

# SRP136558 Data Analysis

First Step is to read in data

```
# DESeq2 readcounts
#-----
norm <- DESeqDataSetFromMatrix(countData <- readcounts_recount, colData <- sampleinfo_recount[, c(1, 3)], design = ~condition, tidy = TRUE)
```

Apply DESeq

Results

```
summary(res)
```

```
##
## out of 37269 with nonzero total read count
## adjusted p-value < 0.1
## LFC > 0 (up)      : 6872, 18%
## LFC < 0 (down)    : 6083, 16%
## outliers [1]       : 5391, 14%
## low counts [2]     : 1683, 4.5%
## (mean count < 8)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

Order by Padj scores

```
res <- res[order(res$padj),]
head(res)
```

```
## log2 fold change (MLE): condition Cold_M1 vs Cold_M2
## Wald test p-value: condition Cold_M1 vs Cold_M2
## DataFrame with 6 rows and 6 columns
##           baseMean log2FoldChange      lfcSE      stat      pvalue
##           <numeric>      <numeric>      <numeric>      <numeric>      <numeric>
## ENSMUSG00000026938.10 1006476.8      -6.66241 0.1680300  -39.6501 0.000000e+00
## ENSMUSG00000072889.9  117439.6      -2.39063 0.0613879  -38.9431 0.000000e+00
## ENSMUSG00000052403.11 81844.9      -6.60011 0.1777407  -37.1334 8.13249e-302
## ENSMUSG00000079164.8   78556.9      -5.10529 0.1486262  -34.3498 1.41592e-258
## ENSMUSG00000031129.9  287666.2      -2.63590 0.0772382  -34.1269 2.93977e-255
## ENSMUSG00000030103.11 396602.6      3.60335 0.1073810   33.5567 7.18107e-247
##           padj
##           <numeric>
## ENSMUSG00000026938.10 0.000000e+00
## ENSMUSG00000072889.9  0.000000e+00
```

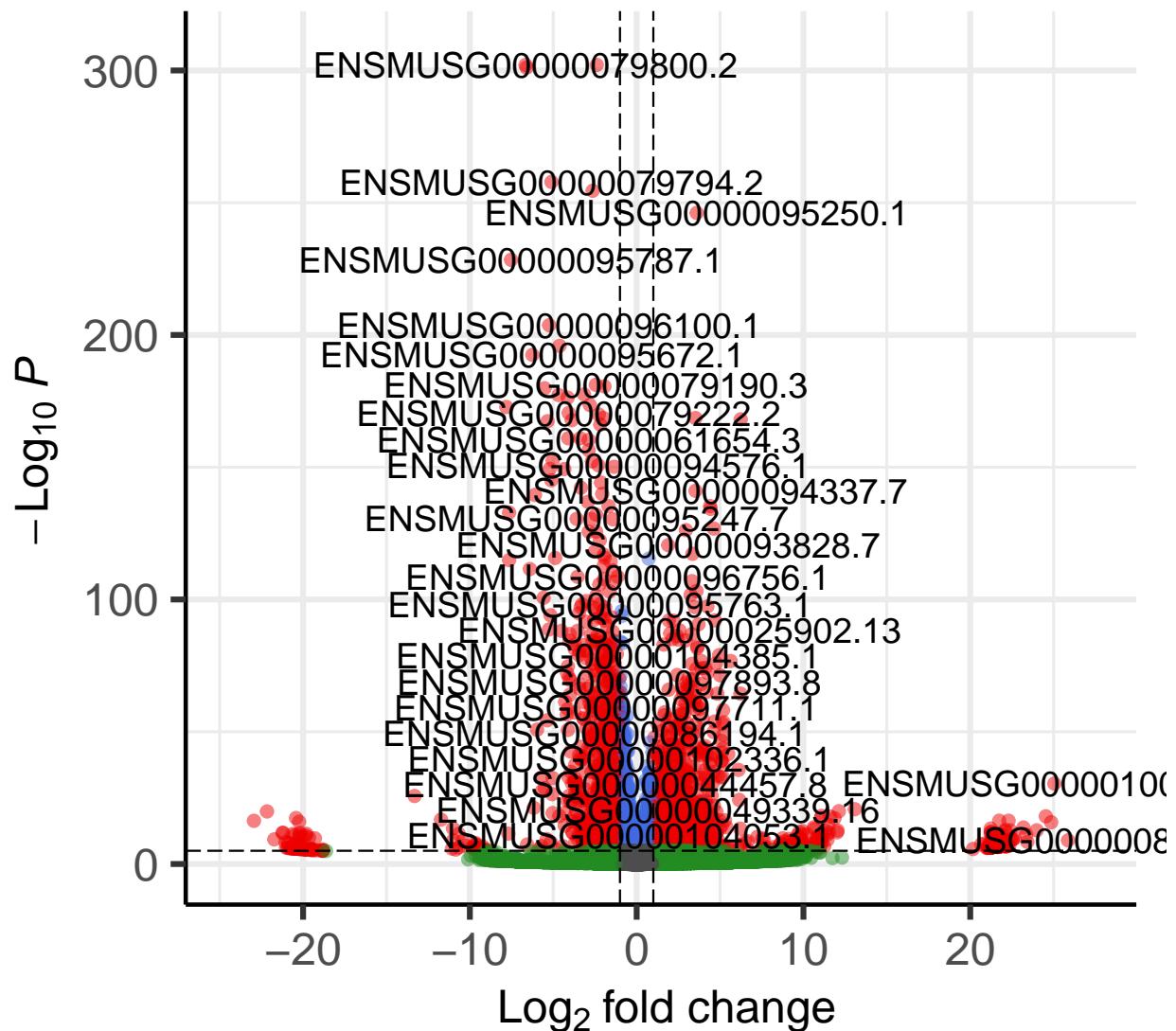
```
## ENSMUSG00000052403.11 8.18535e-298
## ENSMUSG00000079164.8  1.06885e-254
## ENSMUSG00000031129.9  1.77533e-251
## ENSMUSG00000030103.11 3.61387e-243

EnhancedVolcano(res, lab=rownames(m1res), x = "log2FoldChange", y = "pvalue" )
```

# Volcano plot

EnhancedVolcano

● NS ● Log<sub>2</sub> FC ● p-value ● p-value and log<sub>2</sub> FC



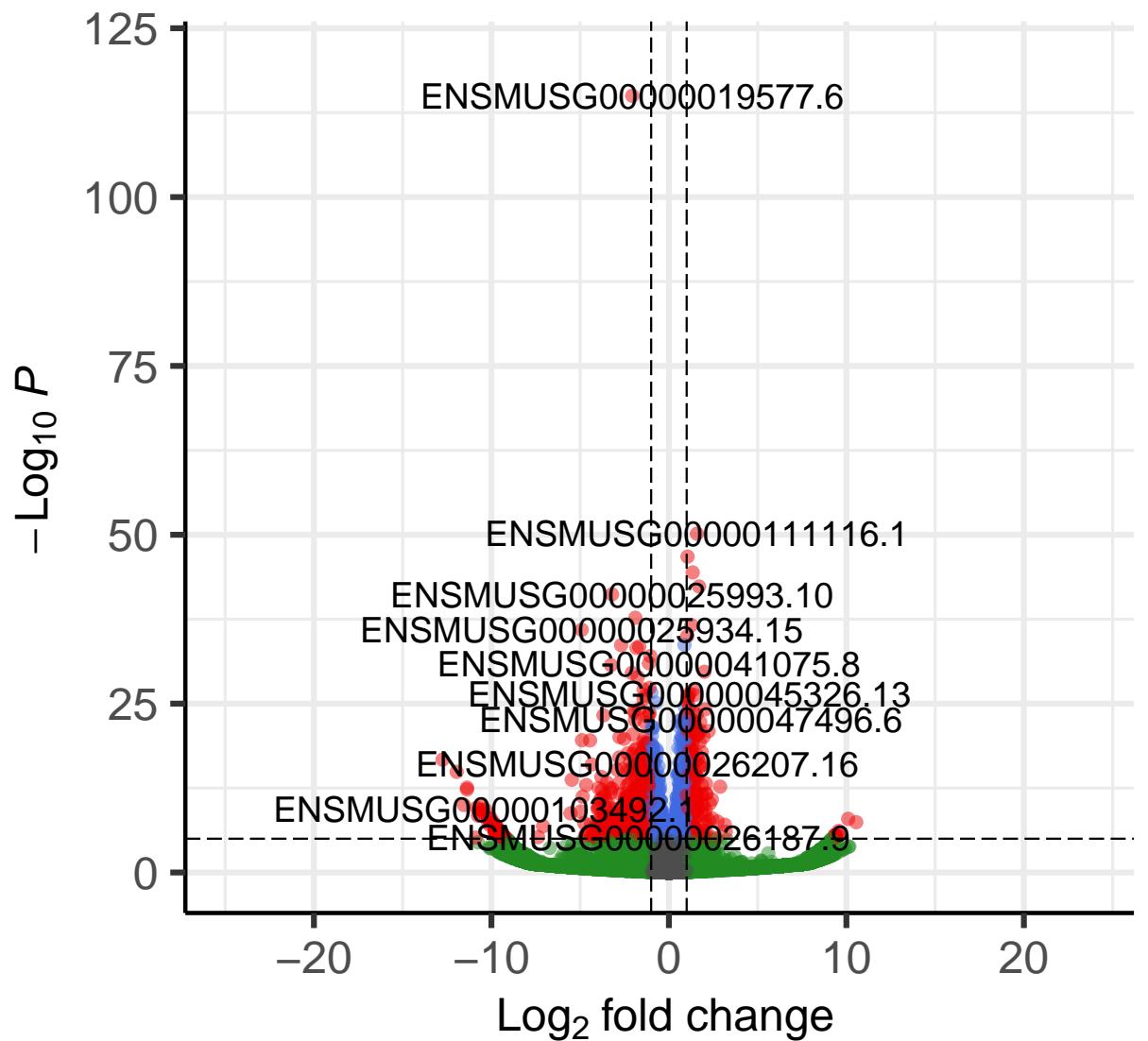
#M1

```
EnhancedVolcano(m1res, lab=rownames(m1res), x = "log2FoldChange", y = "pvalue" )
```

# Volcano plot

*EnhancedVolcano*

● NS ● Log<sub>2</sub> FC ● p-value ● p-value and log<sub>2</sub> FC

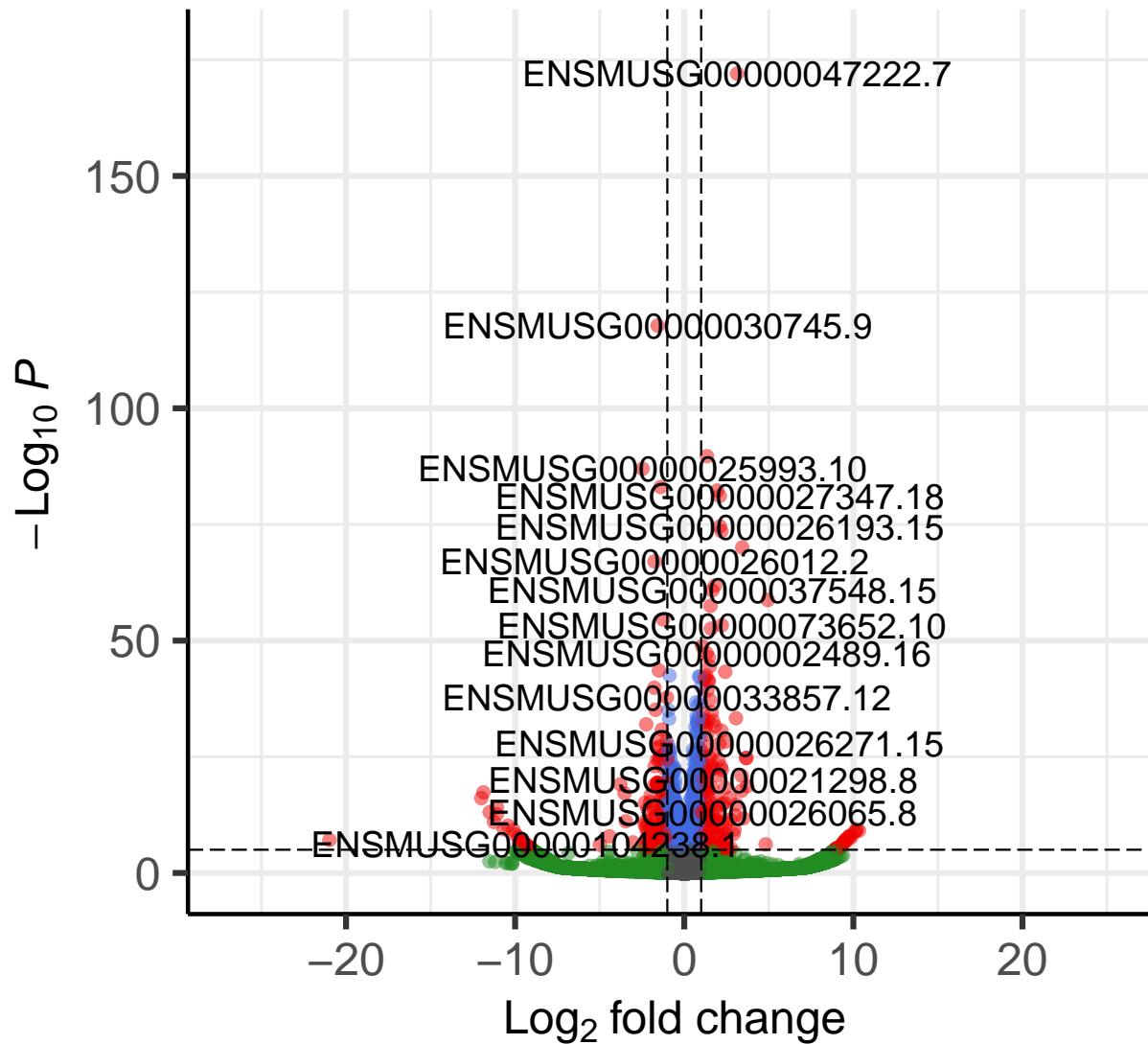


```
#M2  
EnhancedVolcano(m2res, lab=rownames(m1res), x = "log2FoldChange", y = "pvalue" )
```

# Volcano plot

*EnhancedVolcano*

● NS ● Log<sub>2</sub> FC ● p-value ● p-value and log<sub>2</sub> FC



```
## DataFrame with 12 rows and 5 columns
##           external_id condition sizeFactor      cellType      factor
##           <character>  <factor>   <numeric> <character> <character>
## SRR6904747  SRR6904747    Warm_M1    1.113259        M1      Warm
## SRR6904748  SRR6904748    Warm_M2    0.968979        M2      Warm
```

```

## SRR6904749 SRR6904749 Warm_M1 1.250047 M1 Warm
## SRR6904750 SRR6904750 Warm_M2 0.925725 M2 Warm
## SRR6904751 SRR6904751 Warm_M1 1.043344 M1 Warm
## ... ...
## SRR6904754 SRR6904754 Cold_M2 1.155105 M2 Cold
## SRR6904755 SRR6904755 Cold_M1 1.140619 M1 Cold
## SRR6904756 SRR6904756 Cold_M2 0.705684 M2 Cold
## SRR6904757 SRR6904757 Cold_M1 1.254974 M1 Cold
## SRR6904758 SRR6904758 Cold_M2 0.726120 M2 Cold

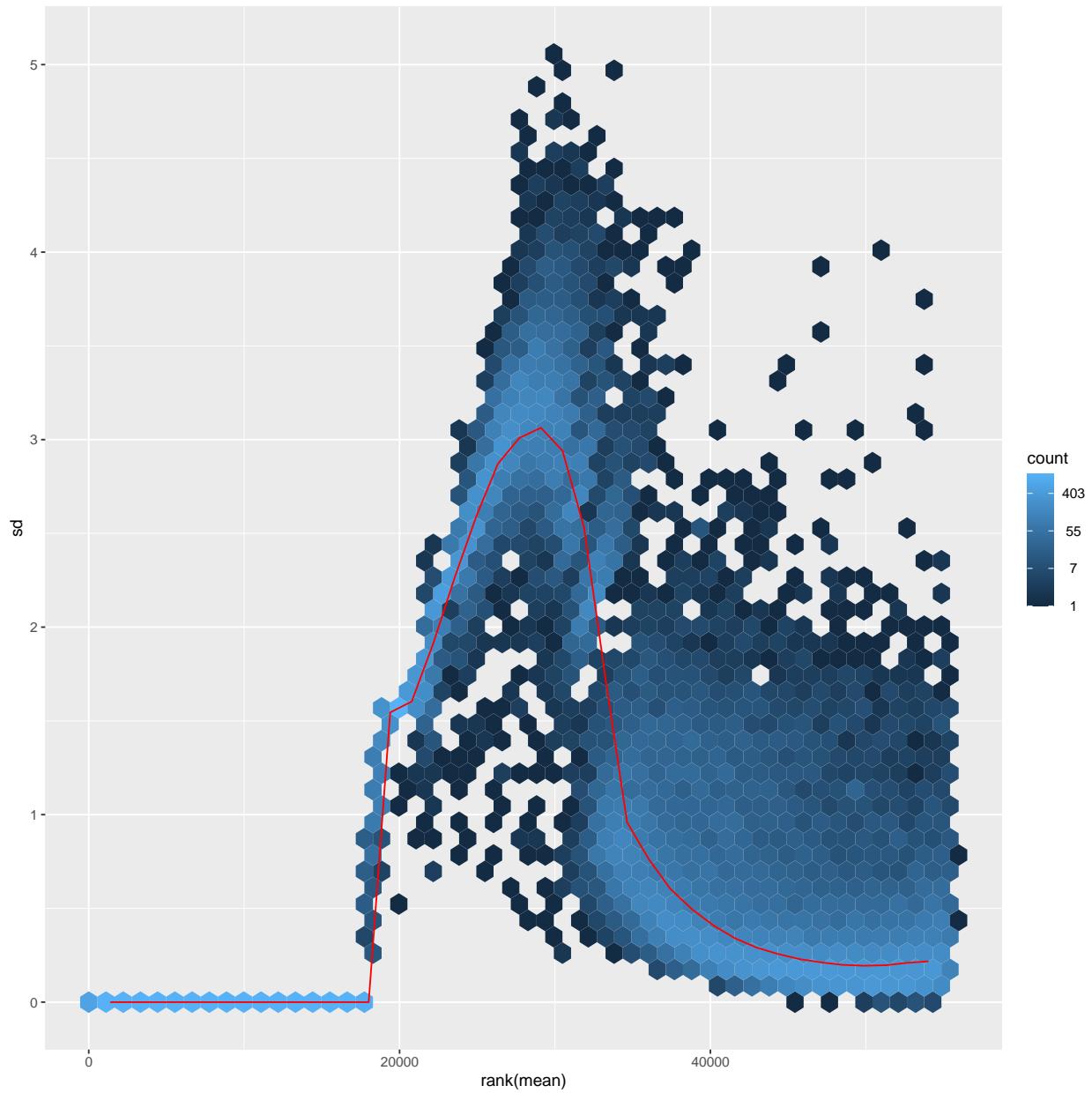
```

## Variance Plots

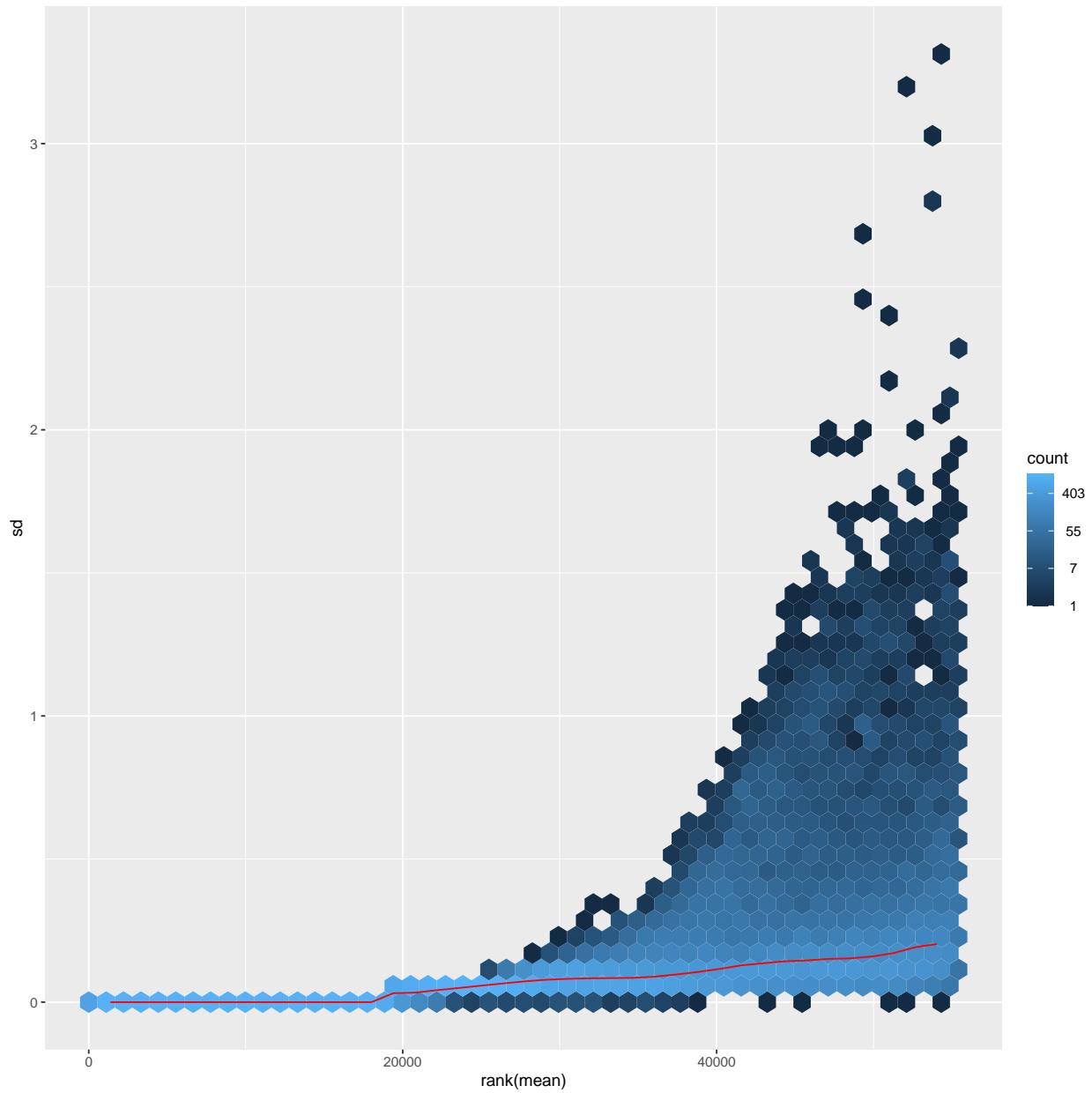
```

#Mean/SD Normal transform
#-----
meanSdPlot(assay(ntd))

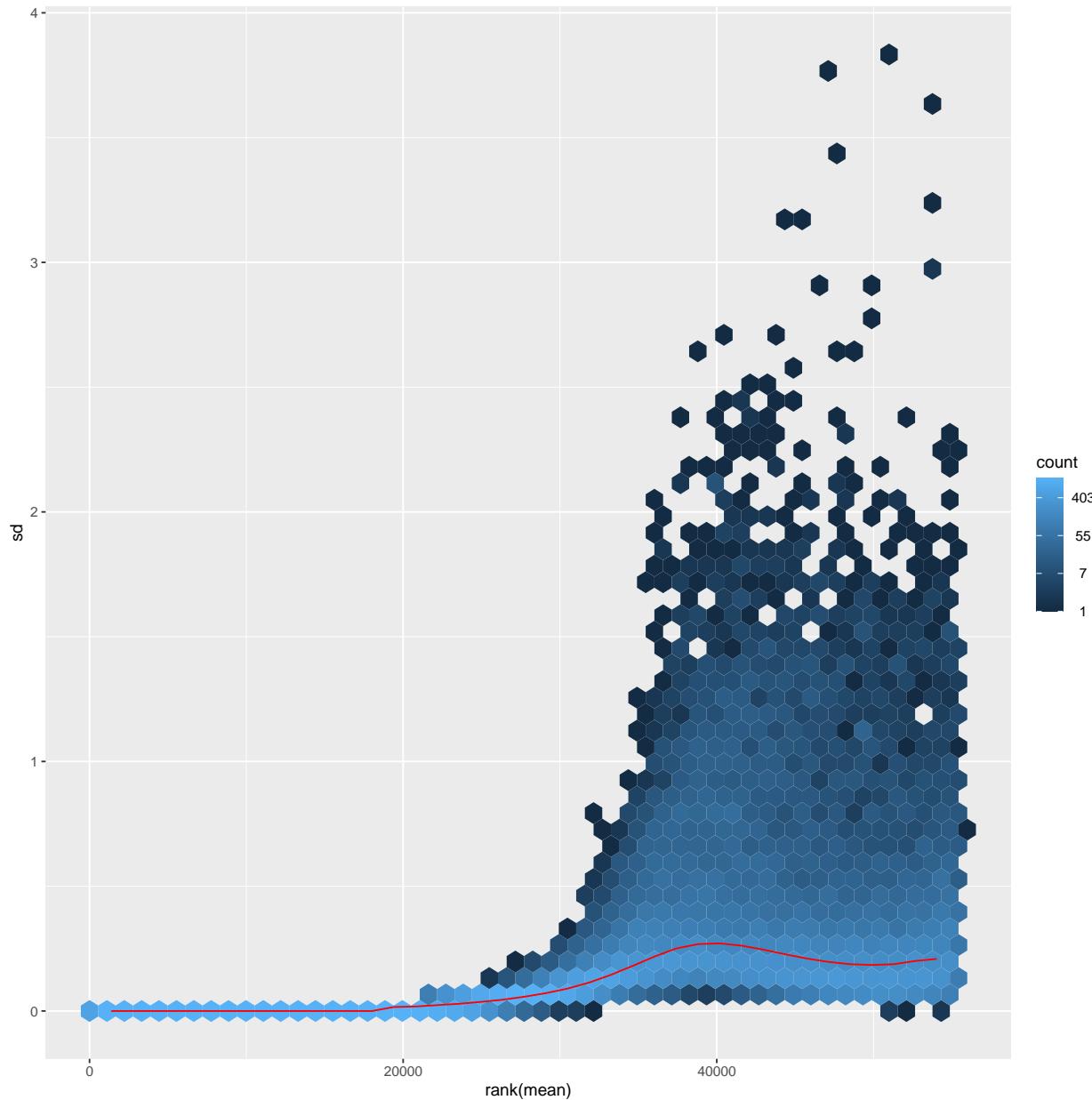
```



```
#Mean/SD Variance stabilizing transformation  
#-----  
meanSdPlot(assay(vsd))
```



```
#Mean/SD Regularized log transform  
#-----  
meanSdPlot(assay(rld))
```



## Heatmaps

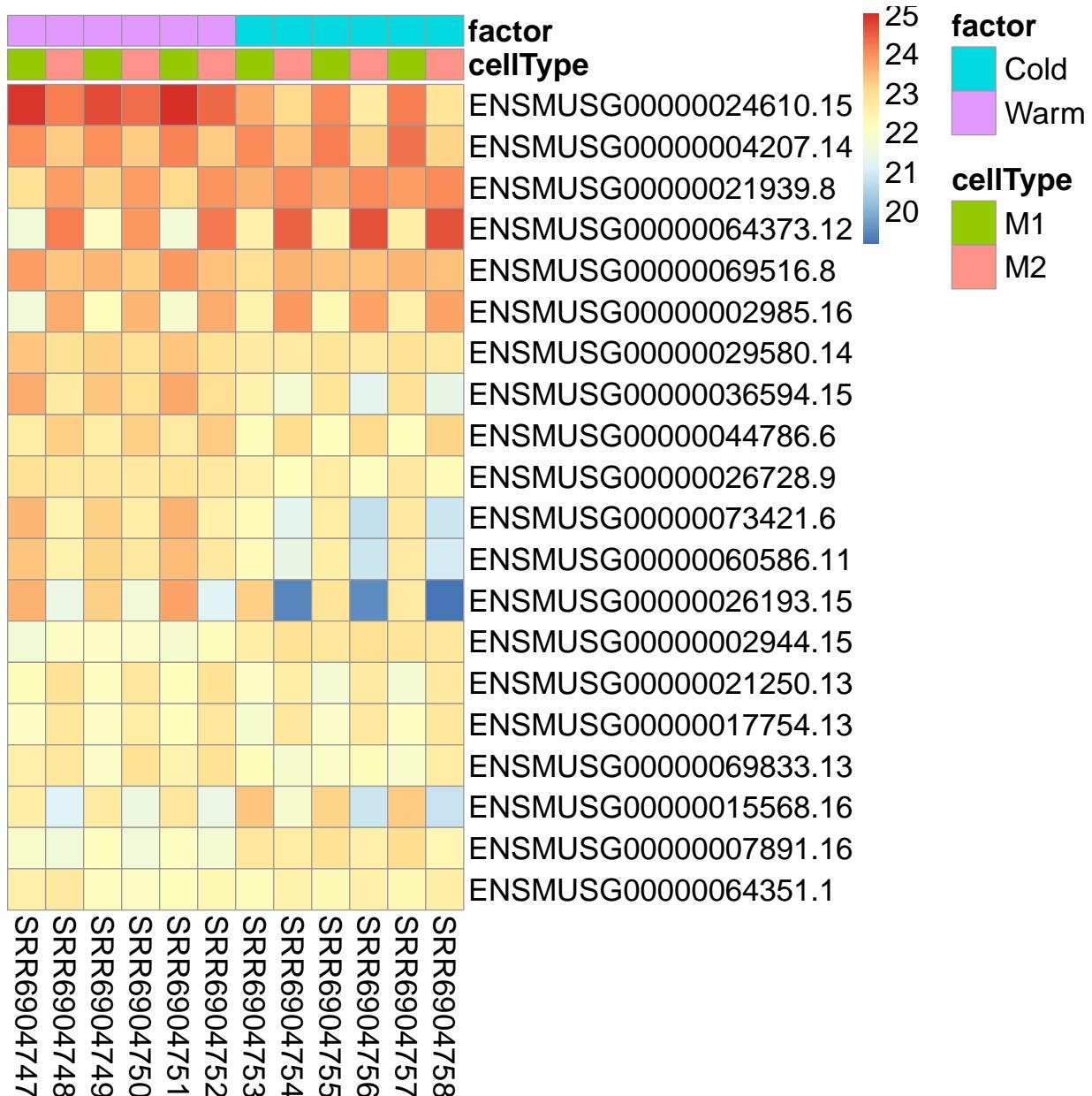
```
# Heatmaps
#-----
# NORM
select <- order(rowMeans(counts(reg, normalized = TRUE)), decreasing = TRUE)[1:20]
df <- as.data.frame(colData(reg)[, c("cellType", "factor")])
pheatmap(assay(ntd)[select, ], cluster_rows = FALSE, show_rownames = TRUE, cluster_cols = FALSE,
         annotation_col = df, fontsize = 20)

## Found more than one class "unit" in cache; using the first, from namespace 'ggbio'

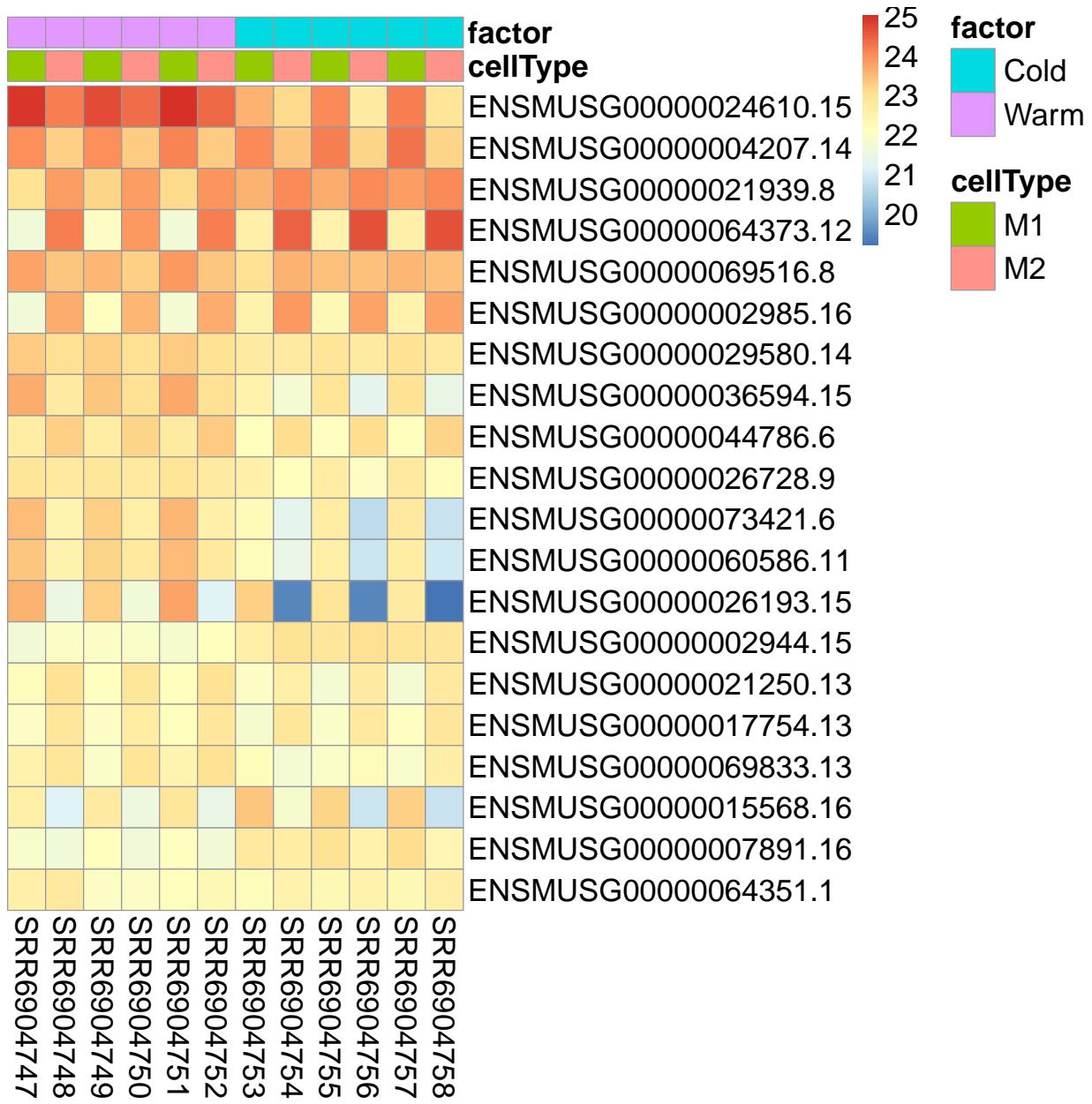
## Also defined by 'hexbin'
```

```
## Found more than one class "simpleUnit" in cache; using the first, from namespace 'ggbio'

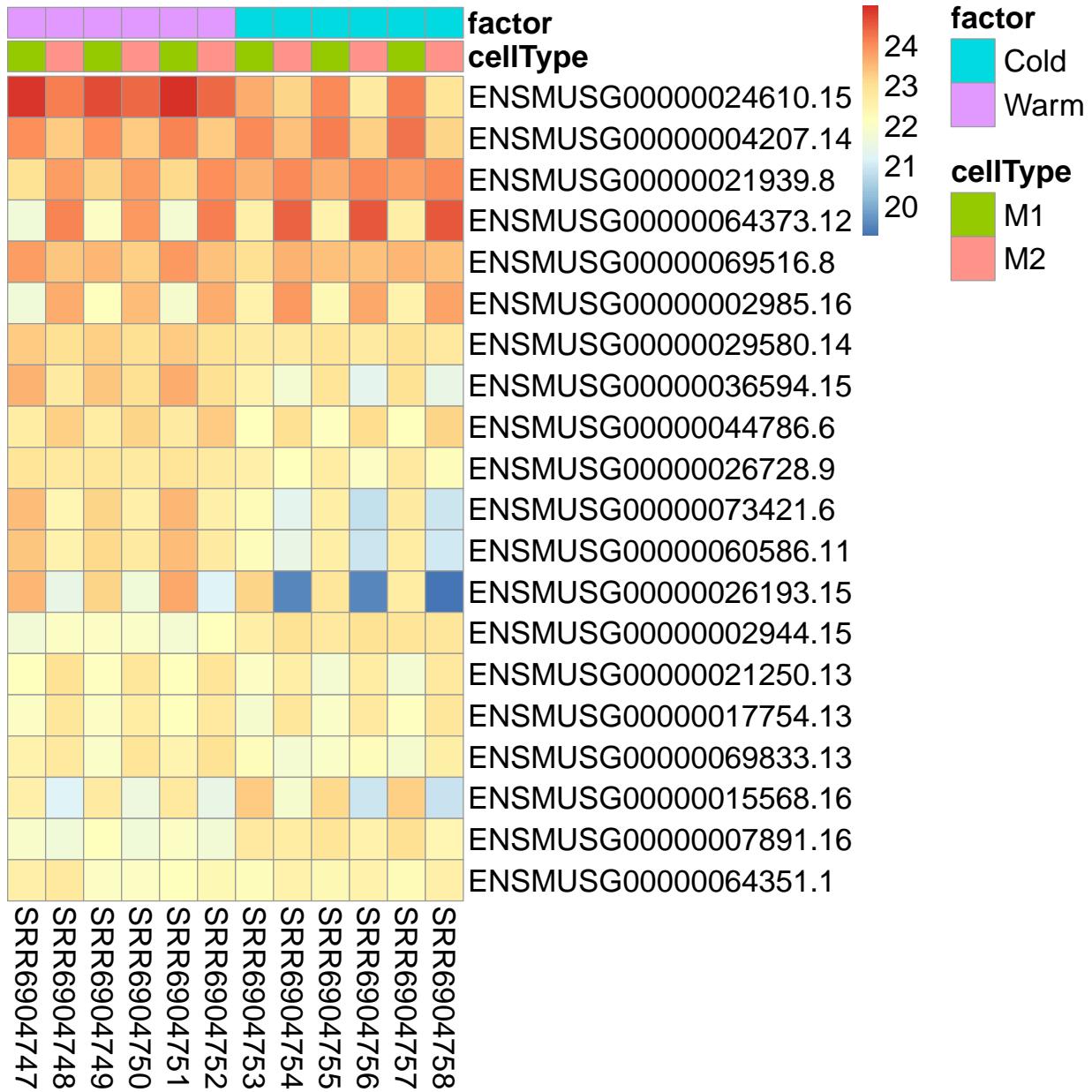
## Also defined by 'hexbin'
```



```
# VSD
select <- order(rowMeans(counts(reg, normalized = TRUE)), decreasing = TRUE)[1:20]
df <- as.data.frame(colData(reg)[, c("cellType", "factor")])
pheatmap(assay(vsd)[select, ], cluster_rows = FALSE, show_rownames = TRUE, cluster_cols = FALSE,
         annotation_col = df, fontsize = 20)
```

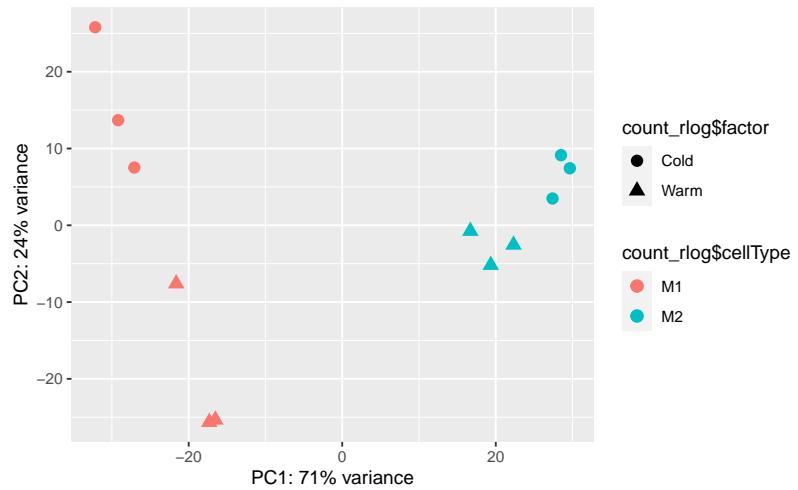


```
# RLD
select <- order(rowMeans(counts(reg, normalized = TRUE)), decreasing = TRUE)[1:20]
df <- as.data.frame(colData(reg)[, c("cellType", "factor")])
pheatmap(assay(rld)[select, ], cluster_rows = FALSE, show_rownames = TRUE, cluster_cols = FALSE,
         annotation_col = df, fontsize = 20)
```

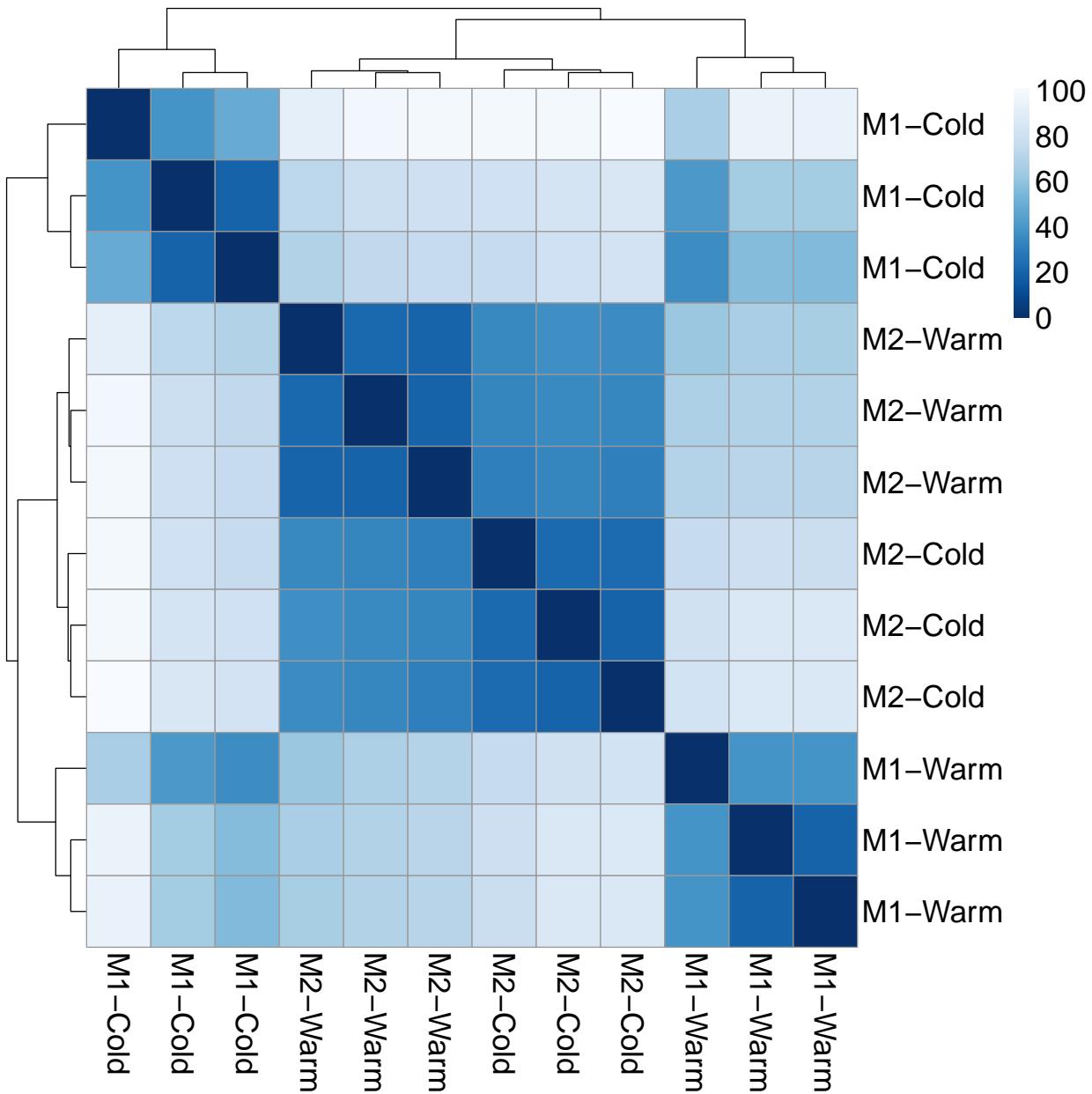


PCA

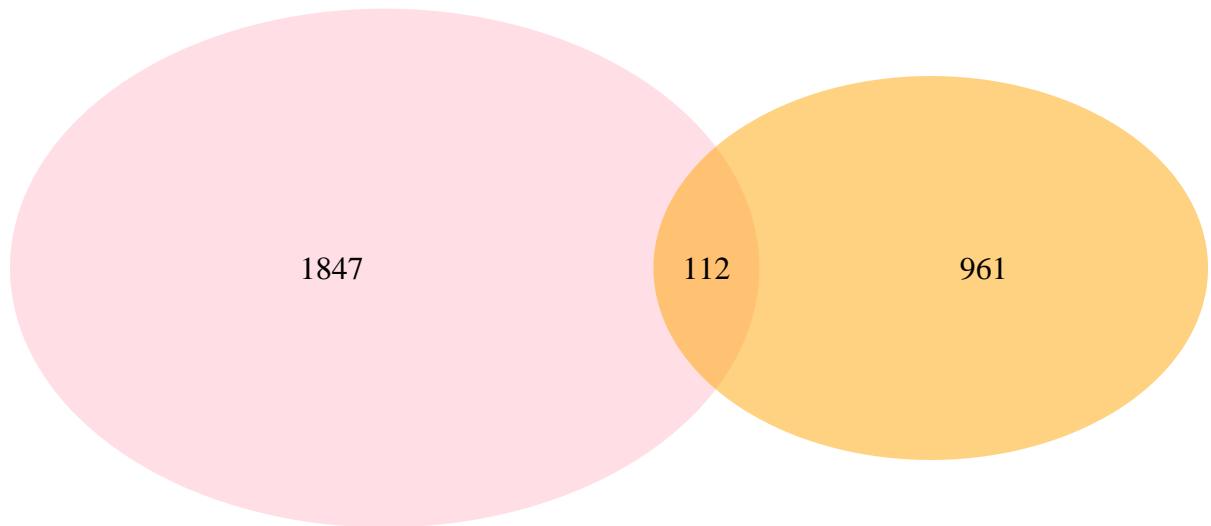
```
ggplot(pcaplot, aes(PC1, PC2, color=count_rlog$cellType, shape=count_rlog$factor)) +
  geom_point(size=3) +
  xlab(paste0("PC1: ", percentVar[1], "% variance")) +
  ylab(paste0("PC2: ", percentVar[2], "% variance")) +
  coord_fixed()
```



Sample distances



```
draw.pairwise.venn(area1 = nrow(da_best_M1), area2 = nrow(da_best_M2), cross.area = nrow(intersection),
```

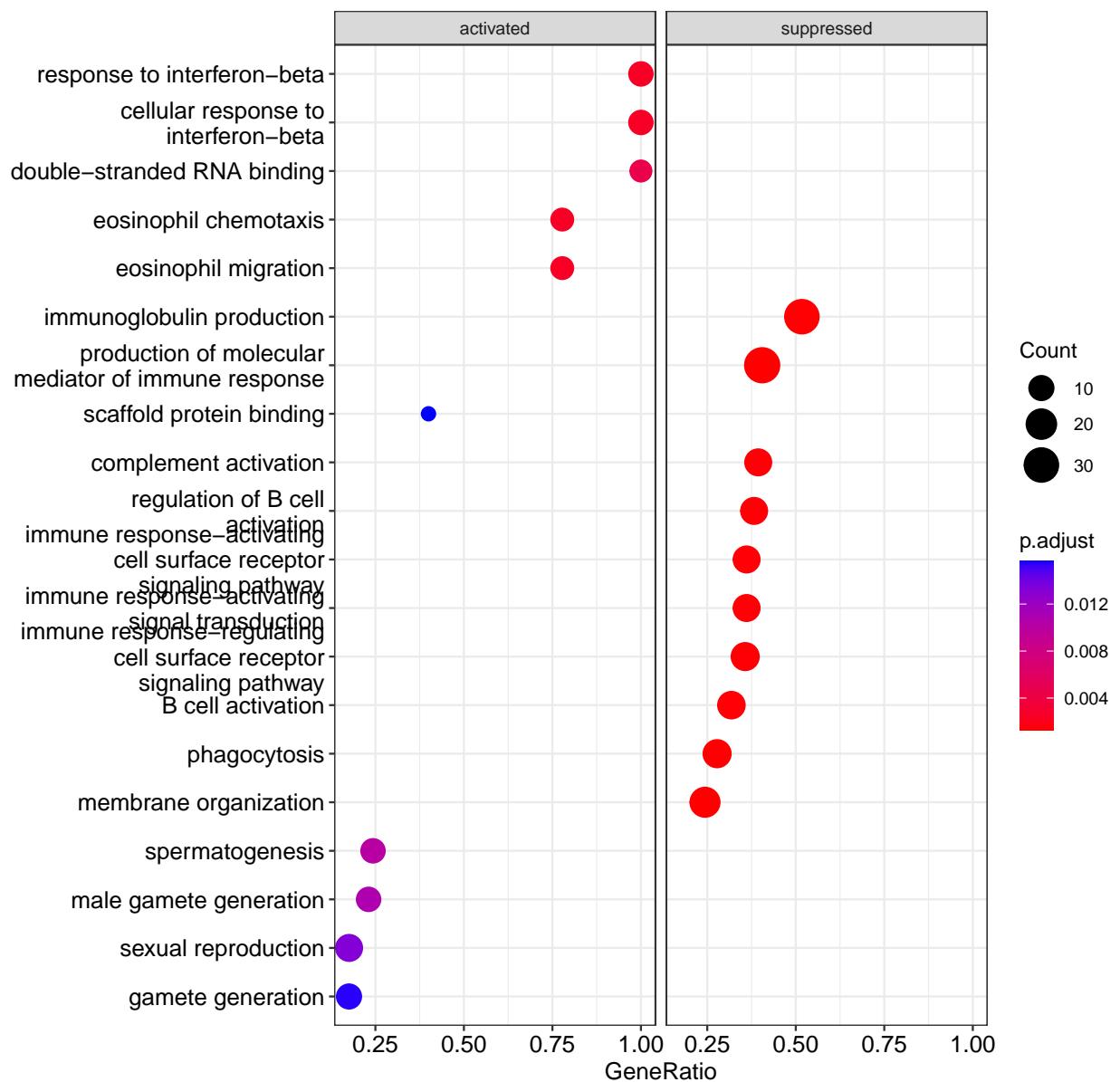


```
## (polygon[GRID.polygon.546], polygon[GRID.polygon.547], polygon[GRID.polygon.548], polygon[GRID.polygon.549])
```

### Cluster Profiler

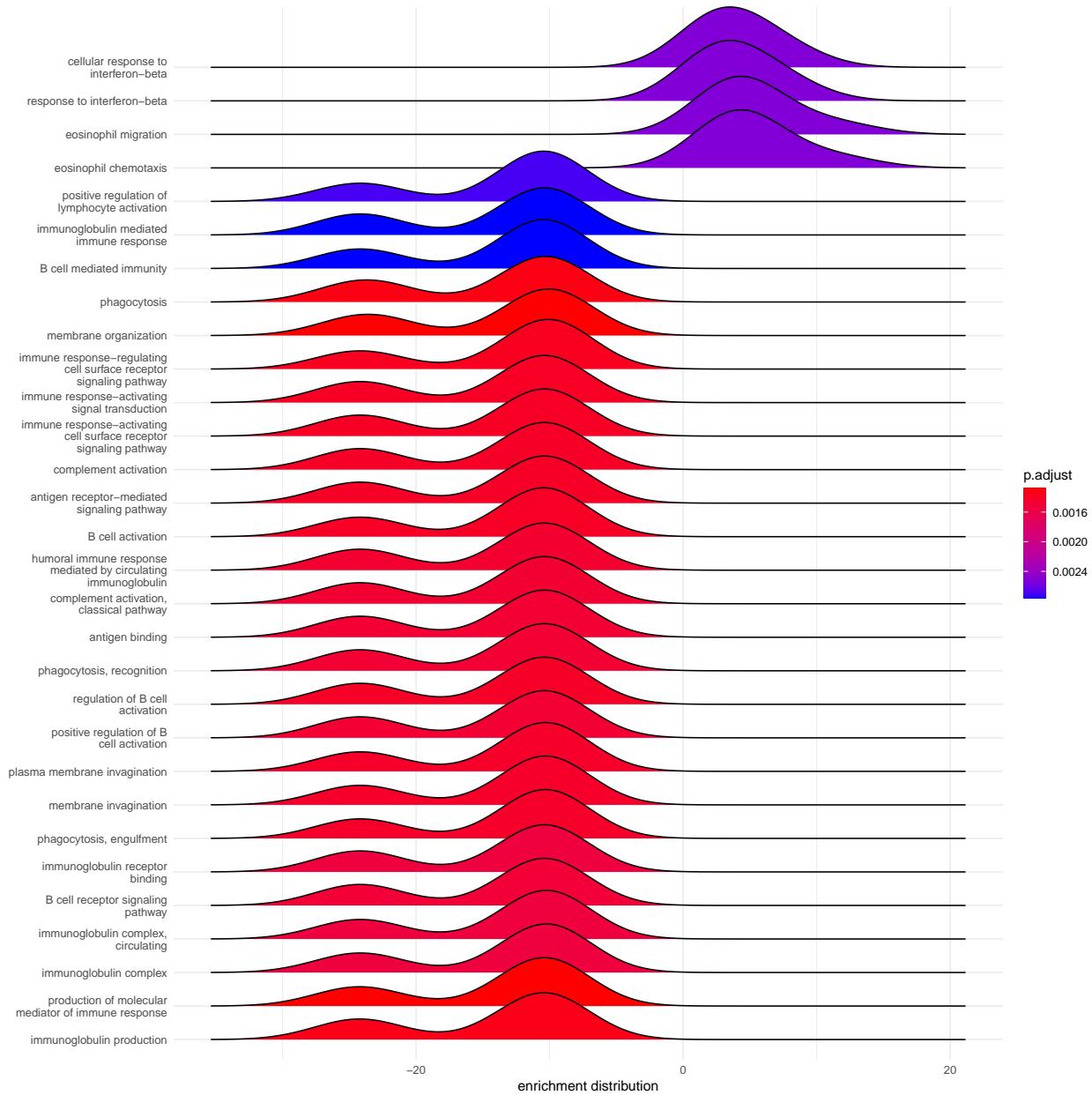
```
gse <- gseGO(geneList=gene_list,
               ont = "ALL",
               keyType = "ENSEMBL",
               nPerm = 1000,
               minGSSize = 5,
               maxGSSize = 100,
               pvalueCutoff = 0.05,
               verbose = TRUE,
               OrgDb = "org.Mm.eg.db",
               pAdjustMethod = "none")

dotplot(gse, showCategory=10, split=".sign") + facet_grid(.~.sign)
```



```
ridgeplot(gse) + labs(x = "enrichment distribution") + theme_minimal()
```

```
## Picking joint bandwidth of 3.17
```



```
gseaplot(gse, by = "all", title = gse$Description[1], geneSetID = 1)
```

## Production of molecular mediator of immune resp

