# GO-Analysis

# Enrichment Analysis

Find pathways that are overrepresented in dataset

## Libraries

For this, we will be using `clusterProfiler`

```
#BiocManager::install(c("clusterProfiler", "org.Mm.eg.db", "AnnotationDbi", "BiocGenerics"))
#BiocManager::install("TxDb.Mmusculus.UCSC.mm10.ensGene")

library(TxDb.Mmusculus.UCSC.mm10.ensGene)
```

```
Loading required package: GenomicFeatures

Loading required package: BiocGenerics

Loading required package: generics


Attaching package: 'generics'

The following objects are masked from 'package:base':

    as.difftime, as.factor, as.ordered, intersect, is.element, setdiff,
    setequal, union


Attaching package: 'BiocGenerics'

The following objects are masked from 'package:stats':

    IQR, mad, sd, var, xtabs

The following objects are masked from 'package:base':

    anyDuplicated, aperm, append, as.data.frame, basename, cbind,
    colnames, dirname, do.call, duplicated, eval, evalq, Filter, Find,
    get, grep, grepl, is.unsorted, lapply, Map, mapply, match, mget,
    order, paste, pmax, pmax.int, pmin, pmin.int, Position, rank,
    rbind, Reduce, rownames, sapply, saveRDS, table, tapply, unique,
    unsplit, which.max, which.min

Loading required package: S4Vectors

Loading required package: stats4
```

Attaching package: 'S4Vectors'

The following object is masked from 'package:utils':

    findMatches

The following objects are masked from 'package:base':

    expand.grid, I, unname

Loading required package: IRanges


Attaching package: 'IRanges'

The following object is masked from 'package:grDevices':

    windows

Loading required package: GenomeInfoDb

Loading required package: GenomicRanges

Loading required package: AnnotationDbi

Loading required package: Biobase

Welcome to Bioconductor

    Vignettes contain introductory material; view with
    'browseVignettes()'. To cite Bioconductor, see
    'citation("Biobase")', and for packages 'citation("pkgname")'.

```r
library(clusterProfiler)
```

clusterProfiler v4.16.0 Learn more at https://yulab-smu.top/contribution-knowledge-mining/

Please cite:

S Xu, E Hu, Y Cai, Z Xie, X Luo, L Zhan, W Tang, Q Wang, B Liu, R Wang,
W Xie, T Wu, L Xie, G Yu. Using clusterProfiler to characterize
multiomics data. Nature Protocols. 2024, 19(11):3292-3320


Attaching package: 'clusterProfiler'

The following object is masked from 'package:AnnotationDbi':

    select

The following object is masked from 'package:IRanges':

```
    slice
```

The following object is masked from 'package:S4Vectors':

```
    rename
```

The following object is masked from 'package:stats':

```
    filter
```

```r
library(org.Mm.eg.db)  #mouse
```

```r
library(dplyr)
```

Attaching package: 'dplyr'

The following object is masked from 'package:AnnotationDbi':

```
    select
```

The following object is masked from 'package:Biobase':

```
    combine
```

The following objects are masked from 'package:GenomicRanges':

```
    intersect, setdiff, union
```

The following object is masked from 'package:GenomeInfoDb':

```
    intersect
```

The following objects are masked from 'package:IRanges':

```
    collapse, desc, intersect, setdiff, slice, union
```

The following objects are masked from 'package:S4Vectors':

```
    first, intersect, rename, setdiff, setequal, union
```

The following objects are masked from 'package:BiocGenerics':

```
    combine, intersect, setdiff, setequal, union
```

The following object is masked from 'package:generics':

```
    explain
```

The following objects are masked from 'package:stats':

```
    filter, lag
```

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union

```r
library(here)
```

here() starts at C:/Users/yjieg/Desktop/GSE169312-Analysis-main

```r
library(tidyverse)
```

── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──
✓ forcats   1.0.0      ✓ readr     2.1.5
✓ ggplot2   3.5.2      ✓ stringr   1.5.1
✓ lubridate 1.9.4      ✓ tibble    3.2.1
✓ purrr     1.0.4      ✓ tidyr     1.3.1

── Conflicts ───────────────────────────────────────── tidyverse_conflicts() ──
✗ lubridate::%within%() masks IRanges::%within%()
✗ dplyr::collapse()     masks IRanges::collapse()
✗ dplyr::combine()      masks Biobase::combine(), BiocGenerics::combine()
✗ dplyr::desc()         masks IRanges::desc()
✗ tidyr::expand()       masks S4Vectors::expand()
✗ dplyr::filter()       masks clusterProfiler::filter(), stats::filter()
✗ dplyr::first()        masks S4Vectors::first()
✗ dplyr::lag()          masks stats::lag()
✗ ggplot2::Position()   masks BiocGenerics::Position(), base::Position()
✗ purrr::reduce()       masks GenomicRanges::reduce(), IRanges::reduce()
✗ dplyr::rename()       masks clusterProfiler::rename(), S4Vectors::rename()
✗ lubridate::second()   masks S4Vectors::second()
✗ lubridate::second<-() masks S4Vectors::second<-()
✗ dplyr::select()       masks clusterProfiler::select(), AnnotationDbi::select()
✗ purrr::simplify()     masks clusterProfiler::simplify()
✗ dplyr::slice()        masks clusterProfiler::slice(), IRanges::slice()
ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become
errors

```r
library(conflicted)
library(DESeq2)
```

Loading required package: SummarizedExperiment
Loading required package: MatrixGenerics
Loading required package: matrixStats

Attaching package: 'matrixStats'

The following object is masked from 'package:dplyr':

    count

The following objects are masked from 'package:Biobase':

    anyMissing, rowMedians

```
Attaching package: 'MatrixGenerics'

The following objects are masked from 'package:matrixStats':

    colAlls, colAnyNAs, colAnys, colAvgsPerRowSet, colCollapse,
    colCounts, colCummaxs, colCummins, colCumprods, colCumsums,
    colDiffs, colIQRDiffs, colIQRs, colLogSumExps, colMadDiffs,
    colMads, colMaxs, colMeans2, colMedians, colMins, colOrderStats,
    colProds, colQuantiles, colRanges, colRanks, colSdDiffs, colSds,
    colSums2, colTabulates, colVarDiffs, colVars, colWeightedMads,
    colWeightedMeans, colWeightedMedians, colWeightedSds,
    colWeightedVars, rowAlls, rowAnyNAs, rowAnys, rowAvgsPerColSet,
    rowCollapse, rowCounts, rowCummaxs, rowCummins, rowCumprods,
    rowCumsums, rowDiffs, rowIQRDiffs, rowIQRs, rowLogSumExps,
    rowMadDiffs, rowMads, rowMaxs, rowMeans2, rowMedians, rowMins,
    rowOrderStats, rowProds, rowQuantiles, rowRanges, rowRanks,
    rowSdDiffs, rowSds, rowSums2, rowTabulates, rowVarDiffs, rowVars,
    rowWeightedMads, rowWeightedMeans, rowWeightedMedians,
    rowWeightedSds, rowWeightedVars

The following object is masked from 'package:Biobase':

    rowMedians
```

```
library(apeglm)
library(ggthemes)
conflicts_prefer(GenomicRanges::setdiff)
```

```
[conflicted] Will prefer GenomicRanges::setdiff over any other package.
```

```
#library(cowplot)
```

## Prepare Dataset

Several steps were lifted from EED-KD-DESeq2 provided from 11Apr Practical

1. Generation of `res` using DESeq
2. `res_df` adds `UP`, `DOWN` and `NC` using `logfoldchange`
3. Final output `aug_res` which filters for only genes which are differentially expressed

```
counts <-
  read_tsv(here("data","GSE169312_ReadCount.txt.gz"),
           col_types = "cnnnnnn") |>
  mutate(ID=str_remove(ID, "\\..*")) |>
  column_to_rownames("ID")

colData <- tibble(ID=names(counts)) |>
  separate(ID,into=c("genotype","replicate"),remove=FALSE) |>
  mutate(genotype=str_sub(genotype,1,-2),
```

```
        replicate=c(1:3,1:3),
        genotype=as_factor(genotype),
        genotype=fct_relevel(genotype,"WT")) |>
  column_to_rownames("ID")


rowRanges <- genes(TxDb.Mmusculus.UCSC.mm10.ensGene)
valid_genes <- rowRanges[names(rowRanges) %in% rownames(counts)]
counts <- counts[names(valid_genes),]

dds <- DESeqDataSetFromMatrix(countData=round(counts),
                    colData=colData,
                    rowRanges=valid_genes,
                    design = ~ genotype,
                    tidy=FALSE)
```

converting counts to integer mode

```
dds <- DESeq(dds)
```

estimating size factors

estimating dispersions

gene-wise dispersion estimates

mean-dispersion relationship

final dispersion estimates

fitting model and testing

```
res <- lfcShrink(dds,2)
```

using 'apeglm' for LFC shrinkage. If used in published research, please cite:
    Zhu, A., Ibrahim, J.G., Love, M.I. (2018) Heavy-tailed prior distributions for
    sequence count data: removing the noise and preserving large differences.
    Bioinformatics. https://doi.org/10.1093/bioinformatics/bty895

```
head(res)
```

log2 fold change (MAP): genotype cKO vs WT
Wald test p-value: genotype cKO vs WT
DataFrame with 6 rows and 5 columns

|                    | baseMean | log2FoldChange | lfcSE | pvalue | padj |
|--------------------|----------|----------------|-------|--------|------|
|                    | \<numeric\> | \<numeric\> | \<numeric\> | \<numeric\> | \<numeric\> |
| ENSMUSG00000000001 | 1137.53663 | 0.2114078 | 0.182723 | 1.88685e-01 | 3.94423e-01 |
| ENSMUSG00000000003 | 0.00000 | NA | NA | NA | NA |
| ENSMUSG00000000028 | 51.94398 | -0.0306503 | 0.275985 | 8.71244e-01 | 9.43170e-01 |
| ENSMUSG00000000031 | 902.56851 | 3.4603002 | 0.172836 | 2.26328e-90 | 2.14367e-86 |
| ENSMUSG00000000037 | 32.20729 | 0.3195098 | 0.396999 | 1.47782e-01 | 3.34377e-01 |
| ENSMUSG00000000049 | 2.66748 | -0.0362098 | 0.370791 | 7.20761e-01 | 8.62445e-01 |

```r
min_lfc <- 1
max_padj <- 0.05

res_df <- rownames_to_column(as.data.frame(res),"gene") |> #Full Table
  as_tibble() |>
  mutate(deg=ifelse(padj< max_padj & log2FoldChange > min_lfc, "UP","NC"),
         deg=ifelse(padj< max_padj & log2FoldChange < -min_lfc, "DOWN",deg),
         deg=as_factor(deg),
         deg=fct_relevel(deg,"NC"))

aug_res <- arrange(res_df,padj) |>  #Table filtered for genes with change
  dplyr::select(gene,log2FoldChange,padj,deg) |>
  dplyr::filter(!is.na(padj))|>
  dplyr::filter(deg!="NC")

dplyr::filter(res_df,str_starts(gene,"ENSMUSG00000022346")) #can be used to search specific ge
```

```
# A tibble: 1 × 7
  gene                baseMean log2FoldChange lfcSE   pvalue    padj deg
  <chr>                  <dbl>          <dbl> <dbl>    <dbl>   <dbl> <fct>
1 ENSMUSG00000022346      107.          0.962 0.301 0.000162 0.00143 NC
```

```r
dplyr::filter(aug_res,str_starts(gene,"ENSMUSG00000030619")) #sanity check for EED
```

```
# A tibble: 1 × 4
  gene                log2FoldChange         padj deg
  <chr>                        <dbl>        <dbl> <fct>
1 ENSMUSG00000030619           -1.06 0.0000000505 DOWN
```

## Convert Ensembl ID to Entrez ID

This step was done for compatibility with Gene Ontology

```r
ensembl_to_entrez <- bitr(aug_res$gene,
                          fromType = "ENSEMBL",
                          toType = "ENTREZID",
                          OrgDb = org.Mm.eg.db)
```

```
'select()' returned 1:many mapping between keys and columns

Warning in bitr(aug_res$gene, fromType = "ENSEMBL", toType = "ENTREZID", :
5.04% of input gene IDs are fail to map...
```

```r
head(ensembl_to_entrez)
```

```
           ENSEMBL ENTREZID
1 ENSMUSG00000020734    14813
2 ENSMUSG00000000031    14955
3 ENSMUSG00000032368    22771
4 ENSMUSG00000031654    12404
```

```
5  ENSMUSG00000002020      16997
6  ENSMUSG00000038255      18013
```

# Enrichment Analysis GO enrichment

```
ego_all <- enrichGO(gene = ensembl_to_entrez$ENTREZID,
                    OrgDb = org.Mm.eg.db,
                    keyType = "ENTREZID",
                    ont = "BP",
                    pAdjustMethod = "BH",
                    pvalueCutoff = 0.05,
                    qvalueCutoff = 0.2)
```

Visualisation of Enrichment analysis

```
dotplot(ego_all, x = "GeneRatio", showCategory = 10, font.size = 8)
```
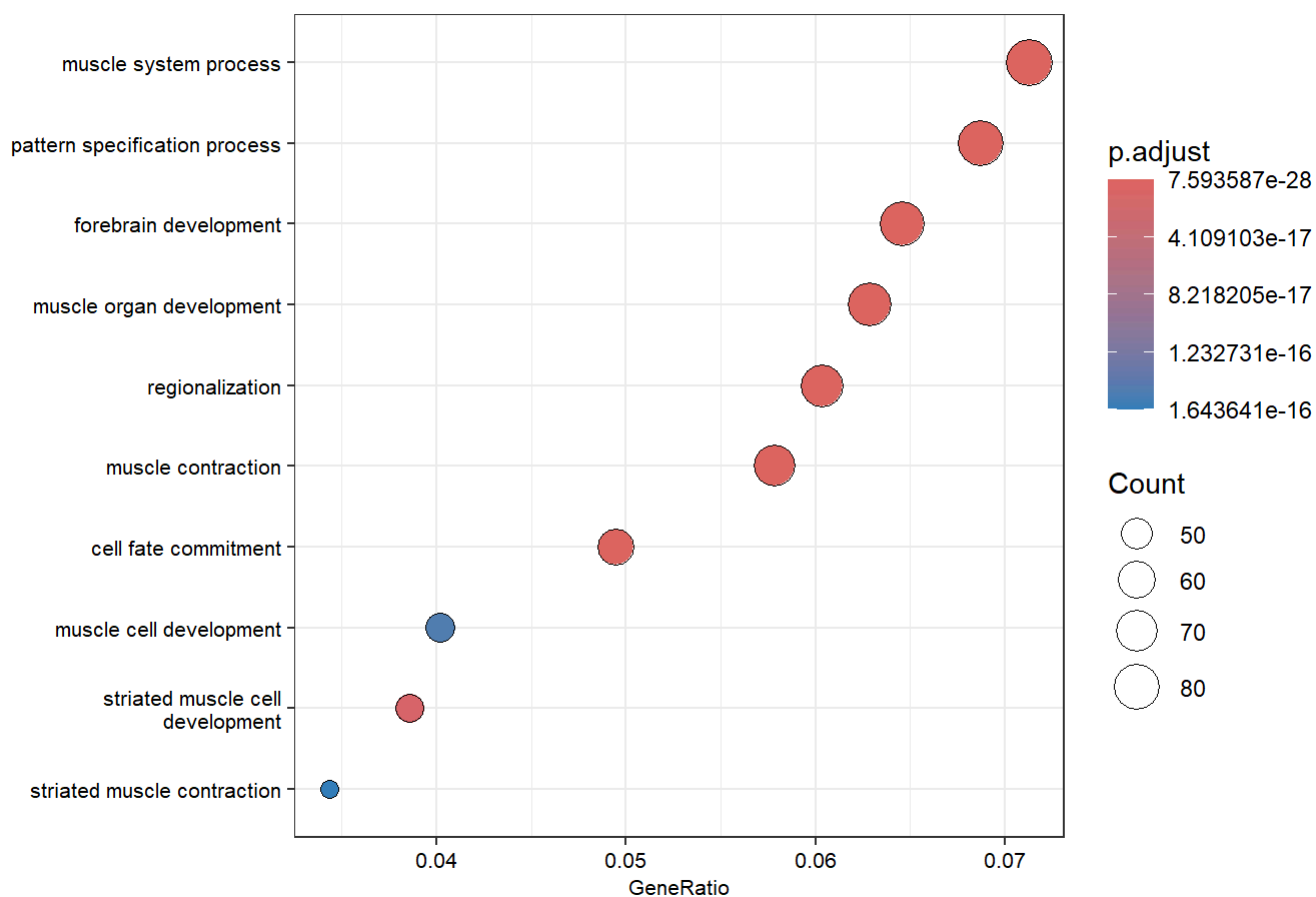


# Find potential GOI

For our report, decided to go with `muscle system process` and `cell fate commitment` This code below singles out these `categories` into `gene_list`

```r
gene_list <- list()

for (category in c(1,6)){
  cat_name <- ego_all$Description[category]
  top_genes <- ego_all$geneID[category]|>
    strsplit("/") |> unlist()

  gene_annot <- bitr(top_genes,
                     fromType = "ENTREZID",
                     toType = c("SYMBOL", "ENSEMBL"),
                     OrgDb = org.Mm.eg.db)

  gene_info <- left_join(gene_annot, aug_res, by = c("ENSEMBL" = "gene"))
  gene_list[[cat_name]] <- gene_info
  print(gene_info)
  rm(gene_info)
  rm(cat_name)
}
```

'select()' returned 1:1 mapping between keys and columns

|    | ENTREZID | SYMBOL | ENSEMBL | log2FoldChange | padj | deg |
|----|----------|--------|---------|----------------|------|-----|
| 1  | 17868 | Mybpc3 | ENSMUSG00000002100 | -3.285037 | 3.533672e-69 | DOWN |
| 2  | 17885 | Myh8 | ENSMUSG00000055775 | 6.098817 | 1.752579e-54 | UP |
| 3  | 140781 | Myh7 | ENSMUSG00000053093 | 3.336189 | 3.118137e-54 | UP |
| 4  | 17883 | Myh3 | ENSMUSG00000020908 | 5.888158 | 6.009965e-54 | UP |
| 5  | 11459 | Acta1 | ENSMUSG00000031972 | 3.659006 | 6.498663e-45 | UP |
| 6  | 21955 | Tnnt1 | ENSMUSG00000064179 | 4.283125 | 9.026646e-40 | UP |
| 7  | 21956 | Tnnt2 | ENSMUSG00000026414 | 3.670247 | 1.519758e-39 | UP |
| 8  | 21844 | Tiam1 | ENSMUSG00000002489 | -1.855328 | 2.841799e-35 | DOWN |
| 9  | 17901 | Myl1 | ENSMUSG00000061816 | 4.479728 | 9.486779e-33 | UP |
| 10 | 11928 | Atp1a1 | ENSMUSG00000033161 | -1.681143 | 1.101418e-25 | DOWN |
| 11 | 22138 | Ttn | ENSMUSG00000051747 | 3.267228 | 2.004270e-24 | UP |
| 12 | 11472 | Actn2 | ENSMUSG00000052374 | 2.579222 | 4.972608e-24 | UP |
| 13 | 12289 | Cacna1d | ENSMUSG00000015968 | -1.188217 | 2.194853e-23 | DOWN |
| 14 | 17907 | Mylpf | ENSMUSG00000030672 | 3.929845 | 2.515710e-22 | UP |
| 15 | 16590 | Kit | ENSMUSG00000005672 | -1.314365 | 3.645757e-21 | DOWN |
| 16 | 114249 | Npnt | ENSMUSG00000040998 | 1.712624 | 3.774109e-21 | UP |
| 17 | 18125 | Nos1 | ENSMUSG00000029361 | -1.517105 | 6.147387e-18 | DOWN |
| 18 | 21916 | Tmod1 | ENSMUSG00000028328 | -1.506226 | 6.939235e-17 | DOWN |
| 19 | 17896 | Myl4 | ENSMUSG00000061086 | 2.642488 | 7.528979e-14 | UP |
| 20 | 12292 | Cacna1s | ENSMUSG00000026407 | 3.116761 | 7.774481e-14 | UP |
| 21 | 21924 | Tnnc1 | ENSMUSG00000091898 | 2.475693 | 3.517298e-13 | UP |
| 22 | 12723 | Clcn1 | ENSMUSG00000029862 | -1.673393 | 3.689008e-12 | DOWN |
| 23 | 11937 | Atp2a1 | ENSMUSG00000030730 | 2.461963 | 1.185178e-11 | UP |
| 24 | 16480 | Jup | ENSMUSG00000001552 | -1.476698 | 2.728058e-11 | DOWN |
| 25 | 21957 | Tnnt3 | ENSMUSG00000061723 | 3.381889 | 3.220108e-11 | UP |
| 26 | 110355 | Grk2 | ENSMUSG00000024858 | -1.056236 | 7.314544e-11 | DOWN |
| 27 | 24046 | Scn11a | ENSMUSG00000034115 | -3.428676 | 9.221885e-11 | DOWN |
| 28 | 11464 | Actc1 | ENSMUSG00000068614 | 4.583171 | 1.636792e-10 | UP |
| 29 | 21937 | Tnfrsf1a | ENSMUSG00000030341 | 1.209811 | 2.135322e-10 | UP |
| 30 | 56484 | Foxo3 | ENSMUSG00000048756 | -1.192702 | 2.414516e-10 | DOWN |
| 31 | 17897 | Myl3 | ENSMUSG00000059741 | 4.096270 | 3.211108e-10 | UP |

| 32 | 17888 | Myh6 | ENSMUSG00000040752 | 1.549622 | 2.622357e-09 | UP |
| 33 | 17930 | Myom2 | ENSMUSG00000031461 | 2.671175 | 5.071345e-09 | UP |
| 34 | 11435 | Chrna1 | ENSMUSG00000027107 | 3.429062 | 9.612077e-09 | UP |
| 35 | 58226 | Cacna1h | ENSMUSG00000024112 | 1.178194 | 1.017027e-08 | UP |
| 36 | 15566 | Htr7 | ENSMUSG00000024798 | 1.087702 | 5.922249e-08 | UP |
| 37 | 16000 | Igf1 | ENSMUSG00000020053 | 1.463565 | 1.461240e-07 | UP |
| 38 | 30937 | Lmcd1 | ENSMUSG00000057604 | 1.340102 | 1.572098e-07 | UP |
| 39 | 11609 | Agtr2 | ENSMUSG00000068122 | 1.402095 | 2.950192e-07 | UP |
| 40 | 22004 | Tpm2 | ENSMUSG00000028464 | 1.001134 | 3.268099e-07 | UP |
| 41 | 18166 | Npy1r | ENSMUSG00000036437 | 1.180944 | 5.626425e-07 | UP |
| 42 | 16521 | Kcnj5 | ENSMUSG00000032034 | 1.840961 | 7.296662e-07 | UP |
| 43 | 18430 | Oxtr | ENSMUSG00000049112 | 1.487266 | 9.221052e-07 | UP |
| 44 | 17928 | Myog | ENSMUSG00000026459 | 4.520883 | 1.304385e-06 | UP |
| 45 | 20271 | Scn5a | ENSMUSG00000032511 | 1.844557 | 2.121191e-06 | UP |
| 46 | 216459 | Myl6b | ENSMUSG00000039824 | 1.009949 | 2.448861e-06 | UP |
| 47 | 15558 | Htr2a | ENSMUSG00000034997 | 1.303366 | 5.084021e-06 | UP |
| 48 | 17929 | Myom1 | ENSMUSG00000024049 | 1.370460 | 8.738822e-06 | UP |
| 49 | 15464 | Hrc | ENSMUSG00000038239 | 2.097225 | 1.016952e-05 | UP |
| 50 | 11447 | Chrnd | ENSMUSG00000026251 | 7.127148 | 1.106995e-05 | UP |
| 51 | 21338 | Tacr3 | ENSMUSG00000028172 | 1.327131 | 1.128574e-05 | UP |
| 52 | 59006 | Myoz2 | ENSMUSG00000028116 | 4.337664 | 4.019760e-05 | UP |
| 53 | 11449 | Chrng | ENSMUSG00000026253 | 5.604829 | 4.779331e-05 | UP |
| 54 | 11553 | Adra2c | ENSMUSG00000045318 | -1.004830 | 4.941429e-05 | DOWN |
| 55 | 21952 | Tnni1 | ENSMUSG00000026418 | 1.409954 | 5.493314e-05 | UP |
| 56 | 13489 | Drd2 | ENSMUSG00000032259 | 1.058501 | 6.303579e-05 | UP |
| 57 | 13009 | Csrp3 | ENSMUSG00000030470 | 6.649938 | 9.039323e-05 | UP |
| 58 | 12818 | Col14a1 | ENSMUSG00000022371 | 1.460632 | 9.814364e-05 | UP |
| 59 | 66139 | Mymk | ENSMUSG00000009214 | 3.050270 | 1.183294e-04 | UP |
| 60 | 93677 | Lmod2 | ENSMUSG00000029683 | 4.728488 | 1.209062e-04 | UP |
| 61 | 15111 | Hand2 | ENSMUSG00000038193 | 6.385182 | 1.324655e-04 | UP |
| 62 | 20855 | Stc1 | ENSMUSG00000014813 | 1.446896 | 1.508048e-04 | UP |
| 63 | 320502 | Lmod3 | ENSMUSG00000044086 | 3.908961 | 2.871457e-04 | UP |
| 64 | 21336 | Tacr1 | ENSMUSG00000030043 | 1.010843 | 4.898653e-04 | UP |
| 65 | 12372 | Casq1 | ENSMUSG00000007122 | 1.363591 | 9.691788e-04 | UP |
| 66 | 14463 | Gata4 | ENSMUSG00000021944 | 4.145141 | 1.059111e-03 | UP |
| 67 | 17906 | Myl2 | ENSMUSG00000013936 | 1.699490 | 1.710423e-03 | UP |
| 68 | 21953 | Tnni2 | ENSMUSG00000031097 | 2.818241 | 2.373658e-03 | UP |
| 69 | 14465 | Gata6 | ENSMUSG00000005836 | 2.286809 | 2.461635e-03 | UP |
| 70 | 21925 | Tnnc2 | ENSMUSG00000017300 | 1.090083 | 2.887646e-03 | UP |
| 71 | 57814 | Kcne4 | ENSMUSG00000047330 | 1.227629 | 3.343012e-03 | UP |
| 72 | 17884 | Myh4 | ENSMUSG00000057003 | 3.270731 | 4.305164e-03 | UP |
| 73 | 208188 | Ghsr | ENSMUSG00000051136 | 1.332655 | 5.715851e-03 | UP |
| 74 | 18091 | Nkx2-5 | ENSMUSG00000015579 | 4.829813 | 5.928879e-03 | UP |
| 75 | 17879 | Myh1 | ENSMUSG00000056328 | 2.143431 | 6.159196e-03 | UP |
| 76 | 110880 | Scn4a | ENSMUSG00000001027 | 2.214332 | 7.590250e-03 | UP |
| 77 | 212937 | Tifab | ENSMUSG00000049625 | 1.014890 | 8.701691e-03 | UP |
| 78 | 11474 | Actn3 | ENSMUSG00000006457 | 1.026342 | 9.088733e-03 | UP |
| 79 | 214384 | Myocd | ENSMUSG00000020542 | -1.532221 | 1.111247e-02 | DOWN |
| 80 | 12391 | Cav3 | ENSMUSG00000062694 | 1.473786 | 1.393800e-02 | UP |
| 81 | 12299 | Cacng1 | ENSMUSG00000020722 | 2.206740 | 1.695219e-02 | UP |
| 82 | 12062 | Bdkrb2 | ENSMUSG00000021070 | 1.042283 | 2.157061e-02 | UP |
| 83 | 20391 | Sgca | ENSMUSG00000001508 | 1.040911 | 2.162585e-02 | UP |
| 84 | 544791 | Myh13 | ENSMUSG00000060180 | -1.128758 | 2.708960e-02 | DOWN |
| 85 | 66402 | Sln | ENSMUSG00000042045 | 1.025031 | 3.791072e-02 | UP |

'select()' returned 1:1 mapping between keys and columns

|    | ENTREZID | SYMBOL   | ENSEMBL            | log2FoldChange | padj         | deg  |
|----|----------|----------|-------------------|----------------|--------------|------|
| 1  | 18508    | Pax6     | ENSMUSG00000027168 | -2.380359      | 1.592559e-64 | DOWN |
| 2  | 18012    | Neurod1  | ENSMUSG00000034701 | -3.090233      | 4.415704e-58 | DOWN |
| 3  | 14369    | Fzd7     | ENSMUSG00000041075 | -2.173615      | 3.564481e-40 | DOWN |
| 4  | 104382   | Barhl2   | ENSMUSG00000034384 | -2.631164      | 1.295091e-39 | DOWN |
| 5  | 12667    | Chrd     | ENSMUSG00000006958 | -1.749798      | 1.817624e-29 | DOWN |
| 6  | 225207   | Zfp521   | ENSMUSG00000024420 | -1.292569      | 1.058092e-24 | DOWN |
| 7  | 18032    | Nfix     | ENSMUSG00000001911 | -1.583572      | 5.945769e-20 | DOWN |
| 8  | 244810   | AW551984 | ENSMUSG00000038112 | 1.871253       | 2.389925e-19 | UP   |
| 9  | 18205    | Ntf3     | ENSMUSG00000049107 | -3.360149      | 3.311733e-19 | DOWN |
| 10 | 22421    | Wnt7a    | ENSMUSG00000030093 | -1.833948      | 2.600566e-15 | DOWN |
| 11 | 27140    | Tlx3     | ENSMUSG00000040610 | -2.907693      | 3.634650e-15 | DOWN |
| 12 | 18509    | Pax7     | ENSMUSG00000028736 | 2.141740       | 1.401780e-14 | UP   |
| 13 | 18028    | Nfib     | ENSMUSG00000008575 | -1.318202      | 3.863658e-14 | DOWN |
| 14 | 18027    | Nfia     | ENSMUSG00000028565 | -1.013502      | 3.040142e-13 | DOWN |
| 15 | 21349    | Tal1     | ENSMUSG00000028717 | 1.859340       | 3.860897e-13 | UP   |
| 16 | 18996    | Pou4f1   | ENSMUSG00000048349 | 2.426690       | 4.267381e-11 | UP   |
| 17 | 21679    | Tead4    | ENSMUSG00000030353 | -2.185924      | 1.009203e-10 | DOWN |
| 18 | 14048    | Eya1     | ENSMUSG00000025932 | 1.262008       | 2.126165e-10 | UP   |
| 19 | 58208    | Bcl11b   | ENSMUSG00000048251 | 1.291644       | 2.157478e-10 | UP   |
| 20 | 14462    | Gata3    | ENSMUSG00000015619 | 1.903745       | 1.115066e-09 | UP   |
| 21 | 20319    | Sfrp2    | ENSMUSG00000027996 | 1.365900       | 5.947973e-09 | UP   |
| 22 | 23872    | Ets2     | ENSMUSG00000022895 | -1.027853      | 1.028244e-07 | DOWN |
| 23 | 252838   | Tox      | ENSMUSG00000041272 | 1.331346       | 1.128903e-07 | UP   |
| 24 | 15398    | Hoxa13   | ENSMUSG00000038203 | 8.026440       | 1.650504e-07 | UP   |
| 25 | 13389    | Dll3     | ENSMUSG00000003436 | 1.840776       | 2.599102e-07 | UP   |
| 26 | 209448   | Hoxc10   | ENSMUSG00000022484 | 7.783423       | 5.626425e-07 | UP   |
| 27 | 13813    | Eomes    | ENSMUSG00000032446 | -1.126230      | 6.428820e-07 | DOWN |
| 28 | 21907    | Nr2e1    | ENSMUSG00000019803 | 3.011551       | 9.600470e-07 | UP   |
| 29 | 242620   | Dmrta2   | ENSMUSG00000047143 | 2.693338       | 3.837707e-06 | UP   |
| 30 | 21384    | Tbx15    | ENSMUSG00000027868 | 1.626185       | 1.107017e-05 | UP   |
| 31 | 11924    | Neurog2  | ENSMUSG00000027967 | 1.653314       | 2.568559e-05 | UP   |
| 32 | 21420    | Tfap2c   | ENSMUSG00000028640 | 4.423814       | 4.098037e-05 | UP   |
| 33 | 53404    | Atoh7    | ENSMUSG00000036816 | -1.475371      | 5.299900e-05 | DOWN |
| 34 | 76365    | Tbx18    | ENSMUSG00000032419 | 1.044394       | 1.016128e-04 | UP   |
| 35 | 22410    | Wnt10b   | ENSMUSG00000022996 | 1.453306       | 2.696575e-04 | UP   |
| 36 | 16814    | Lbx1     | ENSMUSG00000025216 | -1.119467      | 5.653921e-04 | DOWN |
| 37 | 20472    | Six2     | ENSMUSG00000024134 | 2.198035       | 9.569080e-04 | UP   |
| 38 | 14463    | Gata4    | ENSMUSG00000021944 | 4.145141       | 1.059111e-03 | UP   |
| 39 | 17906    | Myl2     | ENSMUSG00000013936 | 1.699490       | 1.710423e-03 | UP   |
| 40 | 17172    | Ascl1    | ENSMUSG00000020052 | 1.027862       | 1.979377e-03 | UP   |
| 41 | 12394    | Runx1    | ENSMUSG00000022952 | 1.277966       | 2.036879e-03 | UP   |
| 42 | 15376    | Foxa2    | ENSMUSG00000037025 | 4.451687       | 2.069659e-03 | UP   |
| 43 | 15430    | Hoxd10   | ENSMUSG00000050368 | 5.314122       | 2.352863e-03 | UP   |
| 44 | 14465    | Gata6    | ENSMUSG00000005836 | 2.286809       | 2.461635e-03 | UP   |
| 45 | 83430    | Il23a    | ENSMUSG00000025383 | -1.000048      | 2.529036e-03 | DOWN |
| 46 | 16392    | Isl1     | ENSMUSG00000042258 | 1.576139       | 3.101229e-03 | UP   |
| 47 | 14836    | Gsc      | ENSMUSG00000021095 | 5.052906       | 3.518054e-03 | UP   |
| 48 | 18740    | Pitx1    | ENSMUSG00000021506 | 4.049199       | 4.473542e-03 | UP   |
| 49 | 12393    | Runx2    | ENSMUSG00000039153 | 1.164319       | 4.707132e-03 | UP   |
| 50 | 18091    | Nkx2-5   | ENSMUSG00000015579 | 4.829813       | 5.928879e-03 | UP   |
| 51 | 109575   | Tbx10    | ENSMUSG00000037477 | 1.798981       | 1.285453e-02 | UP   |

```
52     15379   Onecut1 ENSMUSG00000043013    1.323096 1.341400e-02    UP
53     17085       Ly9 ENSMUSG00000004707    1.987159 1.497071e-02    UP
54     13392      Dlx2 ENSMUSG00000023391    2.815740 2.014856e-02    UP
55    104360      Isl2 ENSMUSG00000032318    2.200315 2.726464e-02    UP
56     15375     Foxa1 ENSMUSG00000035451    1.861775 2.875375e-02    UP
57    207920     Esrp1 ENSMUSG00000040728    1.496011 2.976157e-02    UP
58     21416    Tcf7l2 ENSMUSG00000024985    1.069636 3.418644e-02    UP
59     22423     Wnt8b ENSMUSG00000036961    1.484429 3.461972e-02    UP
```