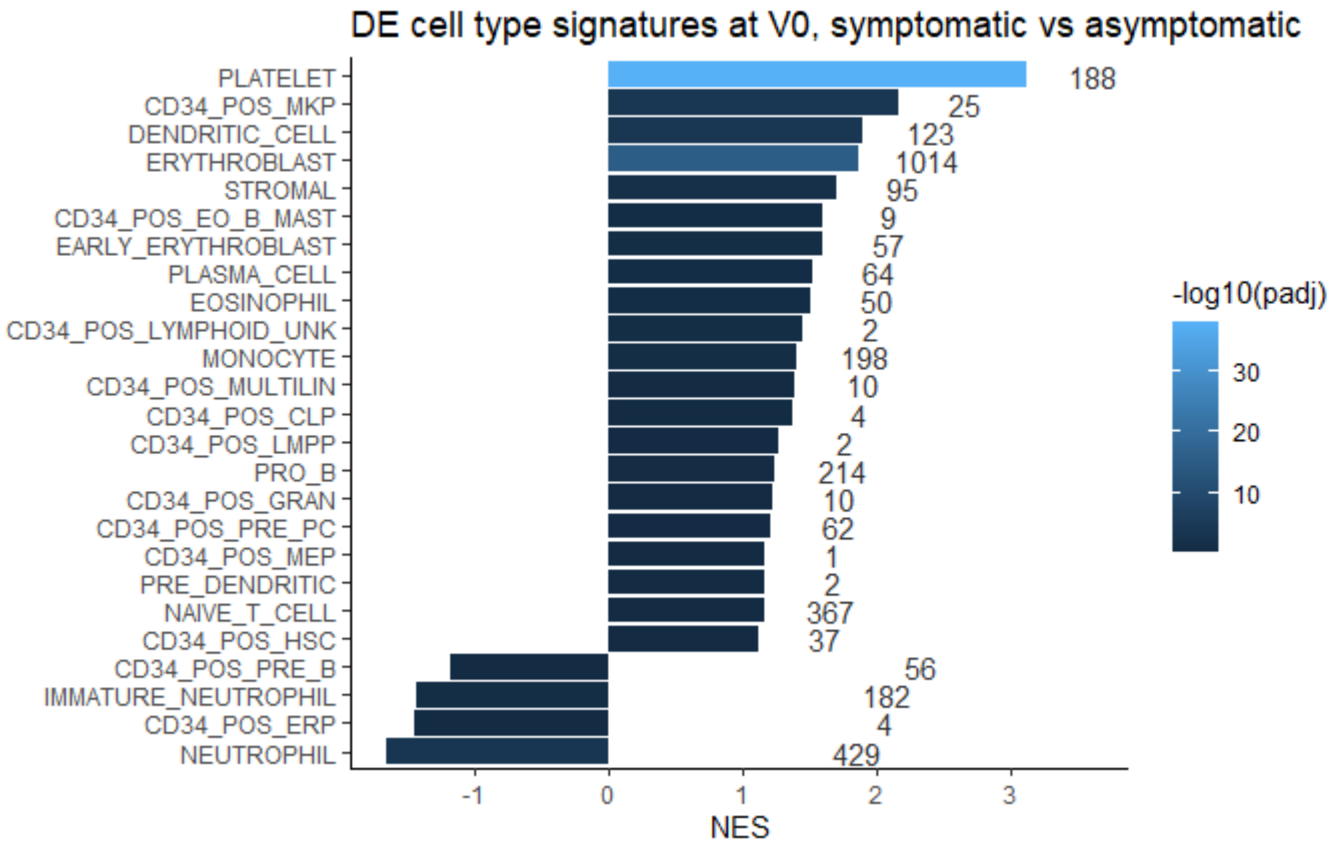


```
rm(ha_top); rm(zzz)
```

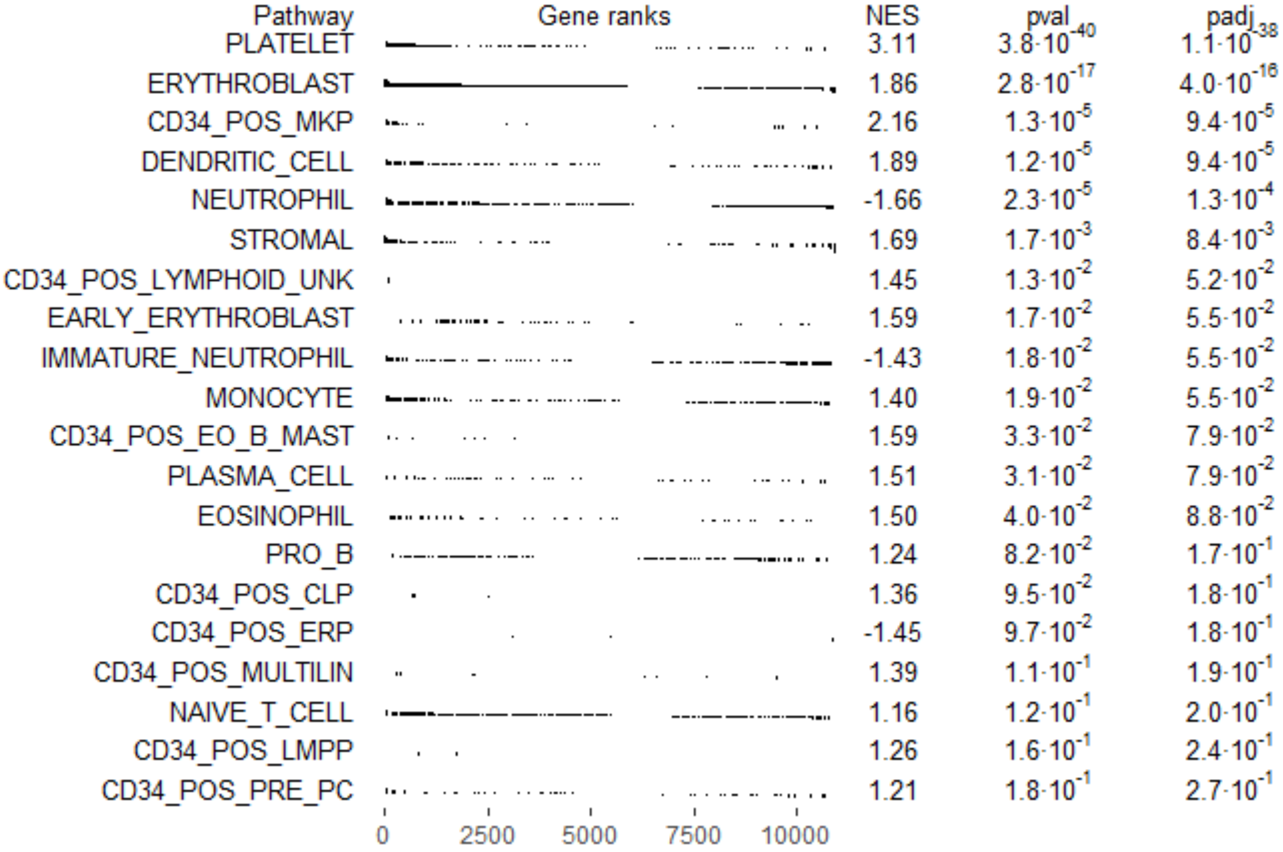
```
ranks <- res.df %>%
  as_tibble() %>%
  dplyr::select(external_gene_name, log2FoldChange) %>%
  na.omit() %>%
  distinct() %>%
  group_by(external_gene_name) %>%
  summarize(log2FoldChange=mean(log2FoldChange))
ranks <- deframe(ranks)
```

```
msig.df = as.data.frame(msigdbr(species = "Homo sapiens", category = "C8")); msig.df = msig.df[grepl('HAY_BONE_MARROW', msig.df$gs_name), ]
msigDB = by(msig.df$gene_symbol, msig.df$gs_name, function(x) as.character(x)); names(msigDB) = gsub('HAY_BONE_MARROW_', '', names(msigDB))
fgseaRes = fgsea(pathways = msigDB, stats = ranks, nproc = 1); fgseaRes = fgseaRes[order(fgseaRes$padj, decreasing = FALSE), ]
```

```
write.csv(subset(fgseaRes, select = -c(leadingEdge)), file = 'results/fgseaRes_C8_marrow_v0.csv', row.names = FALSE)

ggplot(fgseaRes[1:25, ], aes(y = reorder(pathway, NES), x = NES, fill = -log10(padj), label = size)) +
  geom_col() + geom_text(color = 'gray20', nudge_x = ifelse(fgseaRes$NES > 0, 0.5, 3.5)) +
  ylab('') + theme_classic() + ggtitle('DE cell type signatures at V0, symptomatic vs asymptomatic')
```

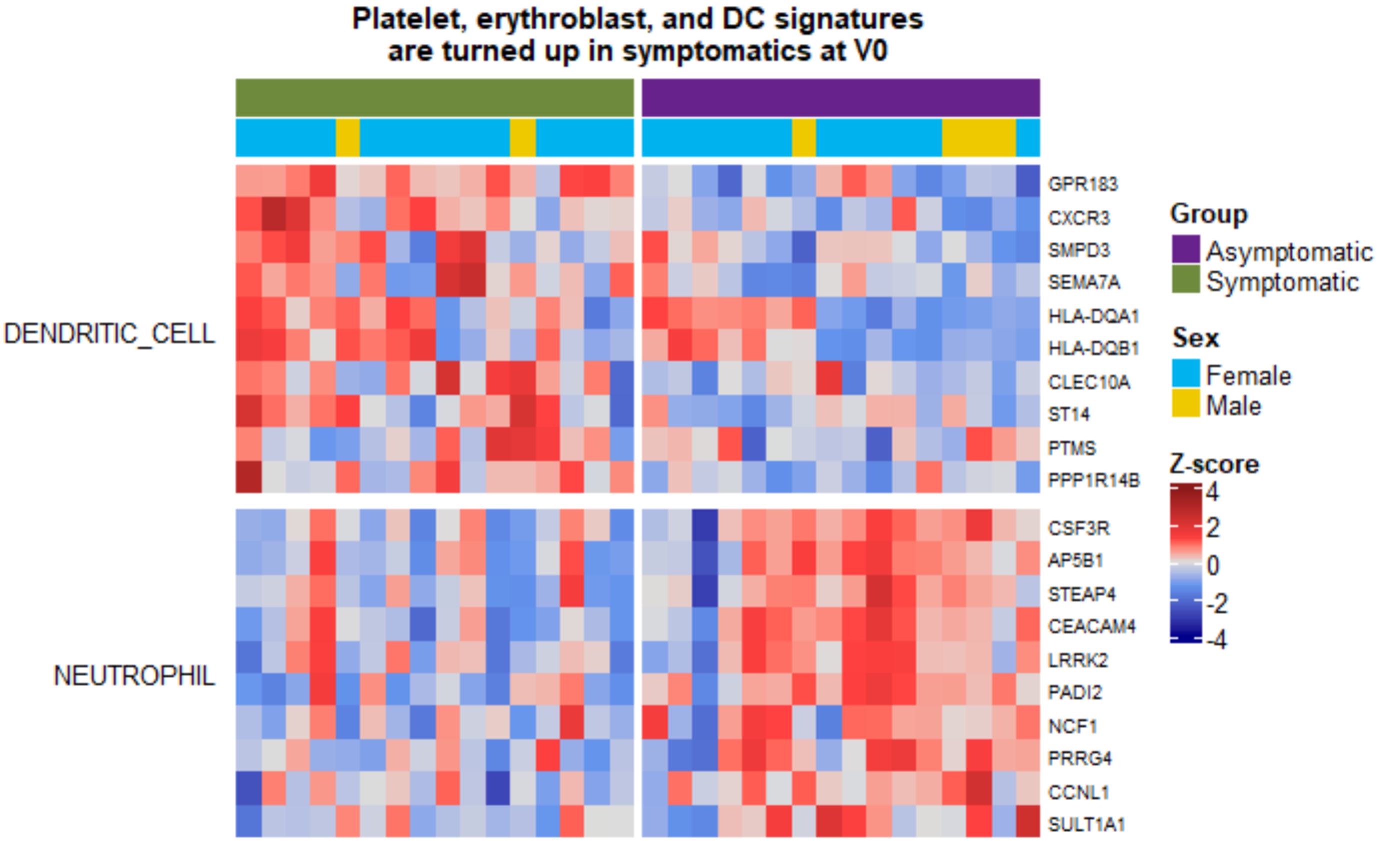


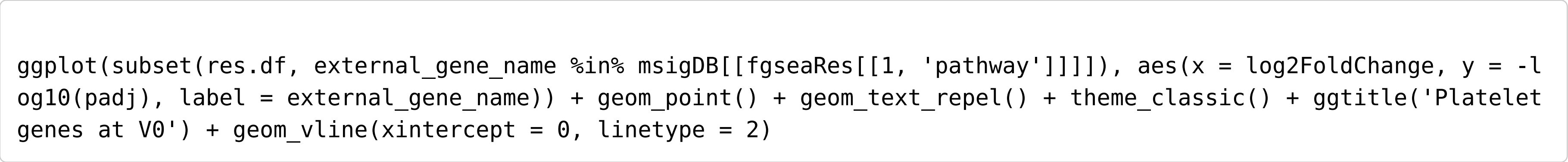
```
plotGseaTable(msigDB[fgseaRes[1:20, ][, pathway]], ranks, fgseaRes[1:20, ], gseaParam = 0.5, colwidths = c(3, 3,
0.8, 1.2, 1.2), pathwayLabelStyle = list(size = 10), headerLabelStyle = list(size = 10), valueStyle = list(size =
10))
```



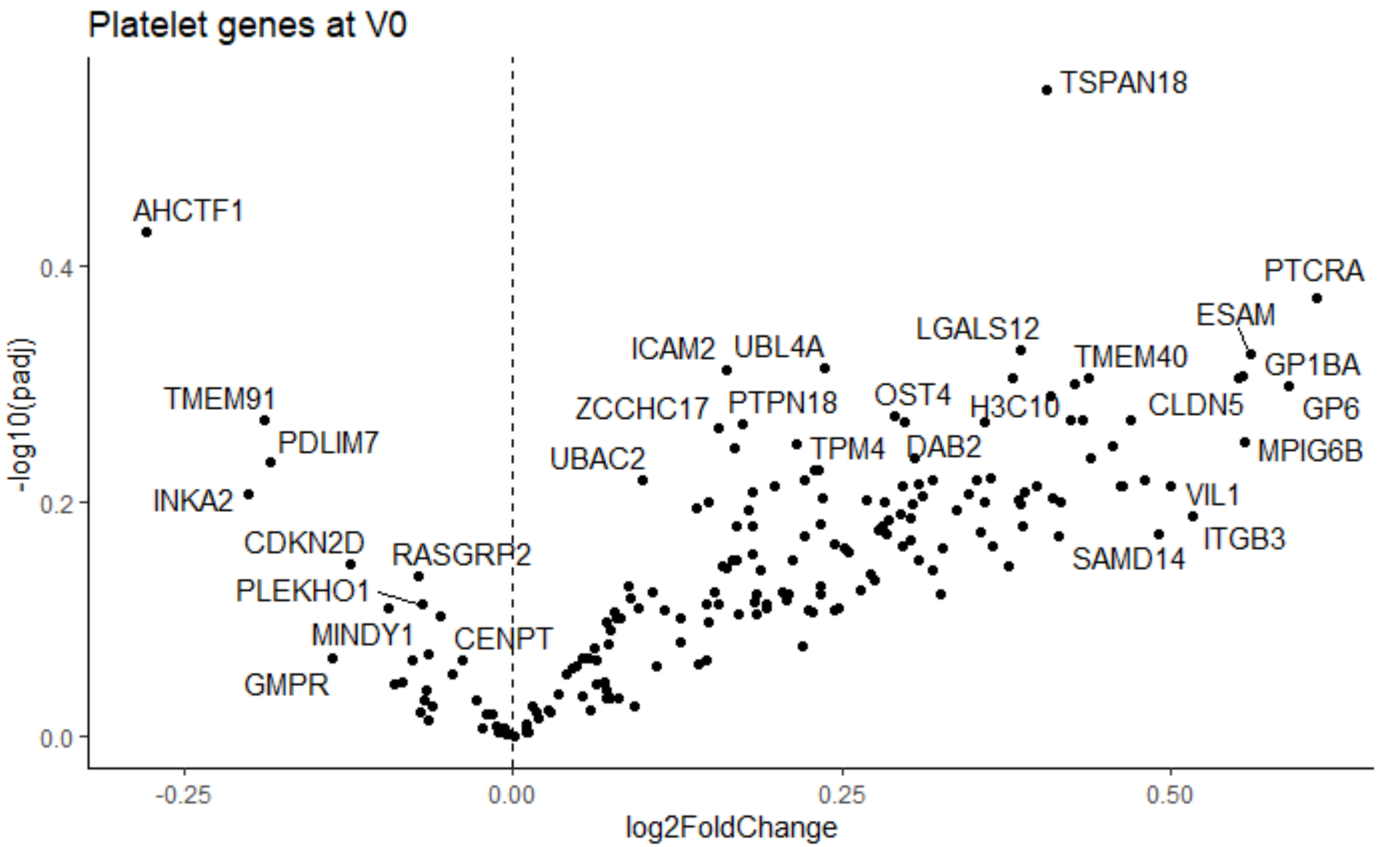
Pathway	NES	pval	padj
PLATELET	3.11	$3.8 \cdot 10^{-40}$	$1.1 \cdot 10^{-38}$
ERYTHROBLAST	1.86	$2.8 \cdot 10^{-17}$	$4.0 \cdot 10^{-16}$
CD34_POS_MKP	2.16	$1.3 \cdot 10^{-5}$	$9.4 \cdot 10^{-5}$
DENDRITIC_CELL	1.89	$1.2 \cdot 10^{-5}$	$9.4 \cdot 10^{-5}$
NEUTROPHIL	-1.66	$2.3 \cdot 10^{-5}$	$1.3 \cdot 10^{-4}$
STROMAL	1.69	$1.7 \cdot 10^{-3}$	$8.4 \cdot 10^{-3}$
CD34_POS_LYMPHOID_UNK	1.45	$1.3 \cdot 10^{-2}$	$5.2 \cdot 10^{-2}$
EARLY_ERYTHROBLAST	1.59	$1.7 \cdot 10^{-2}$	$5.5 \cdot 10^{-2}$
IMMATURE_NEUTROPHIL	-1.43	$1.8 \cdot 10^{-2}$	$5.5 \cdot 10^{-2}$
MONOCYTE	1.40	$1.9 \cdot 10^{-2}$	$5.5 \cdot 10^{-2}$
CD34_POS_EO_B_MAST	1.59	$3.3 \cdot 10^{-2}$	$7.9 \cdot 10^{-2}$
PLASMA_CELL	1.51	$3.1 \cdot 10^{-2}$	$7.9 \cdot 10^{-2}$
EOSINOPHIL	1.50	$4.0 \cdot 10^{-2}$	$8.8 \cdot 10^{-2}$
PRO_B	1.24	$8.2 \cdot 10^{-2}$	$1.7 \cdot 10^{-1}$
CD34_POS_CLP	1.36	$9.5 \cdot 10^{-2}$	$1.8 \cdot 10^{-1}$
CD34_POS_ERP	-1.45	$9.7 \cdot 10^{-2}$	$1.8 \cdot 10^{-1}$
CD34_POS_MULTILIN	1.39	$1.1 \cdot 10^{-1}$	$1.9 \cdot 10^{-1}$
NAIVE_T_CELL	1.16	$1.2 \cdot 10^{-1}$	$2.0 \cdot 10^{-1}$
CD34_POS_LMPP	1.26	$1.6 \cdot 10^{-1}$	$2.4 \cdot 10^{-1}$
CD34_POS_PRE_PC	1.21	$1.8 \cdot 10^{-1}$	$2.7 \cdot 10^{-1}$

```
res.df = res.df[order(res.df$log2FoldChange, decreasing = TRUE), ]
ha_top = HeatmapAnnotation(df = meta_data[, c('Group', 'Sex')], col = list(Group = c('Symptomatic' = 'darkolivegreen4', 'Asymptomatic' = 'darkorchid4'), Sex = c('Male' = 'gold2', 'Female' = 'deepskyblue2')), show_annotation_name = FALSE)
plot1 = vsd.df[rownames(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[4, 'pathway']]]] & log2FoldChange > 0))[1:10], ]
zzz = t(scale(t(plot1))); zzz = pmin(pmax(zzz, -3), 3); rownames(zzz) =  res.df[rownames(zzz), 'external_gene_name']
plot1 = Heatmap(zzz, row_title = fgseaRes[[4, 'pathway']], show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = TRUE, show_row_dend = FALSE, show_column_dend = FALSE, column_split = meta_data$Group, top_annotation = ha_top, heatmap_legend_param = list(title = "Z-score"), col = colorRamp2(c(-4, -1.25, 0, 1.25, 4), c("darkblue", "cornflowerblue", "gainsboro", "brown1", "firebrick4")), row_names_gp = gpar(fontsize = 7), row_title_rot = 0, row_title_gp = gpar(fontsize = 10), column_title_gp = gpar(fontsize = 11, fontface = 2), column_title = 'Platelet, erythroblast, and DC signatures\nare turned up in symptomatics at V0')
res.df = res.df[order(res.df$log2FoldChange, decreasing = FALSE), ]
plot2 = vsd.df[rownames(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[5, 'pathway']]]] & log2FoldChange < 0))[1:10], ]
zzz = t(scale(t(plot2))); zzz = pmin(pmax(zzz, -3), 3); rownames(zzz) =  res.df[rownames(zzz), 'external_gene_name']
plot2 = Heatmap(zzz, show_heatmap_legend = FALSE, row_title = fgseaRes[[5, 'pathway']], show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = TRUE, show_row_dend = FALSE, show_column_dend = FALSE, column_split = meta_data$Group, heatmap_legend_param = list(title = "Z-score"), col = colorRamp2(c(-4, -1.25, 0, 1.25, 4), c("darkblue", "cornflowerblue", "gainsboro", "brown1", "firebrick4")), row_names_gp = gpar(fontsize = 7), row_title_rot = 0, row_title_gp = gpar(fontsize = 10), column_title_gp = gpar(fontsize = 11, fontface = 2))
draw(plot1 %v% plot2); rm(plot1); rm(plot2); rm(zzz); rm(ha_top)
```

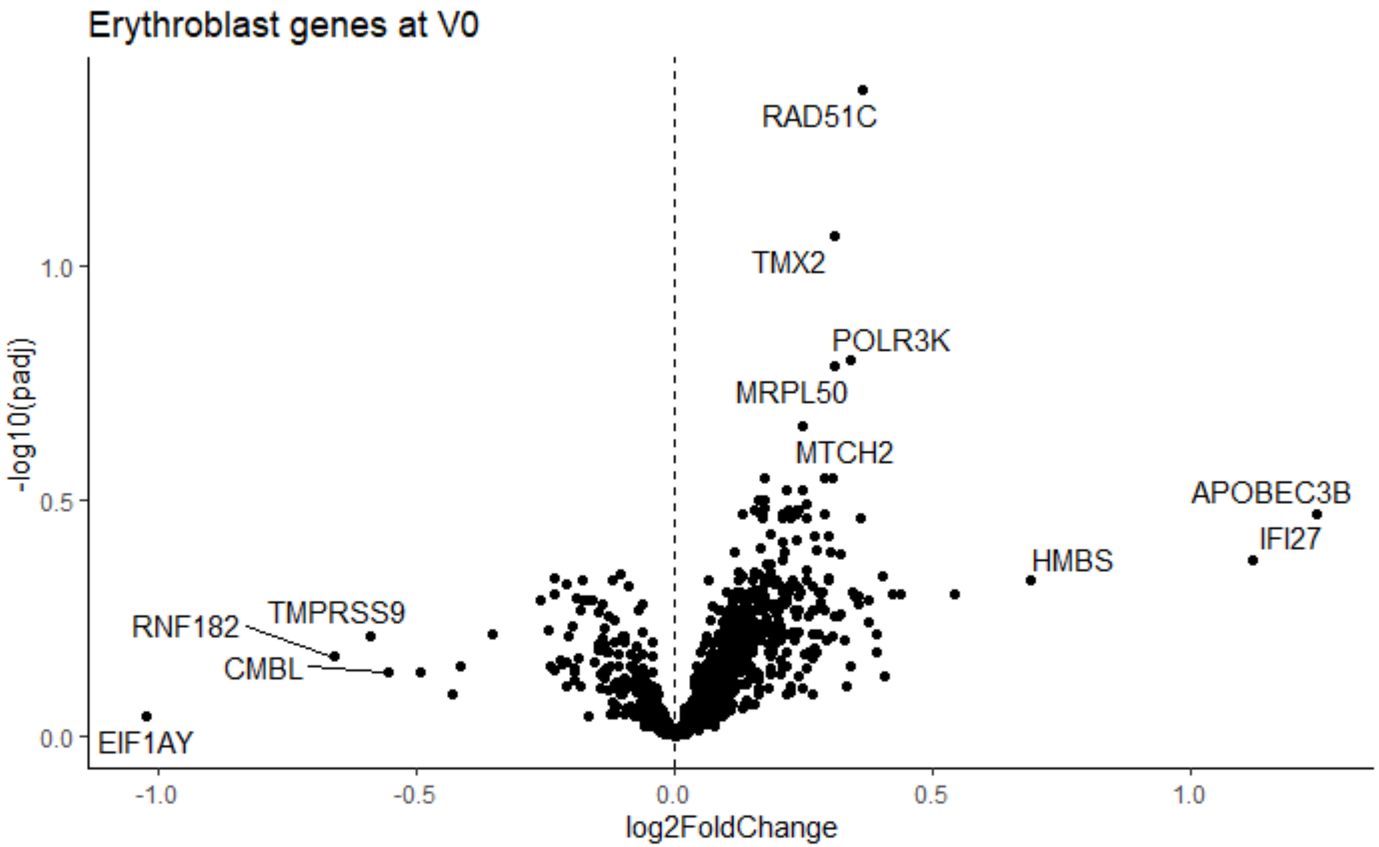




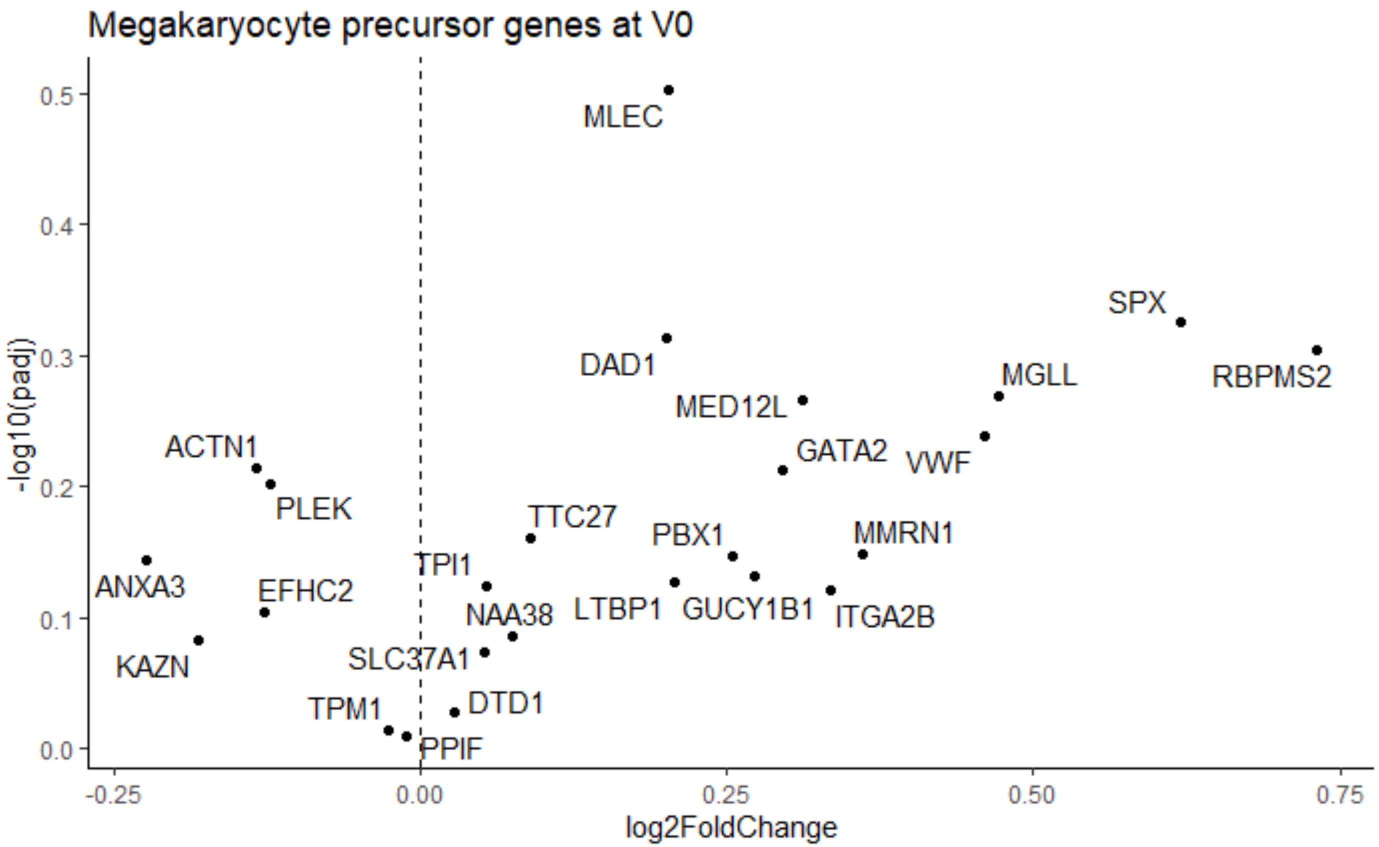
```
ggplot(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[1, 'pathway']]]]), aes(x = log2FoldChange, y = -l
og10(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + ggtitle('Platelet
genes at V0') + geom_vline(xintercept = 0, linetype = 2)
```



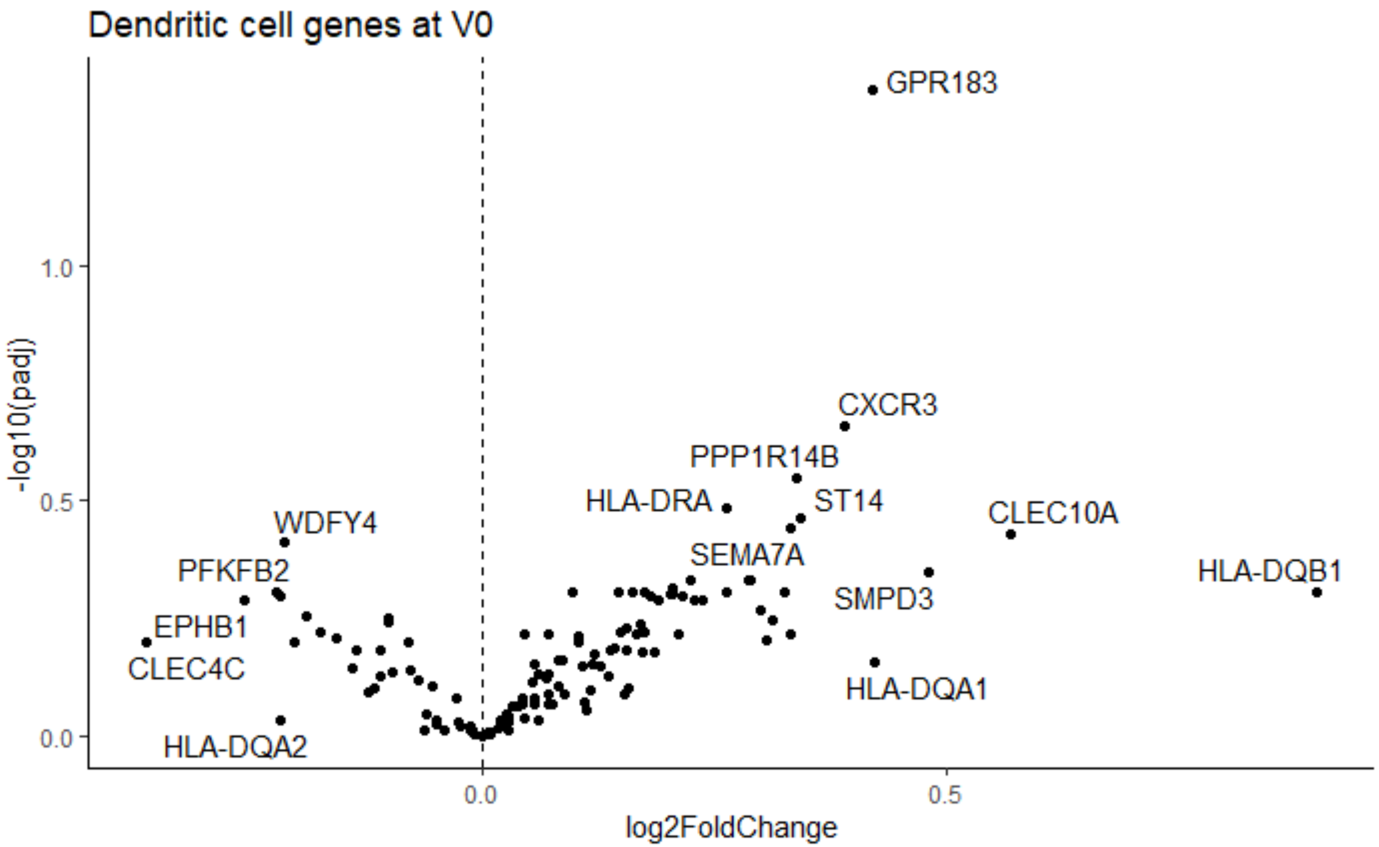
```
ggplot(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[2, 'pathway']]]]), aes(x = log2FoldChange, y = -l
og10(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + ggtitle('Erythrob
last genes at V0') + geom_vline(xintercept = 0, linetype = 2)
```



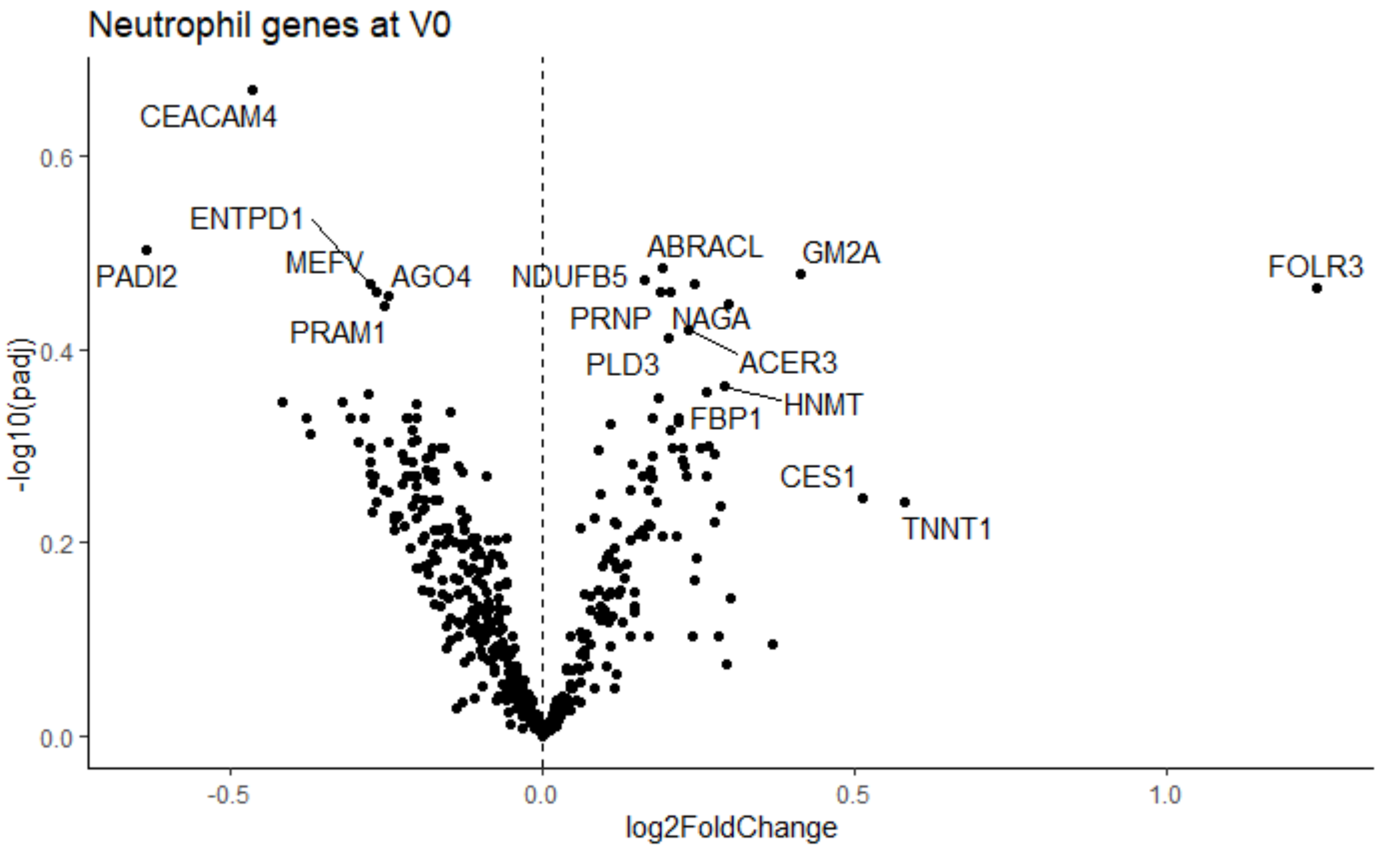
```
ggplot(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[3, 'pathway']]]]), aes(x = log2FoldChange, y = -log10(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + ggtitle('Megakaryocyte precursor genes at V0') + geom_vline(xintercept = 0, linetype = 2)
```



```
ggplot(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[4, 'pathway']]]]), aes(x = log2FoldChange, y = -l
og10(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + ggtitle('Dendriti
c cell genes at V0') + geom_vline(xintercept = 0, linetype = 2)
```



```
ggplot(subset(res.df, external_gene_name %in% msigDB[[fgseaRes[[5, 'pathway']]]]), aes(x = log2FoldChange, y = -log10(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + ggtitle('Neutrophil genes at V0') + geom_vline(xintercept = 0, linetype = 2)
```

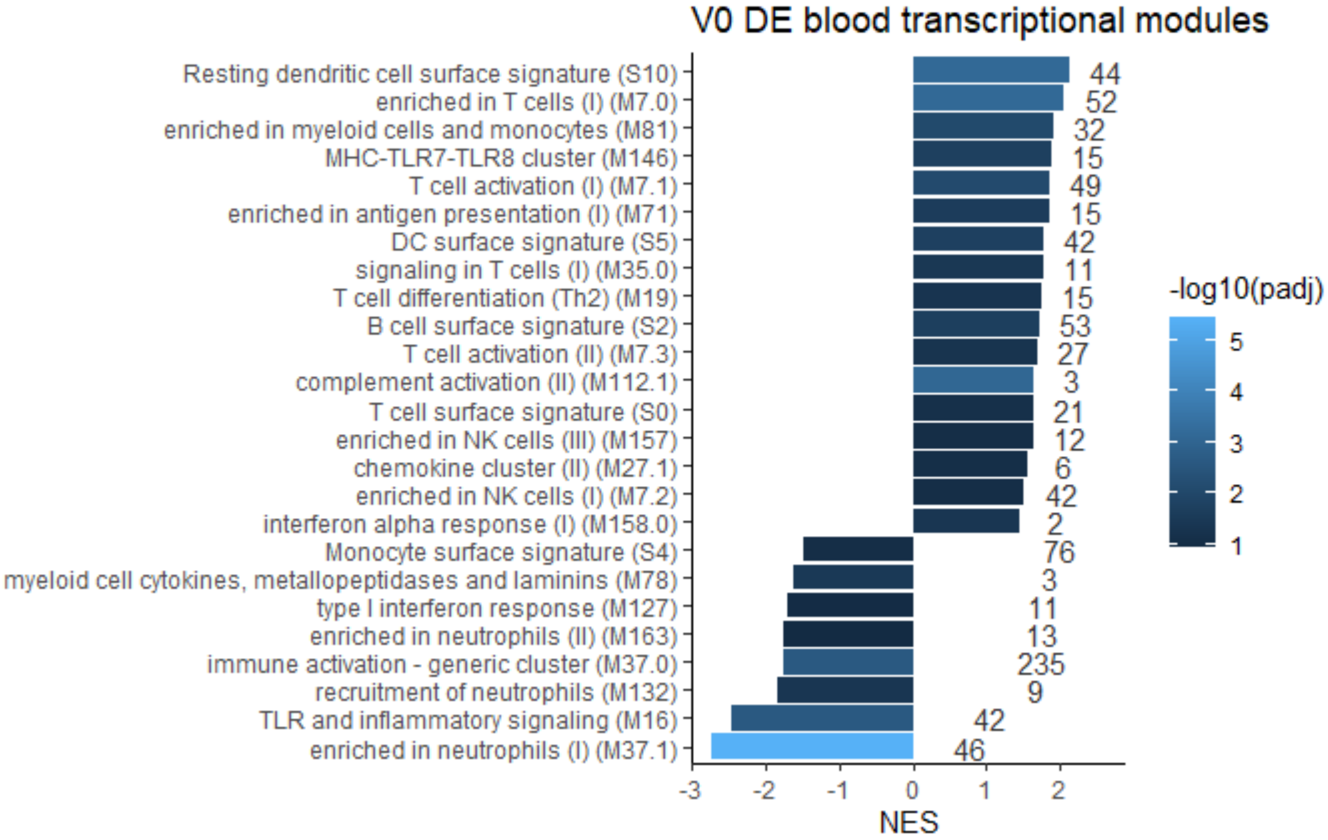


```
btm = readRDS('data/Blood_Transcript_modules.RDS'); btm_annot = read_excel('data/btm_annotation_table.xls')
```

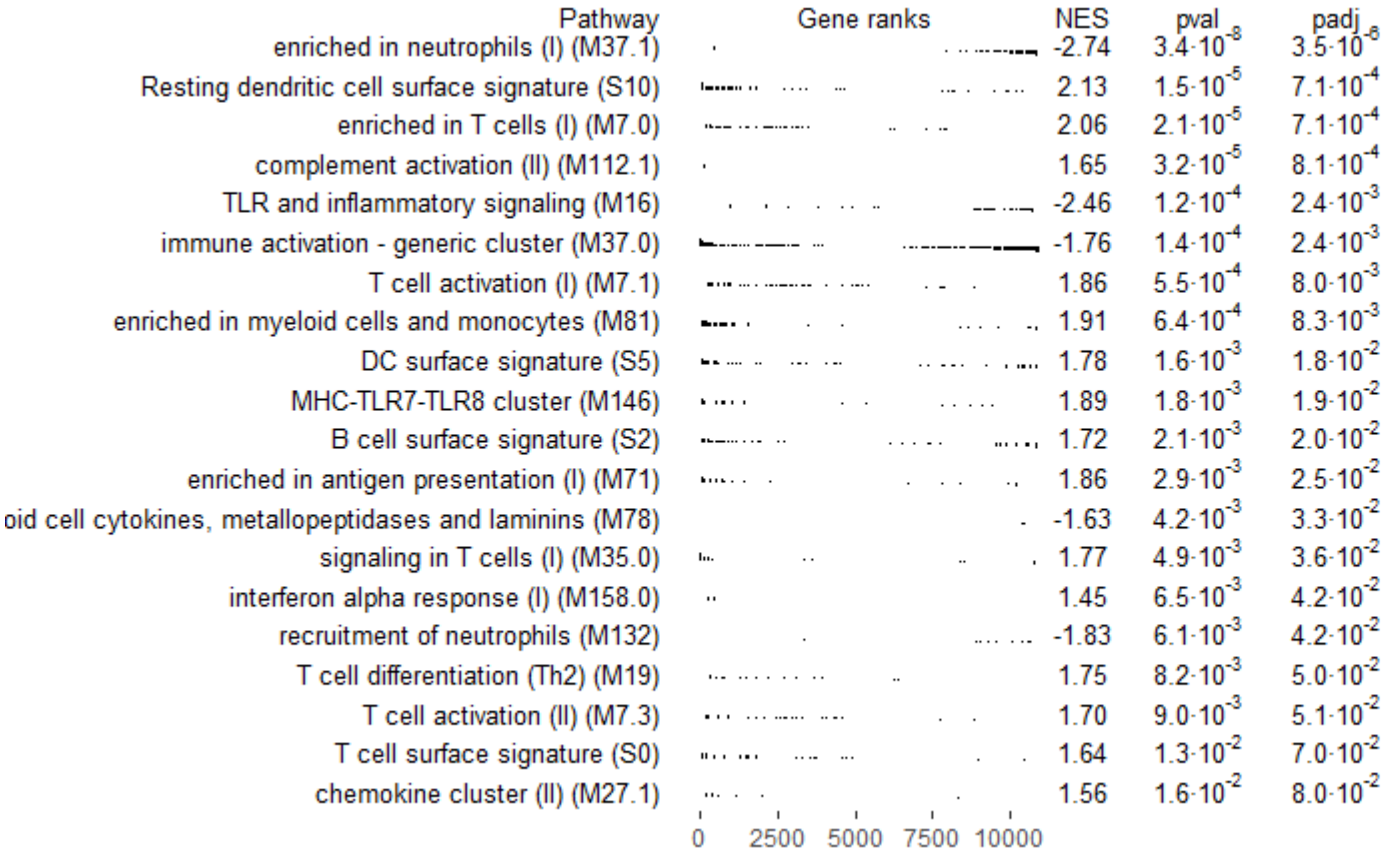
```
New names:
• `Module size` -> `Module size...4`
• `Jaccard Index` -> `Jaccard Index...10`
• `Enrichment p value` -> `Enrichment p value...11`
• `Overlap size` -> `Overlap size...13`
• `Jaccard Index` -> `Jaccard Index...15`
• `Enrichment p value` -> `Enrichment p value...16`
• `Module size` -> `Module size...17`
• `Overlap size` -> `Overlap size...18`
• `Jaccard Index` -> `Jaccard Index...20`
• `Enrichment p value` -> `Enrichment p value...21`
• `Overlap size` -> `Overlap size...23`
```

```
btm_annot = subset(btm_annot, `Module category` == 'immune'); btm = btm[c(btm_annot$ID)]; names(btm) = btm_annot$`Composite name`
fgseaRes = fgsea(pathways = btm, stats = ranks, nproc = 1); fgseaRes = fgseaRes[order(fgseaRes$padj, decreasing = FALSE), ]
```

```
write.csv(subset(fgseaRes, select = -c(leadingEdge)), file = 'results/fgseaRes_btm_immune_v0.csv', row.names = FALSE)

fgseaRes = fgseaRes[pathway %in% subset(btm_annot, `Module category` == 'immune')$`Composite name`, ]
ggplot(fgseaRes[1:25, ], aes(y = reorder(pathway, NES), x = NES, fill = -log10(padj), label = size)) + geom_col() + geom_text(color = 'gray20', nudge_x = ifelse(fgseaRes$NES > 0, 0.5, 3.5)) + ylab('') + theme_classic() + ggtitle('V0 DE blood transcriptional modules')
```



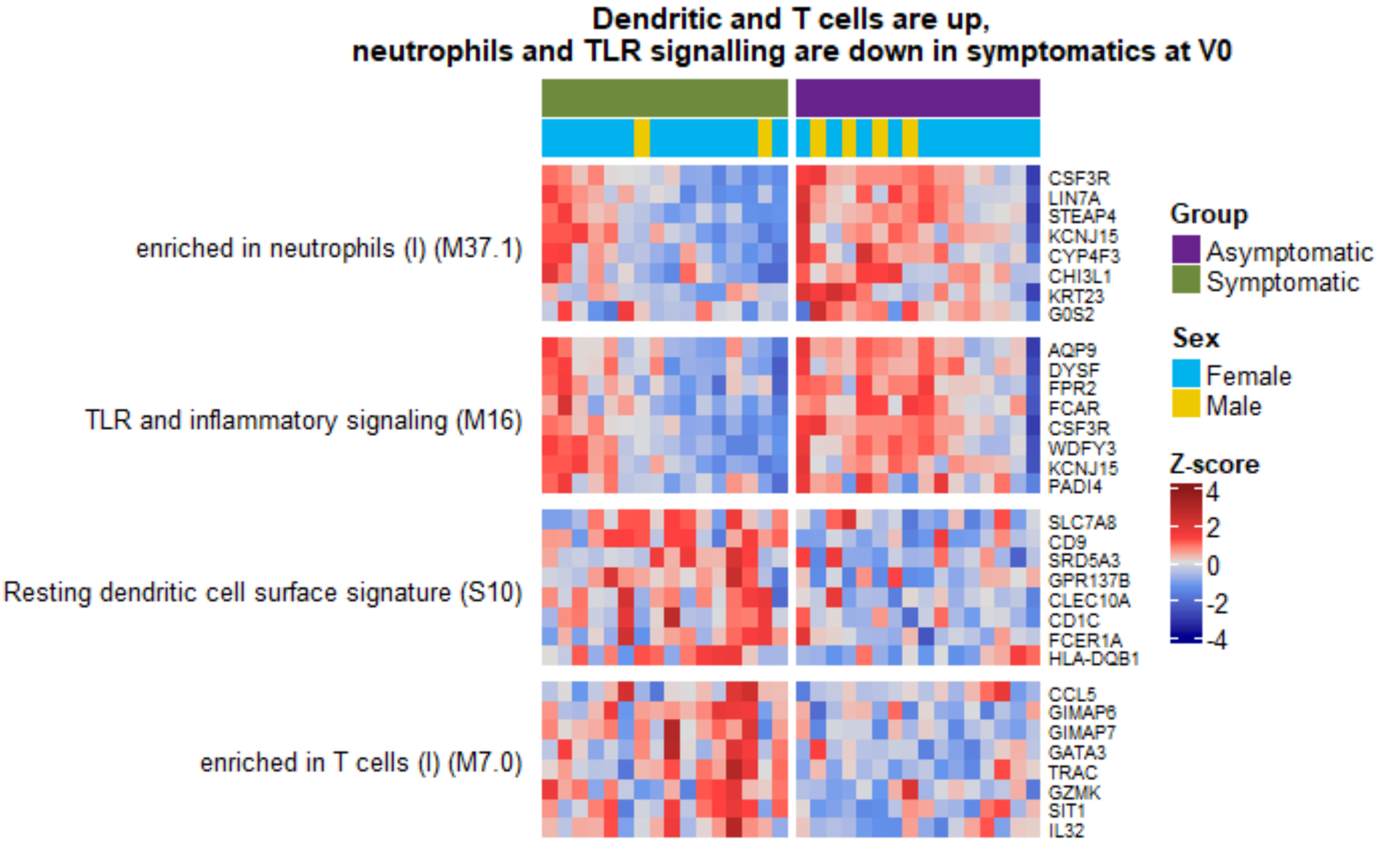
```
plotGseaTable(btm[fgseaRes[1:20, ][, pathway]], ranks, fgseaRes[1:20, ], gseaParam = 0.5, colwidths = c(6, 3, 0.
8, 1.2, 1.2), pathwayLabelStyle = list(size = 10), headerLabelStyle = list(size = 10), valueStyle = list(size = 1
0))
```



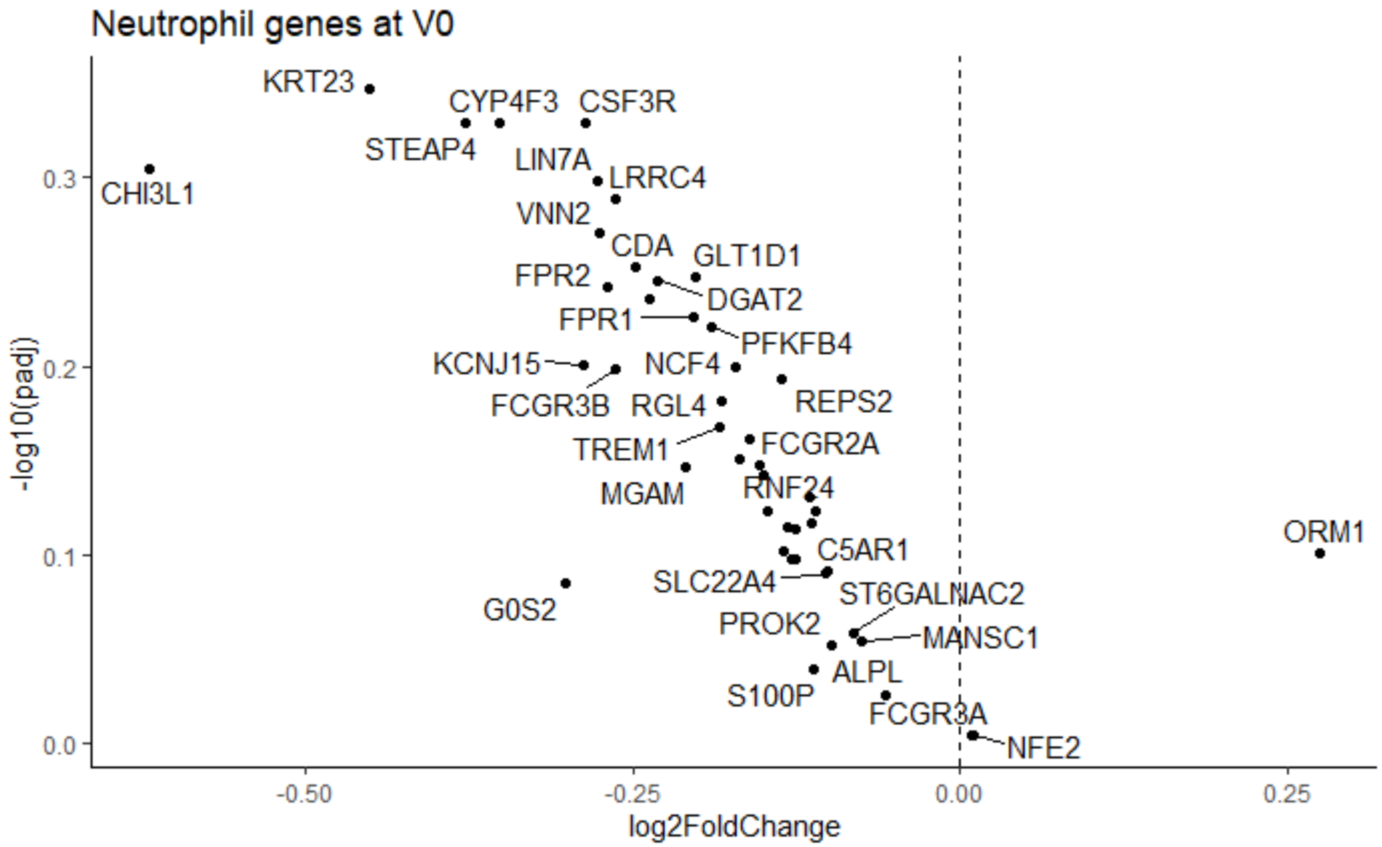
Pathway	Gene ranks	NES	pval	padj
enriched in neutrophils (I) (M37.1)		-2.74	$3.4 \cdot 10^{-8}$	$3.5 \cdot 10^{-6}$
Resting dendritic cell surface signature (S10)		2.13	$1.5 \cdot 10^{-5}$	$7.1 \cdot 10^{-4}$
enriched in T cells (I) (M7.0)		2.06	$2.1 \cdot 10^{-5}$	$7.1 \cdot 10^{-4}$
complement activation (II) (M112.1)		1.65	$3.2 \cdot 10^{-5}$	$8.1 \cdot 10^{-4}$
TLR and inflammatory signaling (M16)		-2.46	$1.2 \cdot 10^{-4}$	$2.4 \cdot 10^{-3}$
immune activation - generic cluster (M37.0)		-1.76	$1.4 \cdot 10^{-4}$	$2.4 \cdot 10^{-3}$
T cell activation (I) (M7.1)		1.86	$5.5 \cdot 10^{-4}$	$8.0 \cdot 10^{-3}$
enriched in myeloid cells and monocytes (M81)		1.91	$6.4 \cdot 10^{-4}$	$8.3 \cdot 10^{-3}$
DC surface signature (S5)		1.78	$1.6 \cdot 10^{-3}$	$1.8 \cdot 10^{-2}$
MHC-TLR7-TLR8 cluster (M146)		1.89	$1.8 \cdot 10^{-3}$	$1.9 \cdot 10^{-2}$
B cell surface signature (S2)		1.72	$2.1 \cdot 10^{-3}$	$2.0 \cdot 10^{-2}$
enriched in antigen presentation (I) (M71)		1.86	$2.9 \cdot 10^{-3}$	$2.5 \cdot 10^{-2}$
oid cell cytokines, metallopeptidases and laminins (M78)		-1.63	$4.2 \cdot 10^{-3}$	$3.3 \cdot 10^{-2}$
signaling in T cells (I) (M35.0)		1.77	$4.9 \cdot 10^{-3}$	$3.6 \cdot 10^{-2}$
interferon alpha response (I) (M158.0)		1.45	$6.5 \cdot 10^{-3}$	$4.2 \cdot 10^{-2}$
recruitment of neutrophils (M132)		-1.83	$6.1 \cdot 10^{-3}$	$4.2 \cdot 10^{-2}$
T cell differentiation (Th2) (M19)		1.75	$8.2 \cdot 10^{-3}$	$5.0 \cdot 10^{-2}$
T cell activation (II) (M7.3)		1.70	$9.0 \cdot 10^{-3}$	$5.1 \cdot 10^{-2}$
T cell surface signature (S0)		1.64	$1.3 \cdot 10^{-2}$	$7.0 \cdot 10^{-2}$
chemokine cluster (II) (M27.1)		1.56	$1.6 \cdot 10^{-2}$	$8.0 \cdot 10^{-2}$

```
res.df = res.df[order(res.df$log2FoldChange, decreasing = FALSE), ]
ha_top = HeatmapAnnotation(df = meta_data[, c('Group', 'Sex')], col = list(Group = c('Symptomatic' = 'darkolivegreen4', 'Asymptomatic' = 'darkorchid4'), Sex = c('Male' = 'gold2', 'Female' = 'deepskyblue2')), show_annotation_name = FALSE)
plot1 = vsd.df[rownames(subset(res.df, external_gene_name %in% btm[[fgseaRes[[1, 'pathway']]]] & log2FoldChange < 0))[1:8], ]
zzz = t(scale(t(plot1))); zzz = pmin(pmax(zzz, -3), 3); rownames(zzz) =  res.df[rownames(zzz), 'external_gene_name']
plot1 = Heatmap(zzz, row_title = fgseaRes[[1, 'pathway']], show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = TRUE, show_row_dend = FALSE, show_column_dend = FALSE, column_split = meta_data$Group, top_annotation = ha_top, heatmap_legend_param = list(title = "Z-score"), col = colorRamp2(c(-4, -1.25, 0, 1.25, 4), c("darkblue", "cornflowerblue", "gainsboro", "brown1", "firebrick4")), row_names_gp = gpar(fontsize = 7), row_title_rot = 0, row_title_gp = gpar(fontsize = 10), column_title_gp = gpar(fontsize = 11, fontface = 2), column_title = 'Dendritic and T cells are up,\nneutrophils and TLR signalling are down in symptomatics at V0')
plot2 = vsd.df[rownames(subset(res.df, external_gene_name %in% btm[[fgseaRes[[5, 'pathway']]]] & log2FoldChange < 0))[1:8], ]
zzz = t(scale(t(plot2))); zzz = pmin(pmax(zzz, -3), 3); rownames(zzz) =  res.df[rownames(zzz), 'external_gene_name']
plot2 = Heatmap(zzz, show_heatmap_legend = FALSE, row_title = fgseaRes[[5, 'pathway']], show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = TRUE, show_row_dend = FALSE, show_column_dend = FALSE, column_split = meta_data$Group, heatmap_legend_param = list(title = "Z-score"), col = colorRamp2(c(-4, -1.25, 0, 1.25, 4), c("darkblue", "cornflowerblue", "gainsboro", "brown1", "firebrick4")), row_names_gp = gpar(fontsize = 7), row_title_rot = 0, row_title_gp = gpar(fontsize = 10), column_title_gp = gpar(fontsize = 11, fontface = 2))
res.df = res.df[order(res.df$log2FoldChange, decreasing = TRUE), ]
plot3 = vsd.df[rownames(subset(res.df, external_gene_name %in% btm[[fgseaRes[[2, 'pathway']]]] & log2FoldChange > 0))[1:8], ]
zzz = t(scale(t(plot3))); zzz = pmin(pmax(zzz, -3), 3); rownames(zzz) =  res.df[rownames(zzz), 'external_gene_name']
plot3 = Heatmap(zzz, show_heatmap_legend = FALSE, row_title = fgseaRes[[2, 'pathway']], show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = TRUE, show_row_dend = FALSE, show_column_dend = FALSE, column_split = meta_data$Group, heatmap_legend_param = list(title = "Z-score"), col = colorRamp2(c(-4, -1.25, 0, 1.25, 4), c("darkblue", "cornflowerblue", "gainsboro", "brown1", "firebrick4")), row_names_gp = gpar(fontsize = 7), row_title_rot = 0, row_title_gp = gpar(fontsize = 10), column_title_gp = gpar(fontsize = 11, fontface = 2))
plot4 = vsd.df[rownames(subset(res.df, external_gene_name %in% btm[[fgseaRes[[3, 'pathway']]]] & log2FoldChange > 0))[1:8], ]
```

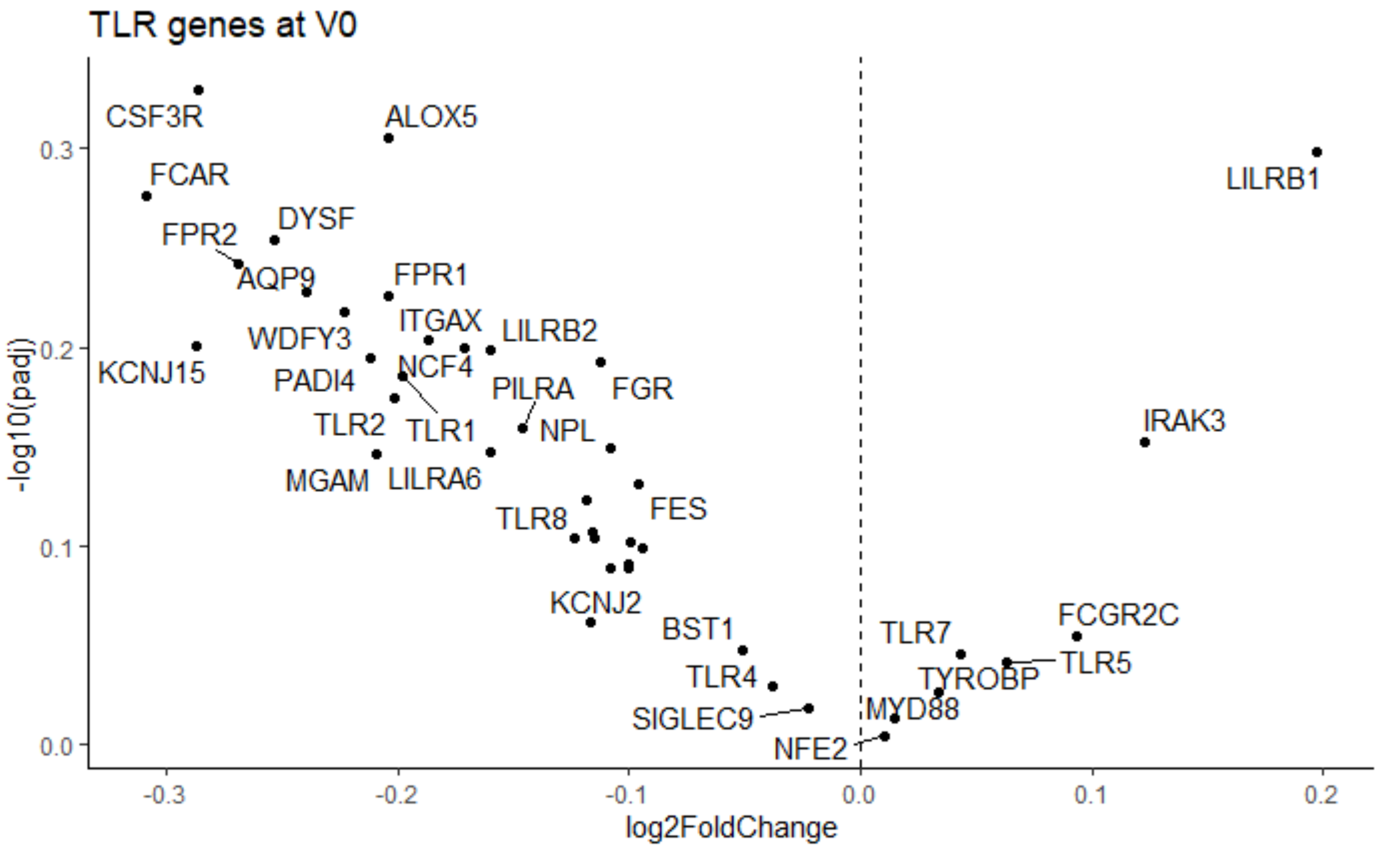
```
zzz = t(scale(t(plot4))); zzz = pmin(pmax(zzz, -3), 3); rownames(zzz) =  res.df[rownames(zzz), 'external_gene_name']
plot4 = Heatmap(zzz, show_heatmap_legend = FALSE, row_title = fgseaRes[[3, 'pathway']], show_row_names = TRUE, show_column_names = FALSE, cluster_rows = TRUE, cluster_columns = TRUE, show_row_dend = FALSE, show_column_dend = FALSE, column_split = meta_data$Group, heatmap_legend_param = list(title = "Z-score"), col = colorRamp2(c(-4, -1.25, 0, 1.25, 4), c("darkblue", "cornflowerblue", "gainsboro", "brown1", "firebrick4")), row_names_gp = gpar(fontsize = 7), row_title_rot = 0, row_title_gp = gpar(fontsize = 10), column_title_gp = gpar(fontsize = 11, fontface = 2))
draw(plot1 %v% plot2 %v% plot3 %v% plot4); rm(plot1); rm(plot2); rm(plot3); rm(plot4); rm(ha_top); rm(zzz); rm(ranks)
```



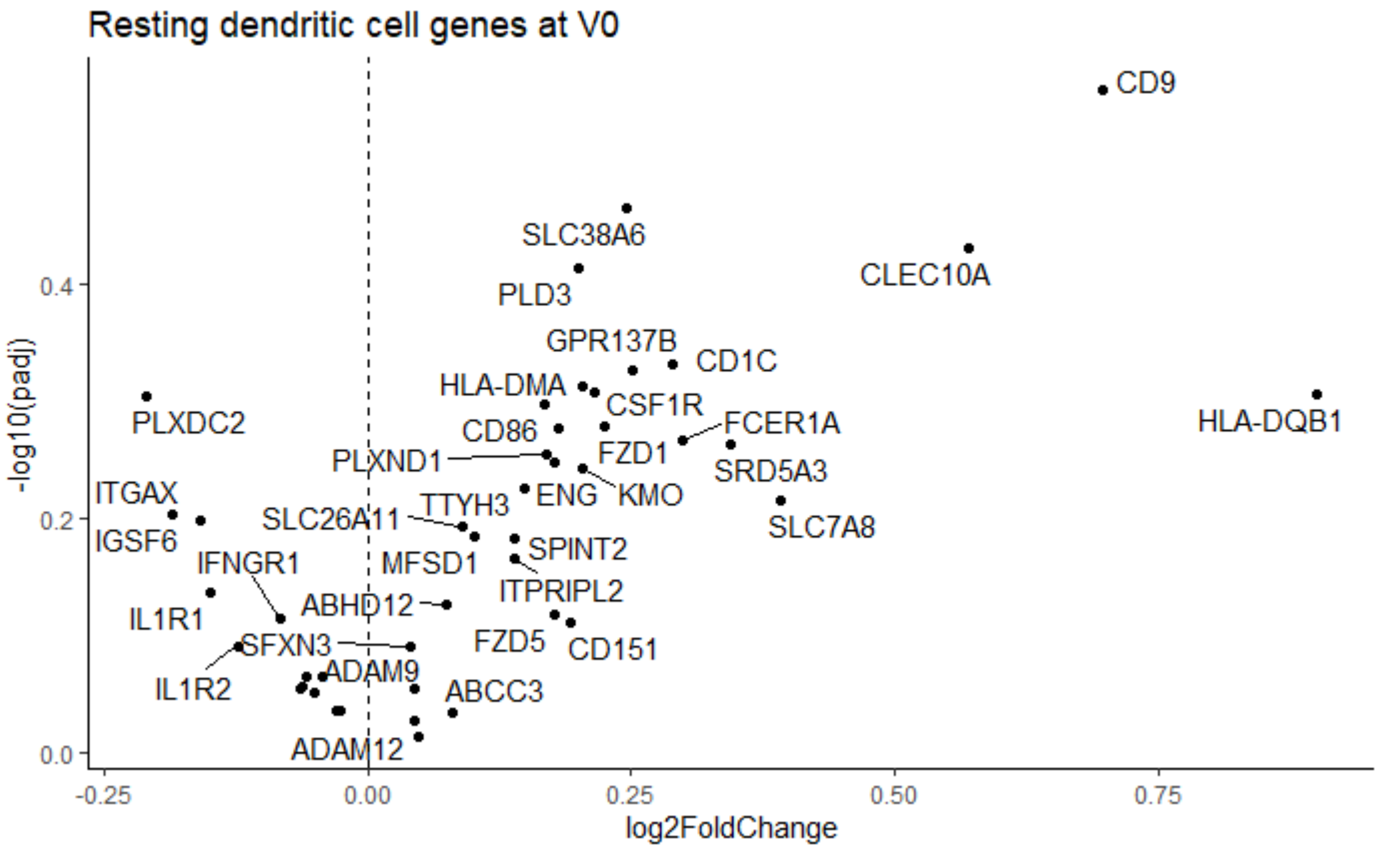
```
ggplot(subset(res.df, external_gene_name %in% btm[[fgseaRes[[1, 'pathway']]]]), aes(x = log2FoldChange, y = -log10(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + ggtitle('Neutrophil genes at V0') + geom_vline(xintercept = 0, linetype = 2)
```



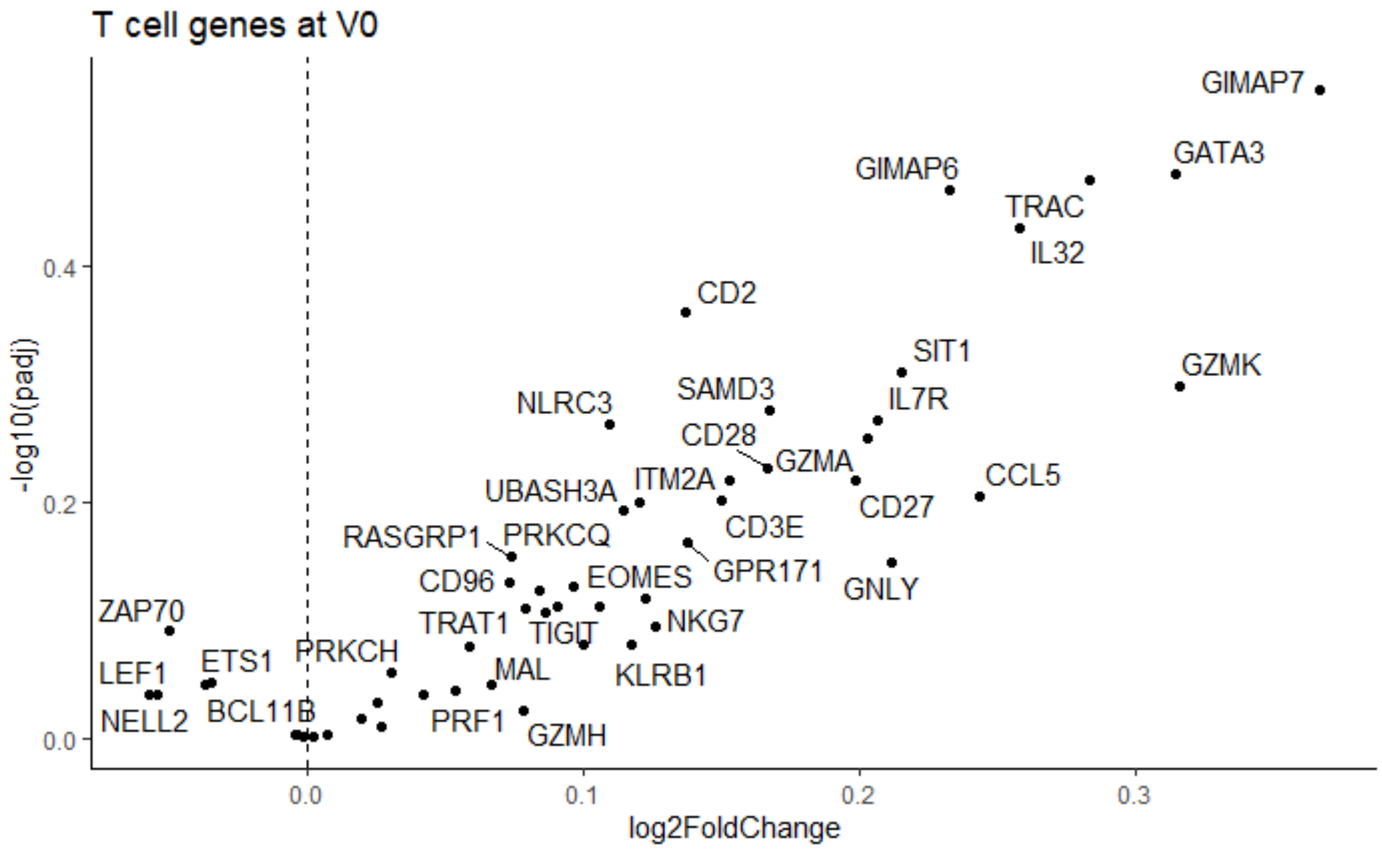
```
ggplot(subset(res.df, external_gene_name %in% btm[[fgseaRes[[5, 'pathway']]]]), aes(x = log2FoldChange, y = -log10(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + ggtitle('TLR genes at V0') + geom_vline(xintercept = 0, linetype = 2)
```



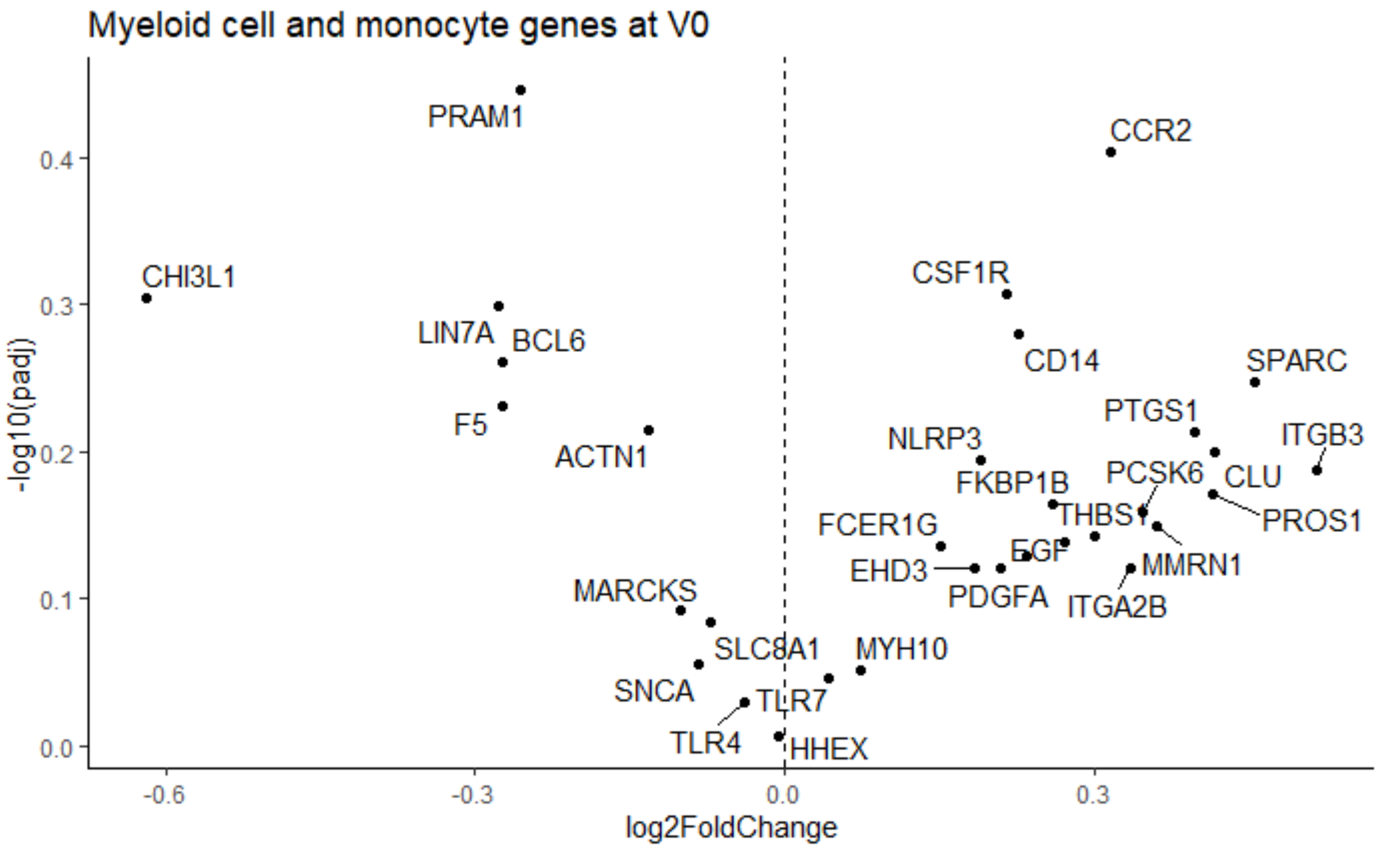
```
ggplot(subset(res.df, external_gene_name %in% btm[[fgseaRes[[2, 'pathway']]]]), aes(x = log2FoldChange, y = -log1
0(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + ggtitle('Resting den
dritic cell genes at V0') + geom_vline(xintercept = 0, linetype = 2)
```



```
ggplot(subset(res.df, external_gene_name %in% btm[[fgseaRes[[3, 'pathway']]]]), aes(x = log2FoldChange, y = -log1
0(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + ggtitle('T cell gene
s at V0') + geom_vline(xintercept = 0, linetype = 2)
```



```
ggplot(subset(res.df, external_gene_name %in% btm[[fgseaRes[[8, 'pathway']]]]), aes(x = log2FoldChange, y = -log1
0(padj), label = external_gene_name)) + geom_point() + geom_text_repel() + theme_classic() + ggtitle('Myeloid cel
l and monocyte genes at V0') + geom_vline(xintercept = 0, linetype = 2)
```


Myeloid cell and monocyte genes at V0
-log10(padj)
0.0
0.1
0.2
0.3
0.4
log2FoldChange
-0.6
-0.3
0.0
0.3
PRAM1
CCR2
CHI3L1
CSF1R
LIN7A
BCL6
CD14
SPARC
F5
ACTN1
PTGS1
ITGB3
NLRP3
FKBP1B
PCSK6
CLU
FCER1G
THBS1
PROS1
EGF
EHD3
MMRN1
PDGFA
ITGA2B
MARCKS
SLC8A1
MYH10
SNCA
TLR7
TLR4
HHEX