

Marques et al 2021 WT1bOE ATACseq

Contents

ATAC seq workflow	2
Preprocessing and analysis overview	2
Analysis	3
Clear memory and set working directory	3
Diffbound analysis	9
Read Peak files	9
Create DBA object for Diffbound	10
QC of samples	11
Count the peaks in samples	13
Set the FDR and pvalue	15
Data Normalization for the analysis	15
Plot PCA and Boxplot for normalized data	19
Annotation of Peaks	23
Get the zebrafish genome data from Ensembl (a TxDB object that can be used for ChIPSeeker)	23
Make .bed files from narrowpeak files (in bash)	24
Read .bed files	25

Annotation of peaks using ChipSeeker	26
Read narrowpeak bed files	26
Annotate the peaks	28
Check the genomic regions assigned to the peaks (promoter, exon, introns, etc)	28
Check the distance to Transcription Start Site (TSS)	30
Compare if the percentage and absolute number of peaks differ in different in Wt1bOE vs Ctrl	32
Check the absolute counts of annotated peaks	42
Annotation of the differential peaks and representation in graphs (promoters, exons, introns, etc)	46
Correct the names of the annotated dataframe of peaks	52
Convert genes names to symbol and entrezid	57
Pathway analysis	61
Convert zebrafish genes to mouse genes and merge to original dataframe	61
Perfroming enrichment analysis on differential peaks/genes.	64
Perform GO analysis on genes common in CHIPT-ATLAS and current ATACseq analysis	106
Save RData and end of program	142

ATAC seq workflow

Preprocessing and analysis overview

The data has been analyzed in the following ways before bringing it to R:

- 1) FASTQC
- 2) Trimming using Trimmomatic
- 3) Alignment using bowtie2
- 4) Remove mitochondria and Fluorescent Proteins
- 5) Remove duplicates using Picard
- 6) Check peak distribution

- 7) Downsample the samples to lowest reads
- 8) Run genrich to call peaks

in R (this notebook)

- 8) Run DiffBind for differential peaks
- 9) Run ChipSeeker for annotation
- 10) Run Clusterprofiler for pathway enrichment

Analysis

Clear memory and set working directory

```
rm(list = ls())  
gc()
```

```
##          used (Mb) gc trigger (Mb) max used (Mb)  
## Ncells 425430 22.8      882741 47.2    665545 35.6  
## Vcells 812841  6.3      8388608 64.0   1819096 13.9
```

```
setwd("/home/prateek/Mercader_Lab/Ines_Data/wt1oe_ATAC_Seq/final")  
getwd()
```

```
## [1] "/home/prateek/Mercader_Lab/Ines_Data/wt1oe_ATAC_Seq/final"
```

```
library(dplyr)
```

```
##  
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':  
##  
##   filter, lag
```

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

```
##
```

```
## intersect, setdiff, setequal, union
```

```
library(DiffBind)
```

```
## Loading required package: GenomicRanges
```

```
## Loading required package: stats4
```

```
## Loading required package: BiocGenerics
```

```
## Loading required package: parallel
```

```
##
```

```
## Attaching package: 'BiocGenerics'
```

```
## The following objects are masked from 'package:parallel':
```

```
##
```

```
## clusterApply, clusterApplyLB, clusterCall, clusterEvalQ,
```

```
## clusterExport, clusterMap, parApply, parCapply, parLapply,
```

```
## parLapplyLB, parRapply, parSapply, parSapplyLB
```

```
## The following objects are masked from 'package:dplyr':
```

```
##
```

```
## combine, intersect, setdiff, union
```

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

```
##
```

```
## IQR, mad, sd, var, xtabs
```

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

```
##
```

```
## anyDuplicated, append, as.data.frame, basename, cbind, colnames,
```

```
## dirname, do.call, duplicated, eval, evalq, Filter, Find, get, grep,
```

```
## grepl, intersect, is.unsorted, lapply, Map, mapply, match, mget,
```

```
## order, paste, pmax, pmax.int, pmin, pmin.int, Position, rank,
```

```
## rbind, Reduce, rownames, sapply, setdiff, sort, table, tapply,
```

```
## union, unique, unsplit, which.max, which.min
```

```

## Loading required package: S4Vectors

##
## Attaching package: 'S4Vectors'

## The following objects are masked from 'package:dplyr':
##
##     first, rename

## The following objects are masked from 'package:base':
##
##     expand.grid, I, unname

## Loading required package: IRanges

##
## Attaching package: 'IRanges'

## The following objects are masked from 'package:dplyr':
##
##     collapse, desc, slice

## Loading required package: GenomeInfoDb

## Loading required package: SummarizedExperiment

## Loading required package: MatrixGenerics

## Loading required package: matrixStats

##
## Attaching package: 'matrixStats'

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

```

```

##
## Attaching package: 'MatrixGenerics'

## The following objects are masked from 'package:matrixStats':
##
##   colAlls, colAnyNAs, colAnys, colAvgPerRowSet, 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, rowAvgPerColSet,
##   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

## 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)".

##
## Attaching package: 'Biobase'

## The following object is masked from 'package:MatrixGenerics':
##
##   rowMedians

## The following objects are masked from 'package:matrixStats':
##
##   anyMissing, rowMedians

```

```
##
```

```
## >>> DiffBind 3.2.
```

```
library(GenomicFeatures)
```

```
## Loading required package: AnnotationDbi
```

```
##
```

```
## Attaching package: 'AnnotationDbi'
```

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

```
##
```

```
##      select
```

```
library(ChIPseeker)
```

```
## ChIPseeker v1.28.3 For help: https://guangchuangyu.github.io/software/ChIPseeker
```

```
##
```

```
## If you use ChIPseeker in published research, please cite:
```

```
## Guangchuang Yu, Li-Gen Wang, Qing-Yu He. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization.
```

```
library(rstatix)
```

```
##
```

```
## Attaching package: 'rstatix'
```

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

```
##
```

```
##      select
```

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

```
##
```

```
##      desc
```

```
## The following object is masked from 'package:stats':  
##  
## filter
```

```
library(ggpubr)
```

```
## Loading required package: ggplot2
```

```
library(RColorBrewer)  
library("viridis")
```

```
## Loading required package: viridisLite
```

```
library(ReactomePA)
```

```
## ReactomePA v1.36.0 For help: https://yulab-smu.top/biomedical-knowledge-mining-book/reactomepa.html
```

```
##
```

```
## If you use ReactomePA in published research, please cite:
```

```
## Guangchuang Yu, Qing-Yu He. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. Molecular BioSystems
```

```
library(clusterProfiler)
```

```
## clusterProfiler v4.0.0 For help: https://guangchuangyu.github.io/software/clusterProfiler
```

```
##
```

```
## If you use clusterProfiler in published research, please cite:
```

```
## Guangchuang Yu, Li-Gen Wang, Yanyan Han, Qing-Yu He. clusterProfiler: an R package for comparing biological themes among gene clusters.
```

```
##
```

```
## 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
```

```
library(org.Mm.eg.db)
```

```
##
```

Diffbound analysis

Read Peak files

```
#make samplesheet to read the peak files
samplesheet_df <- openxlsx::read.xlsx("DiffBind_Samplesheet_generich_downsampled.xlsx")

samplesheet_df$bamReads <- paste0("only_unique_downsampled/", samplesheet_df$bamReads)
samplesheet_df$Peaks <- paste0("Generich_downsampled/", samplesheet_df$Peaks)
(samplesheet_df)
```

```
##      SampleID Condition Replicate
## 1      URV_1      Ctrl          1
## 2      URV_3      Ctrl          3
## 3      URV_4      Ctrl          4
## 4 wt1bOE_1      WT1bOE          1
## 5 wt1bOE_2      WT1bOE          2
## 6 wt1bOE_3      WT1bOE          3
##
##                                     bamReads
## 1  only_unique_downsampled/URV_1_only_unique_downsampled.bam
## 2  only_unique_downsampled/URV_3_only_unique_downsampled.bam
## 3  only_unique_downsampled/URV_4_only_unique_downsampled.bam
## 4  only_unique_downsampled/wt1bOE_1_only_unique_downsampled.bam
```

```
## 5 only_unique_downsampled/wt1bOE_2_only_unique_downsampled.bam
## 6 only_unique_downsampled/wt1bOE_3_only_unique_downsampled.bam
##                               Peaks PeakCaller
## 1   Generich_downsampled/URV_1_downsampled.narrowPeak   narrow
## 2   Generich_downsampled/URV_3_downsampled.narrowPeak   narrow
## 3   Generich_downsampled/URV_4_downsampled.narrowPeak   narrow
## 4   Generich_downsampled/wt1bOE_1_downsampled.narrowPeak   narrow
## 5   Generich_downsampled/wt1bOE_2_downsampled.narrowPeak   narrow
## 6   Generich_downsampled/wt1bOE_3_downsampled.narrowPeak   narrow
```

Create DBA object for Diffbound

```
wt1oe_db <- dba(sampleSheet = samplesheet_df)
```

```
## URV_1   Ctrl  1 narrow
```

```
## URV_3   Ctrl  3 narrow
```

```
## URV_4   Ctrl  4 narrow
```

```
## wt1bOE_1 WT1bOE  1 narrow
```

```
## wt1bOE_2 WT1bOE  2 narrow
```

```
## wt1bOE_3 WT1bOE  3 narrow
```

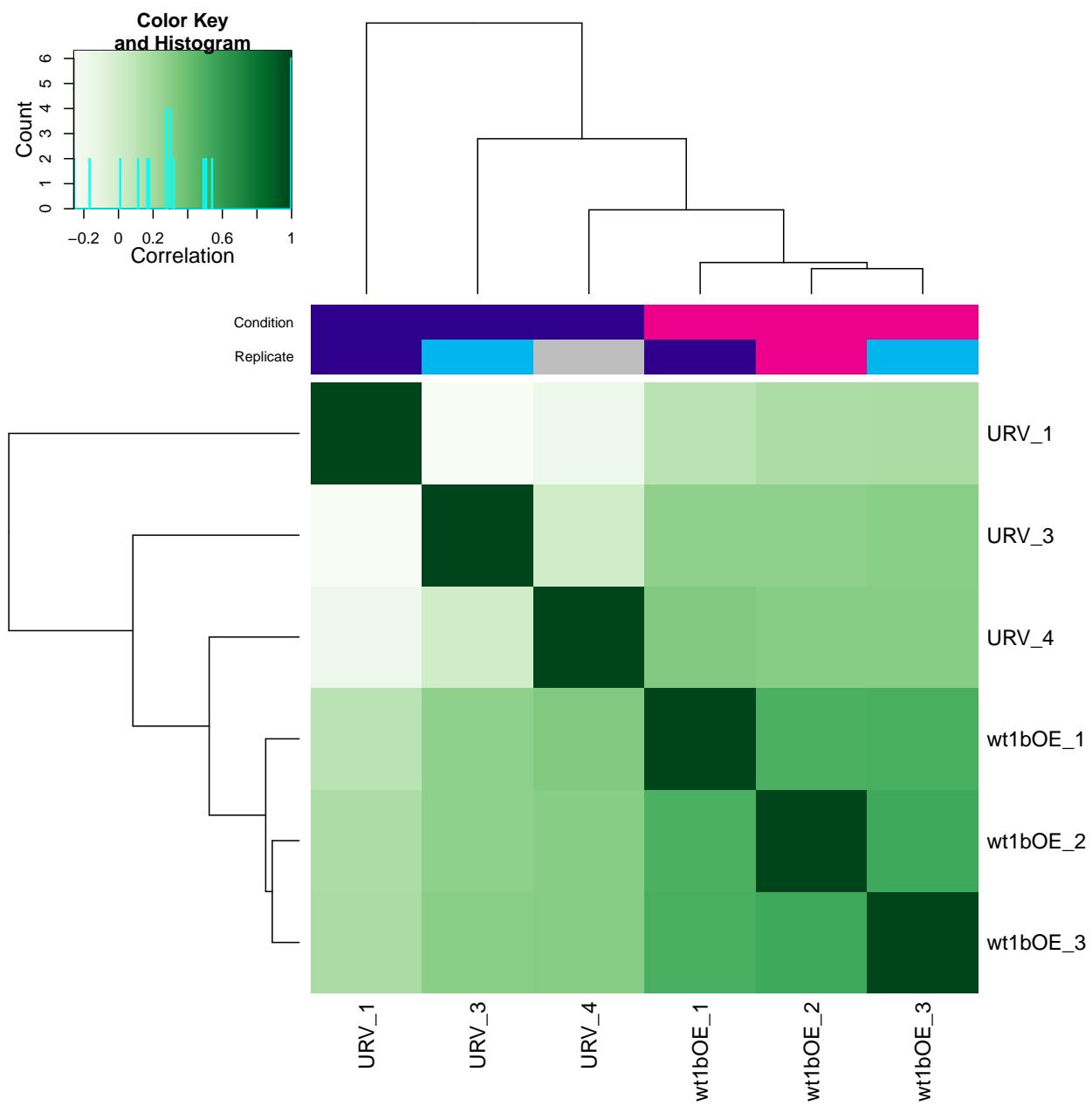
```
wt1oe_db
```

```
## 6 Samples, 12374 sites in matrix (72406 total):
```

```
##      ID Condition Replicate Intervals
## 1   URV_1      Ctrl         1    30372
## 2   URV_3      Ctrl         3    30070
## 3   URV_4      Ctrl         4    23705
## 4 wt1bOE_1    WT1bOE         1     9452
## 5 wt1bOE_2    WT1bOE         2     4212
## 6 wt1bOE_3    WT1bOE         3     4188
```

QC of samples

```
plot(wt1oe_db)
```



Count the peaks in samples

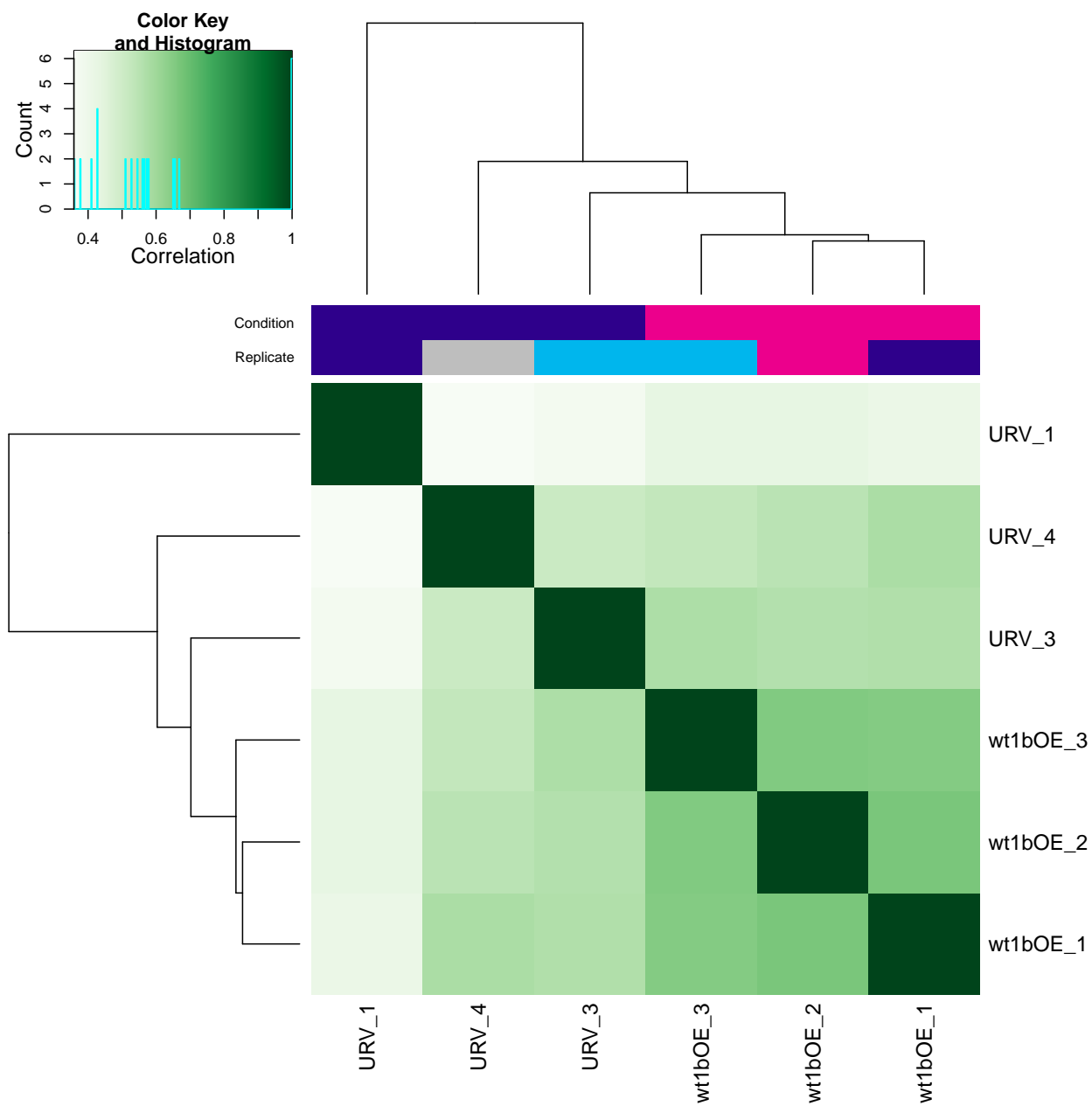
(it counts a peak if the peak is present in more than one sample)

```
wt1oe_db <- dba.count(wt1oe_db, bParallel = TRUE)
```

```
## Computing summits...
```

```
## Re-centering peaks...
```

```
plot(wt1oe_db)
```



```
wt1oe_db
```

```
## 6 Samples, 12169 sites in matrix:
##      ID Condition Replicate  Reads FRiP
## 1   URV_1      Ctrl         1 2434515 0.04
## 2   URV_3      Ctrl         3 2437084 0.05
## 3   URV_4      Ctrl         4 2719793 0.05
## 4 wt1bOE_1    WT1bOE         1 3260957 0.05
## 5 wt1bOE_2    WT1bOE         2 3630472 0.03
## 6 wt1bOE_3    WT1bOE         3 3577586 0.03
```

Set the FDR and pvalue

```
wt1oe_db$config$th = 0.05
```

Data Normalization for the analysis

Using “DeSeq2” and “Background” normalization as it is recommended for ATACSeq (see Diffbind vignette)

```
wt1oe_db <- dba.normalize(wt1oe_db, method=DBA_ALL_METHODS, normalize=DBA_NORM_NATIVE, background=TRUE)
```

```
## Generating background bins...
```

```
wt1oe_db <- dba.analyze(wt1oe_db, method=DBA_ALL_METHODS)
```

```
## Applying Blacklist/Greylists...
```

```
## No genome detected.
```

```
## Forming default model design and contrast(s)...
```

```
## Computing results names...
```

```
## Analyzing...
```

```
dba.show(wt1oe_db,bContrasts=TRUE)
```

```
##      Factor  Group Samples Group2 Samples2 DB.edgeR DB.DESeq2
## 1 Condition WT1bOE      3  Ctrl      3      7      0
```

```
db_deseq2 <- dba.report(wt1oe_db,method=DBA_DESEQ2, bDB=TRUE, bGain=TRUE, bLoss=TRUE,bUsePval = T, bNormalized = T)
```

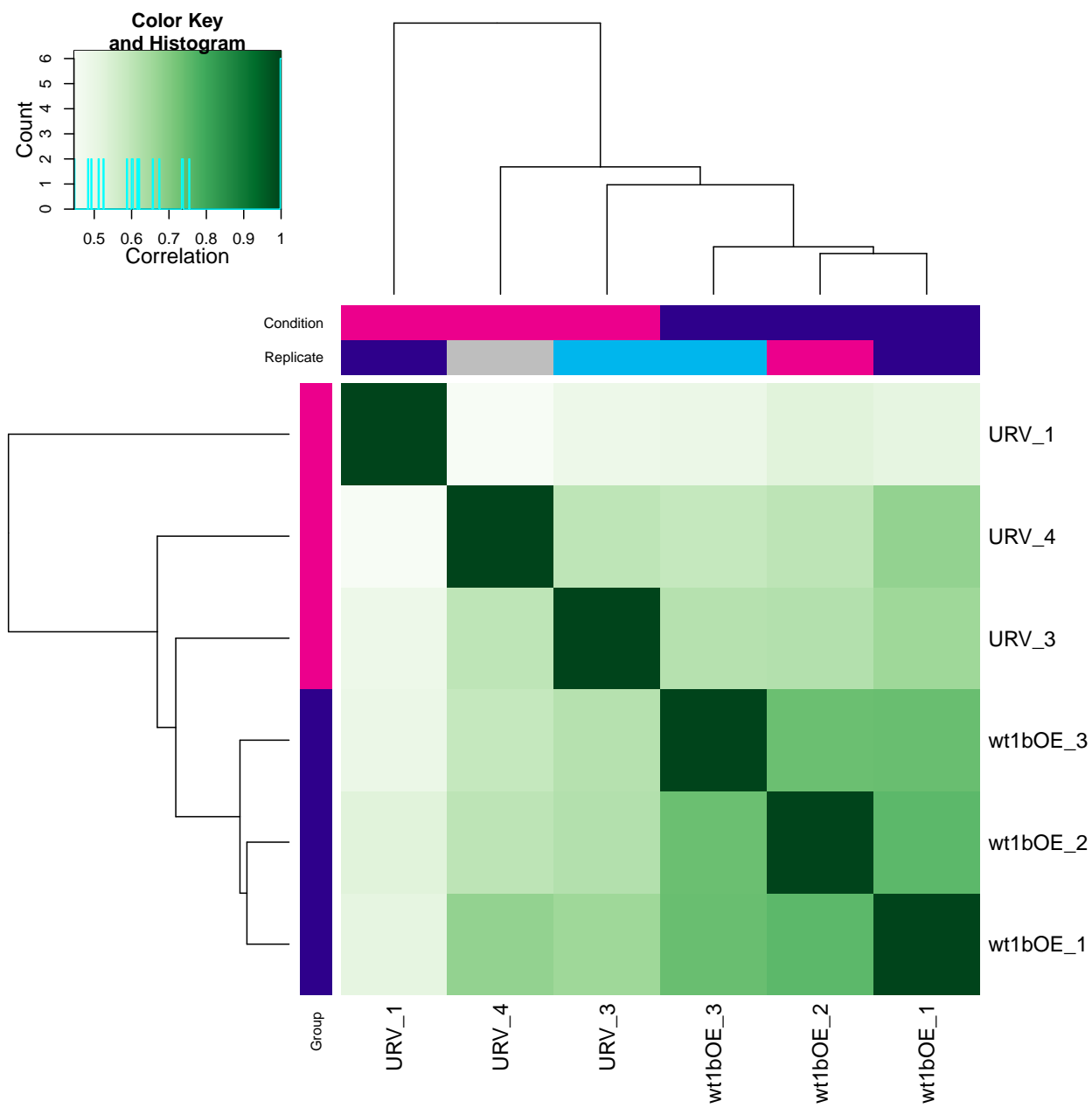
```
## Generating report-based DBA object...
```

```
db_deseq2
```

```
## 3 Samples, 1014 sites in matrix:
```

```
##      Contrast Direction DB Method Intervals
## 1 WT1bOE vs. Ctrl      All DB DESeq2      1014
## 2 WT1bOE vs. Ctrl      Gain DB DESeq2        26
## 3 WT1bOE vs. Ctrl      Loss DB DESeq2      988
```

```
plot(wt1oe_db, contrast=1, bUsePval = T)
```

Write the differential peaks in csv file

```
getwd()
```

```
## [1] "/home/prateek/Mercader_Lab/Ines_Data/WT1bOE_ATAC_Seq/final"
```

```
dba.report(wt1bOE_db,method=DBA_DESEQ2, bCalledDetail = T,file = "./results/DBA_wt1bOE_downsampled_3_samples_generich_deseq_norm",  
           bGain=TRUE, bLoss=TRUE,bUsePval = T, bNormalized = T)
```

```
## Generating report-based DBA object...
```

```
## 3 Samples, 1014 sites in matrix:
```

```
##           Contrast Direction DB Method Intervals  
## 1 WT1bOE vs. Ctrl      All DB DESeq2      1014  
## 2 WT1bOE vs. Ctrl      Gain DB DESeq2       26  
## 3 WT1bOE vs. Ctrl      Loss DB DESeq2      988
```

```
db_deseq2_df <- read.csv("./results/DBA_wt1bOE_downsampled_3_samples_generich_deseq_norm.csv")  
head(db_deseq2_df)
```

```
##   Chr   Start   End   Conc Conc_WT1bOE Conc_Ctrl   Fold p.value  
## 1   2 14006002 14006402 4.256871  2.669184  4.994381 -1.1746329 1.04e-05  
## 2  13 50200079 50200479 3.949543  2.408604  4.677540 -1.1351088 1.82e-05  
## 3  17 53383459 53383859 9.149312  8.826909  9.412664 -0.5387331 2.16e-05  
## 4  13 52114475 52114875 4.578848  3.273357  5.252782 -1.0901966 2.99e-05  
## 5  21 3470696  3471096 4.318552  2.898409  5.020167 -1.0776029 3.96e-05  
## 6   4 77761292 77762031 3.578587  1.695711  4.368405 -1.0402813 4.56e-05  
##           FDR wt1bOE_1 wt1bOE_2 wt1bOE_3 URV_1 URV_3 URV_4 peak_name  
## 1 0.05699248      +      -      +      +      +      + peak_1  
## 2 0.05699248      -      +      -      +      +      + peak_2  
## 3 0.05699248      +      +      +      +      +      + peak_3  
## 4 0.05925655      +      +      -      +      +      + peak_4  
## 5 0.06020326      +      -      -      +      +      + peak_5  
## 6 0.06020326      -      -      -      +      +      + peak_6
```

```
?dba.report
```

```
diff_peaks_pval_norm <- as.data.frame(db_deseq2_df)
head(diff_peaks_pval_norm)
```

```
##   Chr   Start   End   Conc Conc_WT1bOE Conc_Ctrl   Fold p.value
## 1    2 14006002 14006402 4.256871    2.669184  4.994381 -1.1746329 1.04e-05
## 2   13 50200079 50200479 3.949543    2.408604  4.677540 -1.1351088 1.82e-05
## 3   17 53383459 53383859 9.149312    8.826909  9.412664 -0.5387331 2.16e-05
## 4   13 52114475 52114875 4.578848    3.273357  5.252782 -1.0901966 2.99e-05
## 5   21 3470696  3471096 4.318552    2.898409  5.020167 -1.0776029 3.96e-05
## 6    4 77761292 77762031 3.578587    1.695711  4.368405 -1.0402813 4.56e-05
##           FDR wt1bOE_1 wt1bOE_2 wt1bOE_3 URV_1 URV_3 URV_4 peak_name
## 1 0.05699248      +      -      +      +      +      + peak_1
## 2 0.05699248      -      +      -      +      +      + peak_2
## 3 0.05699248      +      +      +      +      +      + peak_3
## 4 0.05925655      +      +      -      +      +      + peak_4
## 5 0.06020326      +      -      -      +      +      + peak_5
## 6 0.06020326      -      -      -      +      +      + peak_6
```

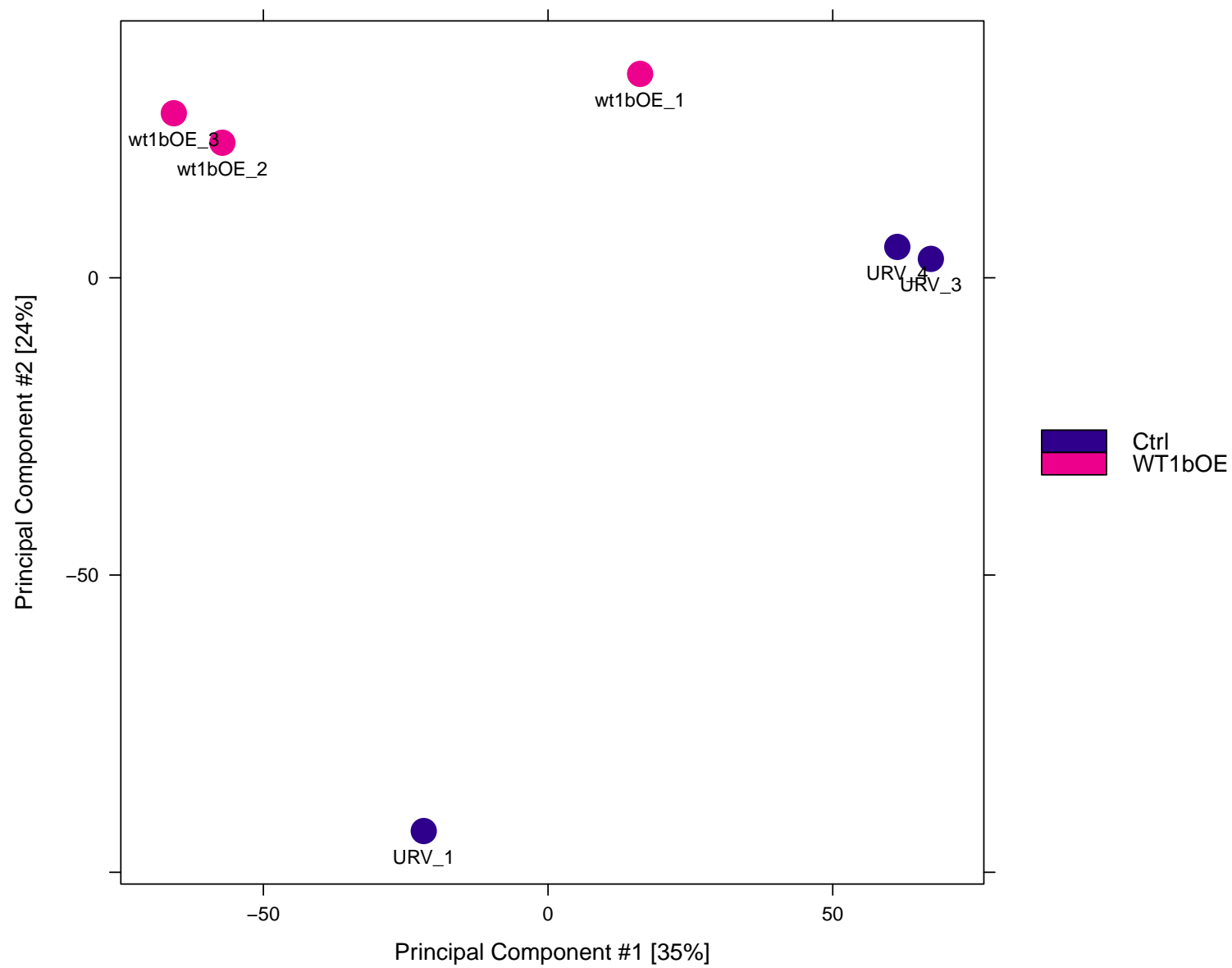
```
write.csv(diff_peaks_pval_norm, "./results/diff_peaks_downsample_generich_3_samples_pval_norm.csv")
write.table(diff_peaks_pval_norm, "./results/diff_peaks_downsample_generich_3_samples_pval_norm.bed",
            quote=F, sep="\t", row.names=F, col.names=F)
```

Plot PCA and Boxplot for normalized data

```
dba.plotPCA(wt1oe_db, DBA_CONDITION, label=DBA_ID)
```

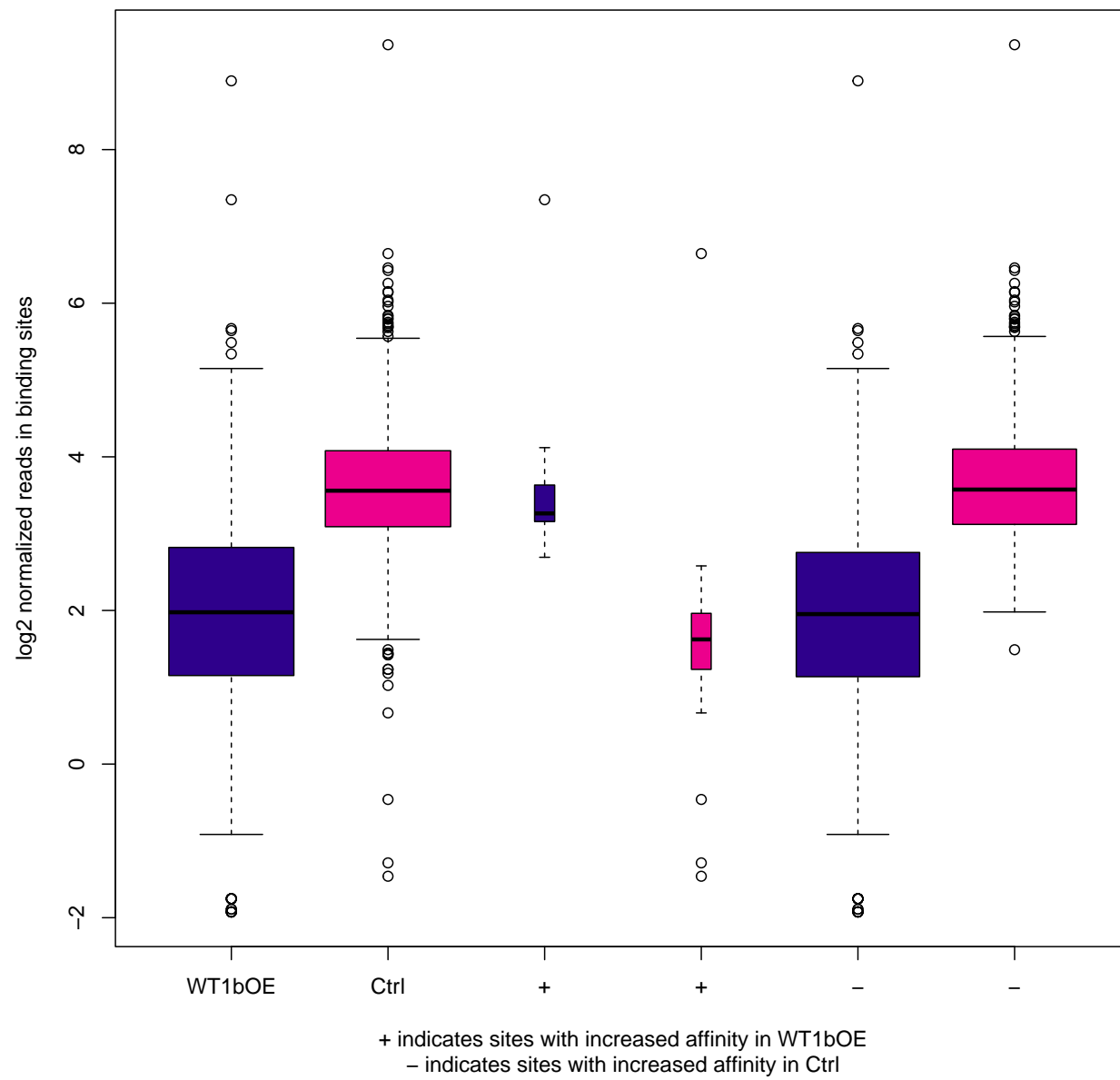
```
## Warning in sprintf("PCA: %s", pv.attrname(class, pv), pvar): one argument not
## used by format 'PCA: %s'
```

PCA: Condition



```
pvals <- dba.plotBox(wt1oe_db, bUsePval = T, notch= FALSE)
```

WT1bOE vs. Ctrl



Annotation of Peaks

Get the zebrafish genome data from Ensembl (a TxDB object that can be used for ChiPSeeker)

```
TxDBDR102 <- makeTxDbFromEnsembl(organism="Danio rerio",  
                                release=102,  
                                circ_seqs=NULL,  
                                server="ensembl.db.ensembl.org",  
                                username="anonymous", password=NULL, port=0L,  
                                tx_attrib=NULL)
```

```
## Fetch transcripts and genes from Ensembl ... OK  
##   (fetched 65905 transcripts from 37241 genes)  
## Fetch exons and CDS from Ensembl ... OK  
## Fetch chromosome names and lengths from Ensembl ...OK  
## Gather the metadata ... OK  
## Make the TxDb object ... OK
```

TxDBDR102

```
## TxDb object:  
## # Db type: TxDb  
## # Supporting package: GenomicFeatures  
## # Data source: Ensembl  
## # Organism: Danio rerio  
## # Ensembl release: 102  
## # Ensembl database: danio_rerio_core_102_11  
## # MySQL server: ensembl.db.ensembl.org  
## # Full dataset: yes  
## # Nb of transcripts: 65905  
## # Db created by: GenomicFeatures package from Bioconductor  
## # Creation time: 2021-07-06 10:00:16 +0200 (Tue, 06 Jul 2021)  
## # GenomicFeatures version at creation time: 1.44.0  
## # RSQLite version at creation time: 2.2.7  
## # DBSCHEMAVERSION: 1.2
```

```
saveDb(TxDBDR102, file = "Danio_rerio_annotationdbi_ensbl_v102.sqlite")
```

```
## TxDb object:
## # Db type: TxDb
## # Supporting package: GenomicFeatures
## # Data source: Ensembl
## # Organism: Danio rerio
## # Ensembl release: 102
## # Ensembl database: danio_rerio_core_102_11
## # MySQL server: ensembl.db.ensembl.org
## # Full dataset: yes
## # Nb of transcripts: 65905
## # Db created by: GenomicFeatures package from Bioconductor
## # Creation time: 2021-07-06 10:00:16 +0200 (Tue, 06 Jul 2021)
## # GenomicFeatures version at creation time: 1.44.0
## # RSQLite version at creation time: 2.2.7
## # DBSCHEMAVERSION: 1.2
```

```
# txdb <- TxDBDR102
txdb <- loadDb("Danio_rerio_annotationdbi_ensbl_v102.sqlite")
```

Make .bed files from narrowpeak files (in bash)

```
inputFiles_R1="./trim_paired/input_files_R1.txt"
array_R1=($(cat inputFiles_R1))

for SAMPLE in {0..5}
do
  #set input file 1 to "FL1", input file 2 to "FL2"
  FL1=${array_R1[SAMPLE]}
  #echo ${FL1}
  FL1_trimmed=${FL1%_L1*}
  echo ${FL1_trimmed}

  cut -f 1-6 ./Generich_downsampled/${FL1_trimmed}_downsampled.narrowPeak > ./Generich_downsampled/${FL1_trimmed}_downsampled_narrowpeak.bed
```



```
done
```

```
## URV_1  
## URV_3  
## URV_4  
## wt1bOE_1  
## wt1bOE_2  
## wt1bOE_3
```

Read .bed files

```
peaks_files <- list.files(path = "./Generich_downsampled/", pattern = "*narrowpeak.bed")  
peaks_files
```

```
## [1] "URV_1_downsampled_narrowpeak.bed"    "URV_3_downsampled_narrowpeak.bed"  
## [3] "URV_4_downsampled_narrowpeak.bed"    "wt1bOE_1_downsampled_narrowpeak.bed"  
## [5] "wt1bOE_2_downsampled_narrowpeak.bed" "wt1bOE_3_downsampled_narrowpeak.bed"
```

```
class(peaks_files)
```

```
## [1] "character"
```

```
remove_files <- peaks_files[grepl(pattern = "bdg", peaks_files)]  
remove_files
```

```
## character(0)
```

```
peaks_files <- peaks_files[!(peaks_files %in% remove_files)]  
peaks_files
```

```
## [1] "URV_1_downsampled_narrowpeak.bed"    "URV_3_downsampled_narrowpeak.bed"  
## [3] "URV_4_downsampled_narrowpeak.bed"    "wt1bOE_1_downsampled_narrowpeak.bed"  
## [5] "wt1bOE_2_downsampled_narrowpeak.bed" "wt1bOE_3_downsampled_narrowpeak.bed"
```

```
sample_names <- gsub("_downsampled_narrowpeak.bed","", peaks_files)
sample_names
```

```
## [1] "URV_1"      "URV_3"      "URV_4"      "wt1b0E_1" "wt1b0E_2" "wt1b0E_3"
```

Annotation of peaks using ChipSeeker

Read narrowpeak bed files

```
peakfiles_granges <- list.files(path = "./Generich_downsampled/", pattern = "narrowpeak.bed")

peakfiles_granges <- peakfiles_granges[!peakfiles_granges %in% remove_files]
peakfiles_granges
```

```
## [1] "URV_1_downsampled_narrowpeak.bed"      "URV_3_downsampled_narrowpeak.bed"
## [3] "URV_4_downsampled_narrowpeak.bed"      "wt1b0E_1_downsampled_narrowpeak.bed"
## [5] "wt1b0E_2_downsampled_narrowpeak.bed"   "wt1b0E_3_downsampled_narrowpeak.bed"
```

```
peakfiles_granges <- paste0("./Generich_downsampled/",peakfiles_granges)
peakfiles_granges
```

```
## [1] "./Generich_downsampled/URV_1_downsampled_narrowpeak.bed"
## [2] "./Generich_downsampled/URV_3_downsampled_narrowpeak.bed"
## [3] "./Generich_downsampled/URV_4_downsampled_narrowpeak.bed"
## [4] "./Generich_downsampled/wt1b0E_1_downsampled_narrowpeak.bed"
## [5] "./Generich_downsampled/wt1b0E_2_downsampled_narrowpeak.bed"
## [6] "./Generich_downsampled/wt1b0E_3_downsampled_narrowpeak.bed"
```

```
read_peaks <- readPeakFile(peakfiles_granges[1])
read_peaks
```

```
## GRanges object with 30372 ranges and 3 metadata columns:
##           seqnames      ranges strand |           V4           V5           V6
```

```
##          <Rle>   <IRanges> <Rle> | <character> <integer> <character>
##      [1]          1   6679-6778    * |      peak_0         1000      .
##      [2]          1 21229-22646    * |      peak_1          664      .
##      [3]          1 27036-27404    * |      peak_2          869      .
##      [4]          1 27560-27873    * |      peak_3         1000      .
##      [5]          1 36576-36771    * |      peak_4         1000      .
##      ...          ...           ...  ...  ...           ...           ...
## [30368] KN150583.1   1386-1918    * | peak_30367          509      .
## [30369] KN150503.1       620-764    * | peak_30368         1000      .
## [30370] KN150437.1       569-1232    * | peak_30369         1000      .
## [30371] KN150621.1       215-890    * | peak_30370          851      .
## [30372] KN150544.1        37-836    * | peak_30371         1000      .
## -----
## seqinfo: 287 sequences from an unspecified genome; no seqlengths
```

```
peakfiles_granges_names <- peakfiles_granges
peakfiles_granges_names <- as.list(peakfiles_granges_names)
names(peakfiles_granges_names) <- sample_names
peakfiles_granges_names
```

```
## $URV_1
## [1] "./Generich_downsampled/URV_1_downsampled_narrowpeak.bed"
##
## $URV_3
## [1] "./Generich_downsampled/URV_3_downsampled_narrowpeak.bed"
##
## $URV_4
## [1] "./Generich_downsampled/URV_4_downsampled_narrowpeak.bed"
##
## $wt1bOE_1
## [1] "./Generich_downsampled/wt1bOE_1_downsampled_narrowpeak.bed"
##
## $wt1bOE_2
## [1] "./Generich_downsampled/wt1bOE_2_downsampled_narrowpeak.bed"
##
## $wt1bOE_3
## [1] "./Generich_downsampled/wt1bOE_3_downsampled_narrowpeak.bed"
```

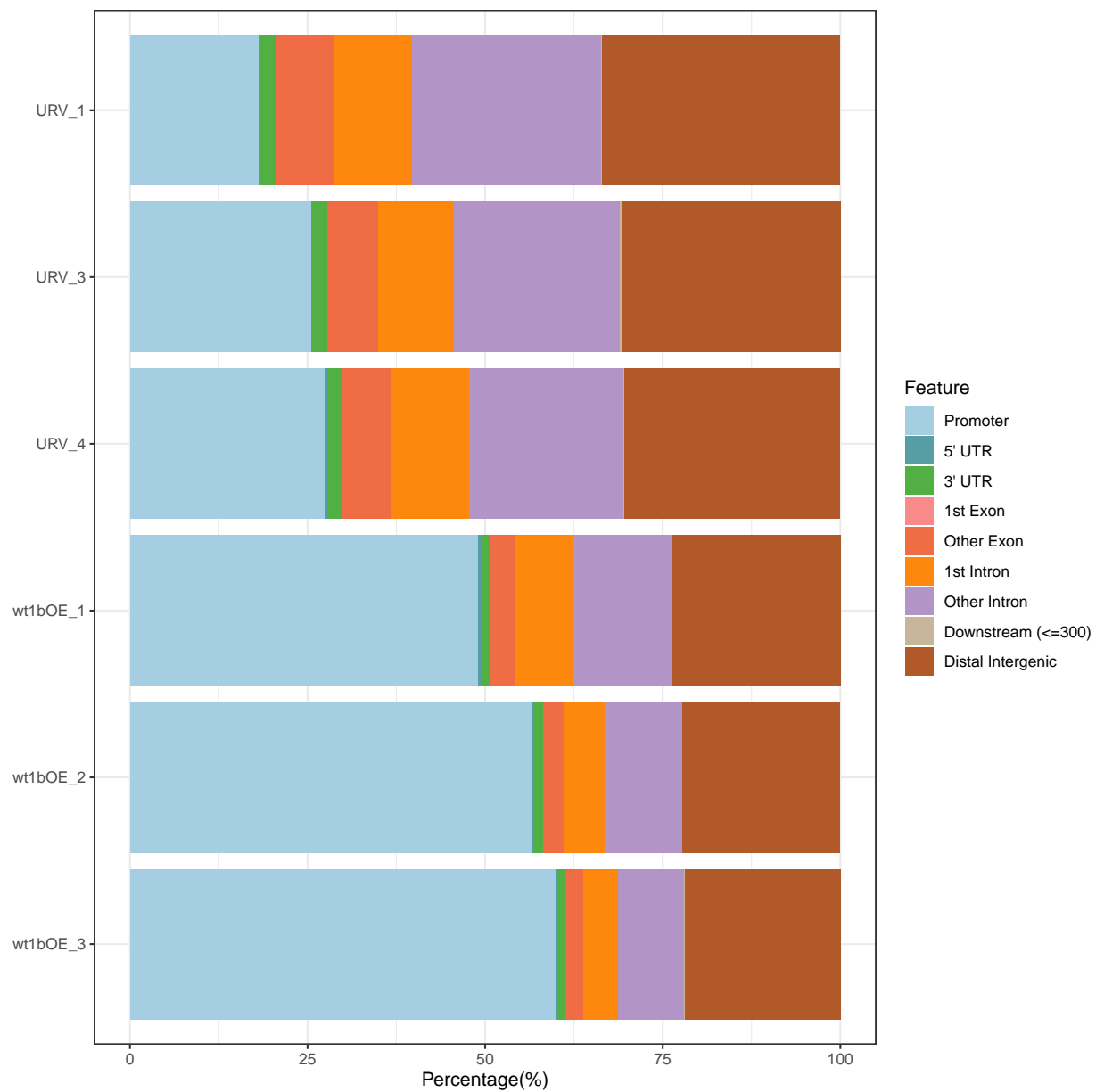
Annotate the peaks

```
peakAnnoList <- lapply(peakfiles_granges_names, annotatePeak, tssRegion=c(-1000, 1000), TxDb=txdb, verbose=FALSE)
```

Check the genomic regions assigned to the peaks (promoter, exon, introns, etc)

```
plotAnnoBar(peakAnnoList)
```

Feature Distribution



```
dev.copy(  
  svg,  
  file = paste0("./results/Peak_Annotation.svg"),  
  width = 10,  
  height = 8  
)
```

```
## svg  
## 3
```

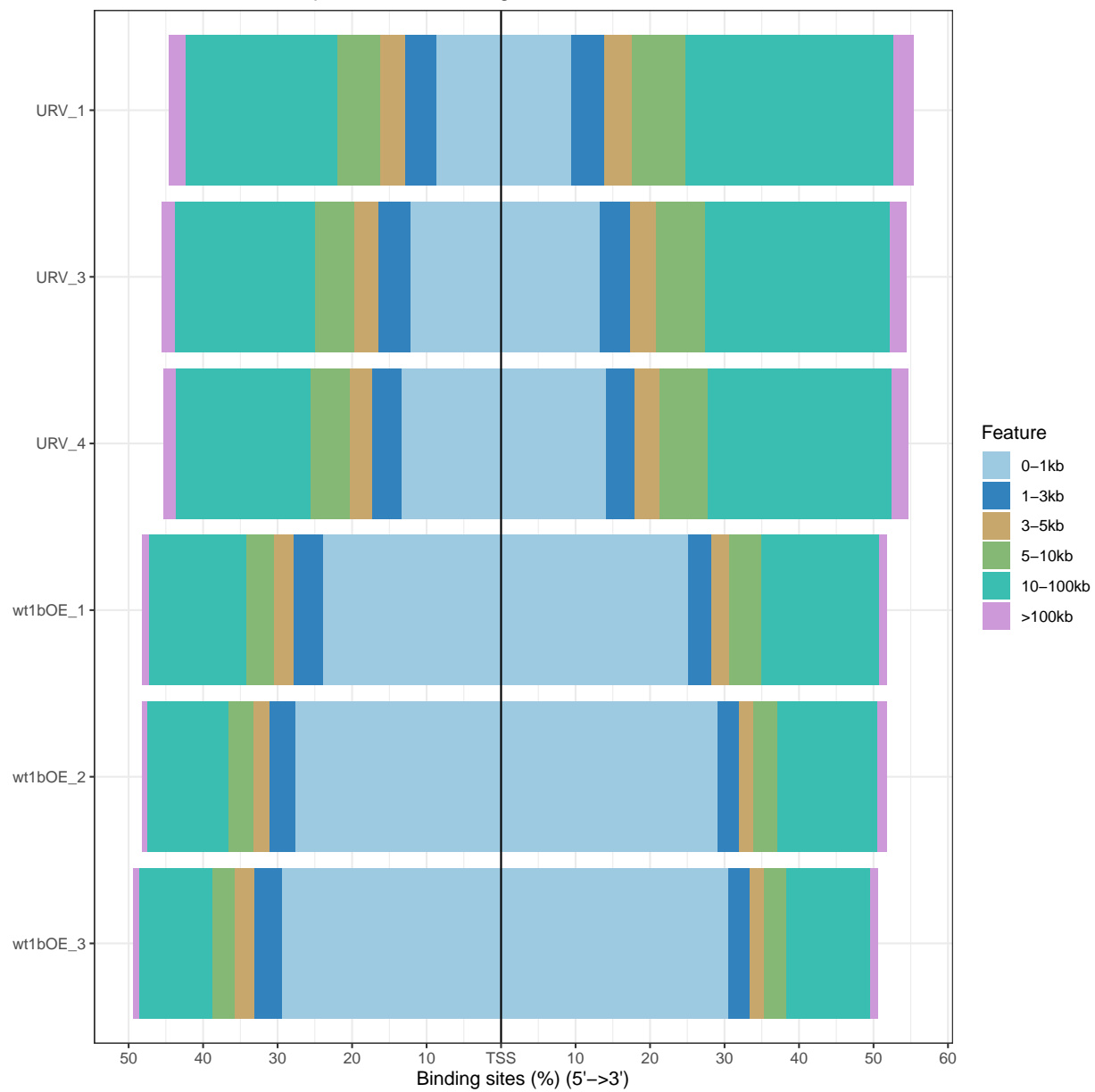
```
dev.off ()
```

```
## pdf  
## 2
```

Check the distance to Transcription Start Site (TSS)

```
plotDistToTSS(peakAnnoList)
```

Distribution of transcription factor-binding loci relative to TSS



```
dev.copy(
  svg,
  file = paste0("./results/Distance_to_TSS.svg"),
  width = 10,
  height = 8
)
```

```
## svg
## 3
```

```
dev.off ()
```

```
## pdf
## 2
```

Compare if the percentage and absolute number of peaks differ in different in Wt1bOE vs Ctrl

```
# Convert peak stats to dataframe
peakAnnoList %>% head
```

```
## $URV_1
## Annotated peaks generated by ChIPseeker
## 30110/30372 peaks were annotated
## Genomic Annotation Summary:
##           Feature   Frequency
## 9           Promoter 18.15675855
## 4             5' UTR  0.22583859
## 3             3' UTR  2.21521089
## 1           1st Exon  0.06310196
## 7           Other Exon 8.06044503
## 2           1st Intron 11.02955829
## 8           Other Intron 26.55596147
## 6 Downstream (<=300) 0.20259050
## 5 Distal Intergenic 33.49053471
##
```



```

## $URV_3
## Annotated peaks generated by ChIPseeker
## 29802/30070 peaks were annotated
## Genomic Annotation Summary:
##           Feature    Frequency
## 9          Promoter 25.46473391
## 4           5' UTR  0.24159452
## 3           3' UTR  2.05019797
## 1          1st Exon  0.07382055
## 7          Other Exon 7.13710489
## 2          1st Intron 10.71069056
## 8          Other Intron 23.29709415
## 6 Downstream (<=300) 0.20803973
## 5 Distal Intergenic 30.81672371
##
## $URV_4
## Annotated peaks generated by ChIPseeker
## 23506/23705 peaks were annotated
## Genomic Annotation Summary:
##           Feature    Frequency
## 9          Promoter 27.47809070
## 4           5' UTR  0.28503361
## 3           3' UTR  2.07181145
## 1          1st Exon  0.05955926
## 7          Other Exon 6.98119629
## 2          1st Intron 10.89509062
## 8          Other Intron 21.67531694
## 6 Downstream (<=300) 0.17867778
## 5 Distal Intergenic 30.37522335
##
## $wt1bOE_1
## Annotated peaks generated by ChIPseeker
## 9383/9452 peaks were annotated
## Genomic Annotation Summary:
##           Feature    Frequency
## 9          Promoter 49.07812000
## 4           5' UTR  0.30906959
## 3           3' UTR  1.22562080
## 1          1st Exon  0.07460301

```

```

## 7      Other Exon  3.48502611
## 2      1st Intron  8.11041245
## 8      Other Intron 14.00404988
## 6 Downstream (<=300) 0.10657572
## 5 Distal Intergenic 23.60652243
##
## $wt1bOE_2
## Annotated peaks generated by ChIPseeker
## 4170/4212 peaks were annotated
## Genomic Annotation Summary:
##           Feature  Frequency
## 9           Promoter 56.69064748
## 4             5' UTR  0.11990408
## 3             3' UTR  1.41486811
## 1             1st Exon 0.02398082
## 7           Other Exon 2.78177458
## 2           1st Intron 5.82733813
## 8           Other Intron 10.83932854
## 6 Downstream (<=300) 0.11990408
## 5 Distal Intergenic 22.18225420
##
## $wt1bOE_3
## Annotated peaks generated by ChIPseeker
## 4142/4188 peaks were annotated
## Genomic Annotation Summary:
##           Feature  Frequency
## 8           Promoter 59.9710285
## 3             5' UTR  0.2172863
## 2             3' UTR  1.1105746
## 6           Other Exon 2.5591502
## 1           1st Intron 4.8768711
## 7           Other Intron 9.2467407
## 5 Downstream (<=300) 0.1448576
## 4 Distal Intergenic 21.8734911

```

```
length(peakAnnoList)
```

```
## [1] 6
```

```

annotstat_df <- data.frame()
for (i in 1:length(peakAnnoList)) {
  tmp_df <- as.data.frame(peakAnnoList[[i]]@annoStat)
  tmp_df$sample <- names(peakAnnoList)[i]
  annotstat_df <- bind_rows(annotstat_df, tmp_df)
}
(annotstat_df)

```

	Feature	Frequency	sample
## 1	Promoter	18.15675855	URV_1
## 2	5' UTR	0.22583859	URV_1
## 3	3' UTR	2.21521089	URV_1
## 4	1st Exon	0.06310196	URV_1
## 5	Other Exon	8.06044503	URV_1
## 6	1st Intron	11.02955829	URV_1
## 7	Other Intron	26.55596147	URV_1
## 8	Downstream (<=300)	0.20259050	URV_1
## 9	Distal Intergenic	33.49053471	URV_1
## 10	Promoter	25.46473391	URV_3
## 11	5' UTR	0.24159452	URV_3
## 12	3' UTR	2.05019797	URV_3
## 13	1st Exon	0.07382055	URV_3
## 14	Other Exon	7.13710489	URV_3
## 15	1st Intron	10.71069056	URV_3
## 16	Other Intron	23.29709415	URV_3
## 17	Downstream (<=300)	0.20803973	URV_3
## 18	Distal Intergenic	30.81672371	URV_3
## 19	Promoter	27.47809070	URV_4
## 20	5' UTR	0.28503361	URV_4
## 21	3' UTR	2.07181145	URV_4
## 22	1st Exon	0.05955926	URV_4
## 23	Other Exon	6.98119629	URV_4
## 24	1st Intron	10.89509062	URV_4
## 25	Other Intron	21.67531694	URV_4
## 26	Downstream (<=300)	0.17867778	URV_4
## 27	Distal Intergenic	30.37522335	URV_4
## 28	Promoter	49.07812000	wt1bOE_1
## 29	5' UTR	0.30906959	wt1bOE_1

```
## 30          3' UTR  1.22562080 wt1bOE_1
## 31          1st Exon 0.07460301 wt1bOE_1
## 32          Other Exon 3.48502611 wt1bOE_1
## 33          1st Intron 8.11041245 wt1bOE_1
## 34          Other Intron 14.00404988 wt1bOE_1
## 35 Downstream (<=300) 0.10657572 wt1bOE_1
## 36 Distal Intergenic 23.60652243 wt1bOE_1
## 37          Promoter 56.69064748 wt1bOE_2
## 38          5' UTR  0.11990408 wt1bOE_2
## 39          3' UTR  1.41486811 wt1bOE_2
## 40          1st Exon 0.02398082 wt1bOE_2
## 41          Other Exon 2.78177458 wt1bOE_2
## 42          1st Intron 5.82733813 wt1bOE_2
## 43          Other Intron 10.83932854 wt1bOE_2
## 44 Downstream (<=300) 0.11990408 wt1bOE_2
## 45 Distal Intergenic 22.18225420 wt1bOE_2
## 46          Promoter 59.97102849 wt1bOE_3
## 47          5' UTR  0.21728634 wt1bOE_3
## 48          3' UTR  1.11057460 wt1bOE_3
## 49          Other Exon 2.55915017 wt1bOE_3
## 50          1st Intron 4.87687108 wt1bOE_3
## 51          Other Intron 9.24674070 wt1bOE_3
## 52 Downstream (<=300) 0.14485756 wt1bOE_3
## 53 Distal Intergenic 21.87349107 wt1bOE_3
```

```
annotstat_df$condition <- NA
annotstat_df$condition[grepl(pattern = "URV",annotstat_df$sample)] <- "Control"
annotstat_df$condition[grepl(pattern = "wt1b",annotstat_df$sample)] <- "WT1bOE"
annotstat_df
```

```
##          Feature  Frequency  sample condition
## 1          Promoter 18.15675855   URV_1   Control
## 2           5' UTR  0.22583859   URV_1   Control
## 3           3' UTR  2.21521089   URV_1   Control
## 4          1st Exon 0.06310196   URV_1   Control
## 5          Other Exon 8.06044503   URV_1   Control
## 6          1st Intron 11.02955829   URV_1   Control
## 7          Other Intron 26.55596147   URV_1   Control
## 8 Downstream (<=300) 0.20259050   URV_1   Control
```

## 9	Distal Intergenic	33.49053471	URV_1	Control
## 10	Promoter	25.46473391	URV_3	Control
## 11	5' UTR	0.24159452	URV_3	Control
## 12	3' UTR	2.05019797	URV_3	Control
## 13	1st Exon	0.07382055	URV_3	Control
## 14	Other Exon	7.13710489	URV_3	Control
## 15	1st Intron	10.71069056	URV_3	Control
## 16	Other Intron	23.29709415	URV_3	Control
## 17	Downstream (<=300)	0.20803973	URV_3	Control
## 18	Distal Intergenic	30.81672371	URV_3	Control
## 19	Promoter	27.47809070	URV_4	Control
## 20	5' UTR	0.28503361	URV_4	Control
## 21	3' UTR	2.07181145	URV_4	Control
## 22	1st Exon	0.05955926	URV_4	Control
## 23	Other Exon	6.98119629	URV_4	Control
## 24	1st Intron	10.89509062	URV_4	Control
## 25	Other Intron	21.67531694	URV_4	Control
## 26	Downstream (<=300)	0.17867778	URV_4	Control
## 27	Distal Intergenic	30.37522335	URV_4	Control
## 28	Promoter	49.07812000	wt1bOE_1	WT1bOE
## 29	5' UTR	0.30906959	wt1bOE_1	WT1bOE
## 30	3' UTR	1.22562080	wt1bOE_1	WT1bOE
## 31	1st Exon	0.07460301	wt1bOE_1	WT1bOE
## 32	Other Exon	3.48502611	wt1bOE_1	WT1bOE
## 33	1st Intron	8.11041245	wt1bOE_1	WT1bOE
## 34	Other Intron	14.00404988	wt1bOE_1	WT1bOE
## 35	Downstream (<=300)	0.10657572	wt1bOE_1	WT1bOE
## 36	Distal Intergenic	23.60652243	wt1bOE_1	WT1bOE
## 37	Promoter	56.69064748	wt1bOE_2	WT1bOE
## 38	5' UTR	0.11990408	wt1bOE_2	WT1bOE
## 39	3' UTR	1.41486811	wt1bOE_2	WT1bOE
## 40	1st Exon	0.02398082	wt1bOE_2	WT1bOE
## 41	Other Exon	2.78177458	wt1bOE_2	WT1bOE
## 42	1st Intron	5.82733813	wt1bOE_2	WT1bOE
## 43	Other Intron	10.83932854	wt1bOE_2	WT1bOE
## 44	Downstream (<=300)	0.11990408	wt1bOE_2	WT1bOE
## 45	Distal Intergenic	22.18225420	wt1bOE_2	WT1bOE
## 46	Promoter	59.97102849	wt1bOE_3	WT1bOE
## 47	5' UTR	0.21728634	wt1bOE_3	WT1bOE

## 48	3' UTR	1.11057460	wt1bOE_3	WT1bOE
## 49	Other Exon	2.55915017	wt1bOE_3	WT1bOE
## 50	1st Intron	4.87687108	wt1bOE_3	WT1bOE
## 51	Other Intron	9.24674070	wt1bOE_3	WT1bOE
## 52	Downstream (<=300)	0.14485756	wt1bOE_3	WT1bOE
## 53	Distal Intergenic	21.87349107	wt1bOE_3	WT1bOE

annotstat_df

##	Feature	Frequency	sample	condition
## 1	Promoter	18.15675855	URV_1	Control
## 2	5' UTR	0.22583859	URV_1	Control
## 3	3' UTR	2.21521089	URV_1	Control
## 4	1st Exon	0.06310196	URV_1	Control
## 5	Other Exon	8.06044503	URV_1	Control
## 6	1st Intron	11.02955829	URV_1	Control
## 7	Other Intron	26.55596147	URV_1	Control
## 8	Downstream (<=300)	0.20259050	URV_1	Control
## 9	Distal Intergenic	33.49053471	URV_1	Control
## 10	Promoter	25.46473391	URV_3	Control
## 11	5' UTR	0.24159452	URV_3	Control
## 12	3' UTR	2.05019797	URV_3	Control
## 13	1st Exon	0.07382055	URV_3	Control
## 14	Other Exon	7.13710489	URV_3	Control
## 15	1st Intron	10.71069056	URV_3	Control
## 16	Other Intron	23.29709415	URV_3	Control
## 17	Downstream (<=300)	0.20803973	URV_3	Control
## 18	Distal Intergenic	30.81672371	URV_3	Control
## 19	Promoter	27.47809070	URV_4	Control
## 20	5' UTR	0.28503361	URV_4	Control
## 21	3' UTR	2.07181145	URV_4	Control
## 22	1st Exon	0.05955926	URV_4	Control
## 23	Other Exon	6.98119629	URV_4	Control
## 24	1st Intron	10.89509062	URV_4	Control
## 25	Other Intron	21.67531694	URV_4	Control
## 26	Downstream (<=300)	0.17867778	URV_4	Control
## 27	Distal Intergenic	30.37522335	URV_4	Control
## 28	Promoter	49.07812000	wt1bOE_1	WT1bOE
## 29	5' UTR	0.30906959	wt1bOE_1	WT1bOE

```
## 30      3' UTR  1.22562080 wt1bOE_1  WT1bOE
## 31      1st Exon 0.07460301 wt1bOE_1  WT1bOE
## 32      Other Exon 3.48502611 wt1bOE_1  WT1bOE
## 33      1st Intron 8.11041245 wt1bOE_1  WT1bOE
## 34      Other Intron 14.00404988 wt1bOE_1  WT1bOE
## 35 Downstream (<=300) 0.10657572 wt1bOE_1  WT1bOE
## 36 Distal Intergenic 23.60652243 wt1bOE_1  WT1bOE
## 37      Promoter 56.69064748 wt1bOE_2  WT1bOE
## 38      5' UTR  0.11990408 wt1bOE_2  WT1bOE
## 39      3' UTR  1.41486811 wt1bOE_2  WT1bOE
## 40      1st Exon 0.02398082 wt1bOE_2  WT1bOE
## 41      Other Exon 2.78177458 wt1bOE_2  WT1bOE
## 42      1st Intron 5.82733813 wt1bOE_2  WT1bOE
## 43      Other Intron 10.83932854 wt1bOE_2  WT1bOE
## 44 Downstream (<=300) 0.11990408 wt1bOE_2  WT1bOE
## 45 Distal Intergenic 22.18225420 wt1bOE_2  WT1bOE
## 46      Promoter 59.97102849 wt1bOE_3  WT1bOE
## 47      5' UTR  0.21728634 wt1bOE_3  WT1bOE
## 48      3' UTR  1.11057460 wt1bOE_3  WT1bOE
## 49      Other Exon 2.55915017 wt1bOE_3  WT1bOE
## 50      1st Intron 4.87687108 wt1bOE_3  WT1bOE
## 51      Other Intron 9.24674070 wt1bOE_3  WT1bOE
## 52 Downstream (<=300) 0.14485756 wt1bOE_3  WT1bOE
## 53 Distal Intergenic 21.87349107 wt1bOE_3  WT1bOE
```

```
#some samples do not have 1st exon,. this needs to be added manually to help conduct ttest
```

```
add_df_rows_exon <- data.frame(Feature= c("1st Exon","1st Exon"), Frequency= c(0,0),
                               sample= c("wt1bOE_2", "wt1bOE_3"), condition= c("WT1bOE", "WT1bOE"))
annotstat_df <- rbind(annotstat_df, add_df_rows_exon)

annotstat.ttest <- annotstat_df %>%
  group_by(Feature) %>%
  t_test(Frequency ~ condition) %>%
  add_significance()
annotstat.ttest
```

```
## # A tibble: 9 x 10
##   Feature    .y.  group1 group2    n1    n2 statistic    df      p p.signif
```

```
## * <fct>      <chr>  <chr>  <chr>  <int> <int>      <dbl> <dbl>  <dbl> <chr>
## 1 Promoter   Freque~ Contr~ WT1bOE    3    3    -7.35  3.93  1.95e-3 **
## 2 5' UTR     Freque~ Contr~ WT1bOE    3    3     0.617 2.42  5.91e-1 ns
## 3 3' UTR     Freque~ Contr~ WT1bOE    3    3     8.39  3.22  2.7 e-3 **
## 4 1st Exon   Freque~ Contr~ WT1bOE    3    4     2.26  3.35  1 e-1 ns
## 5 Other Exon Freque~ Contr~ WT1bOE    3    3    10.2   3.87  6.23e-4 ***
## 6 1st Intron Freque~ Contr~ WT1bOE    3    3     4.78  2.04  3.96e-2 *
## 7 Other Intr~ Freque~ Contr~ WT1bOE    3    3     6.23  4.00  3.39e-3 **
## 8 Downstream~ Freque~ Contr~ WT1bOE    3    3     5.05  3.82  8.16e-3 **
## 9 Distal Int~ Freque~ Contr~ WT1bOE    3    3     8.11  3.10  3.45e-3 **
```

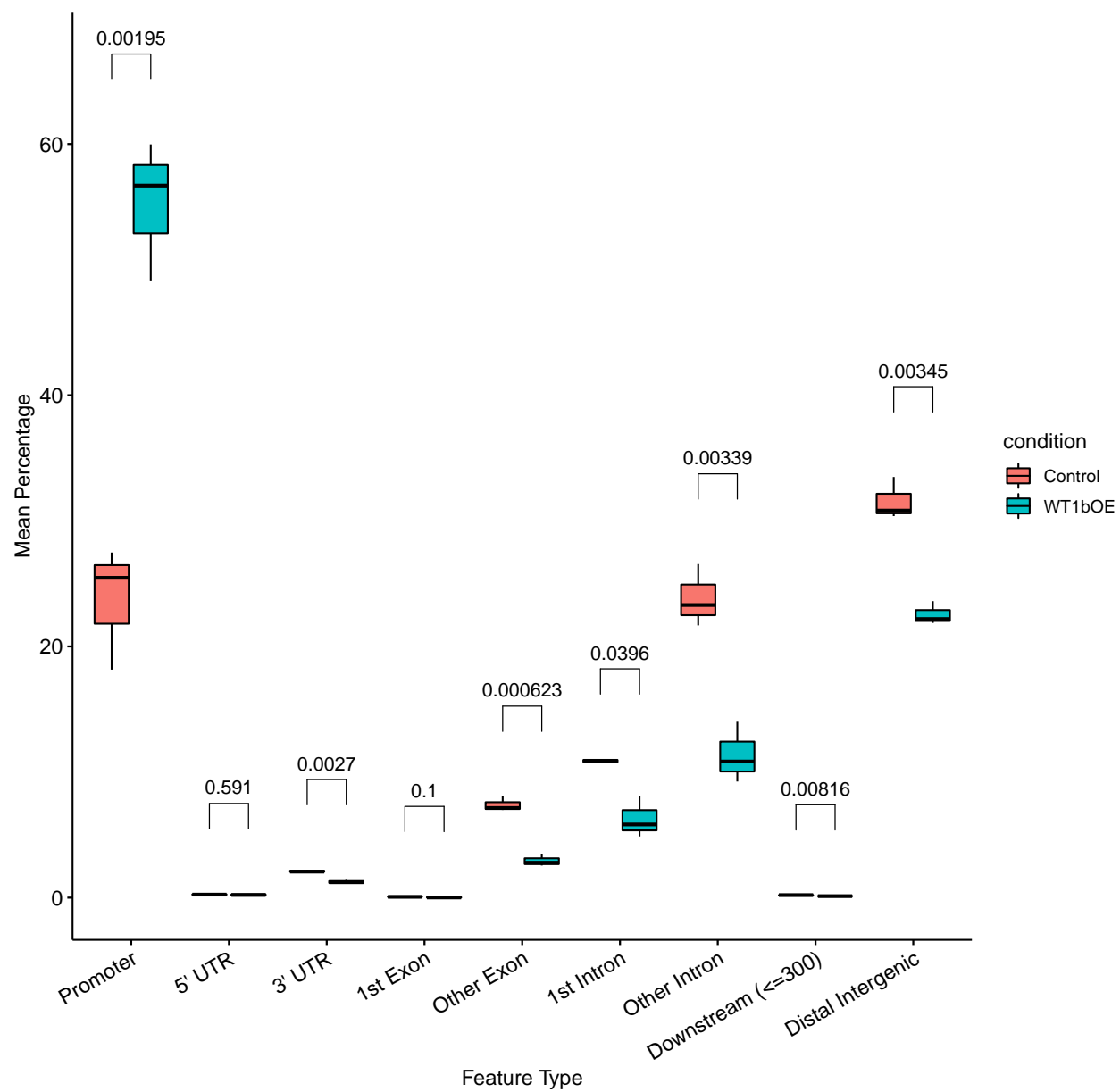
```
annotstat.ttest <- annotstat.ttest %>% add_xy_position(x = "Feature")
annotstat.ttest
```

```
## # A tibble: 9 x 15
##   Feature      .y.    group1 group2    n1    n2 statistic    df      p p.signif
##   <fct>      <chr>  <chr>  <chr>  <int> <int>      <dbl> <dbl>  <dbl> <chr>
## 1 Promoter   Freque~ Contr~ WT1bOE    3    3    -7.35  3.93  1.95e-3 **
## 2 5' UTR     Freque~ Contr~ WT1bOE    3    3     0.617 2.42  5.91e-1 ns
## 3 3' UTR     Freque~ Contr~ WT1bOE    3    3     8.39  3.22  2.7 e-3 **
## 4 1st Exon   Freque~ Contr~ WT1bOE    3    4     2.26  3.35  1 e-1 ns
## 5 Other Exon Freque~ Contr~ WT1bOE    3    3    10.2   3.87  6.23e-4 ***
## 6 1st Intron Freque~ Contr~ WT1bOE    3    3     4.78  2.04  3.96e-2 *
## 7 Other Intr~ Freque~ Contr~ WT1bOE    3    3     6.23  4.00  3.39e-3 **
## 8 Downstream~ Freque~ Contr~ WT1bOE    3    3     5.05  3.82  8.16e-3 **
## 9 Distal Int~ Freque~ Contr~ WT1bOE    3    3     8.11  3.10  3.45e-3 **
## # ... with 5 more variables: y.position <dbl>, groups <named list>, x <dbl>,
## #   xmin <dbl>, xmax <dbl>
```

```
q <- ggboxplot(
  annotstat_df, x = "Feature", y = "Frequency",
  fill = "condition", legend= "right") +
  theme(axis.text.x = element_text(angle = 30, hjust = 1))+
  stat_pvalue_manual(annotstat.ttest, label = "p")+
  labs(x= "Feature Type", y= "Mean Percentage", title = "Various Genomic features in Control vs WT1bOE")

print(q)
```


Various Genomic features in Control vs WT1bOE



Write the annotated peaks for each sample

```
for(i in 1:length(peakfiles_granges_names)){  
  # for(i in 1:1){  
    tmp_df <- as.data.frame(peakAnnoList[[i]]@anno)  
    write.csv(tmp_df, file=paste0("./annotated_peaks/",  
                                  sample_names[i], "_annotations", ".csv"),  
              row.names=F, col.names=T)  
  }  
}
```

```
## Warning in write.csv(tmp_df, file = paste0("./annotated_peaks/",  
## sample_names[i], : attempt to set 'col.names' ignored
```

```
## Warning in write.csv(tmp_df, file = paste0("./annotated_peaks/",  
## sample_names[i], : attempt to set 'col.names' ignored
```

```
## Warning in write.csv(tmp_df, file = paste0("./annotated_peaks/",  
## sample_names[i], : attempt to set 'col.names' ignored
```

```
## Warning in write.csv(tmp_df, file = paste0("./annotated_peaks/",  
## sample_names[i], : attempt to set 'col.names' ignored
```

```
## Warning in write.csv(tmp_df, file = paste0("./annotated_peaks/",  
## sample_names[i], : attempt to set 'col.names' ignored
```

```
## Warning in write.csv(tmp_df, file = paste0("./annotated_peaks/",  
## sample_names[i], : attempt to set 'col.names' ignored
```

Check the absolute counts of annotated peaks

```
read_peaks <- function(filename){  
  print(filename)  
  df <- read.csv(paste0("./annotated_peaks/", filename))  
  df$sample_name <- gsub("_annotations.csv", "", filename)  
  if((grepl(pattern = "URV", filename))){df$condition <- "Control"  
  } else if(grepl(pattern = "wt1", filename)) {df$condition <- "WT1bOE"}  
}
```

```
df$annotation <- gsub("\\(EN.*", "", df$annotation)
df
}
```

```
peaks_files <- list.files(path = "./annotated_peaks/", pattern = "csv")
```

```
combine_peaks_df<- lapply(X = peaks_files, FUN = read_peaks) %>% bind_rows()
```

```
## [1] "URV_1_annotations.csv"
## [1] "URV_3_annotations.csv"
## [1] "URV_4_annotations.csv"
## [1] "wt1bOE_1_annotations.csv"
## [1] "wt1bOE_2_annotations.csv"
## [1] "wt1bOE_3_annotations.csv"
```

```
head(combine_peaks_df)
```

```
##   seqnames start   end width strand    V4   V5 V6 annotation geneChr geneStart
## 1      1    6679   6778   100      * peak_0 1000 .   Promoter    931    6642
## 2      1   21229  22646  1418      * peak_1  664 .     Exon     931   18786
## 3      1   27036  27404   369      * peak_2  869 .   Promoter    931   27690
## 4      1   27560  27873   314      * peak_3 1000 .   Promoter    931   27690
## 5      1   36576  36771   196      * peak_4 1000 .   Promoter    931   36727
## 6      1   36888  37124   237      * peak_5 1000 .   Promoter    931   36727
##   geneEnd geneLength geneStrand      geneId      transcriptId
## 1    7335      694          2 ENSDARG00000099104 ENSDART000000158290
## 2   23173     4388          1 ENSDARG000000102097 ENSDART000000161190
## 3   34330     6641          1 ENSDARG000000099640 ENSDART000000162928
## 4   34330     6641          1 ENSDARG000000099640 ENSDART000000162928
## 5   38446     1720          1 ENSDARG000000104071 ENSDART000000166557
## 6   38446     1720          1 ENSDARG000000104071 ENSDART000000166557
##   distanceToTSS sample_name condition
## 1          557      URV_1   Control
## 2         2443      URV_1   Control
## 3         -286      URV_1   Control
```

```
## 4          0      URV_1  Control
## 5          0      URV_1  Control
## 6        161      URV_1  Control
```

```
combine_peaks_df_counts <- combine_peaks_df %>% group_by(condition, sample_name, annotation) %>%
  summarise(counts= n()) %>% group_by(condition, annotation) %>%
  mutate(sd=sd(counts), mean=mean(counts)) %>%
  arrange(desc(counts))
```

'summarise()' has grouped output by 'condition', 'sample_name'. You can override using the '.groups' argument.

```
combine_peaks_df_counts$annotation <- factor(combine_peaks_df_counts$annotation,
                                              levels = unique(combine_peaks_df_counts$annotation))

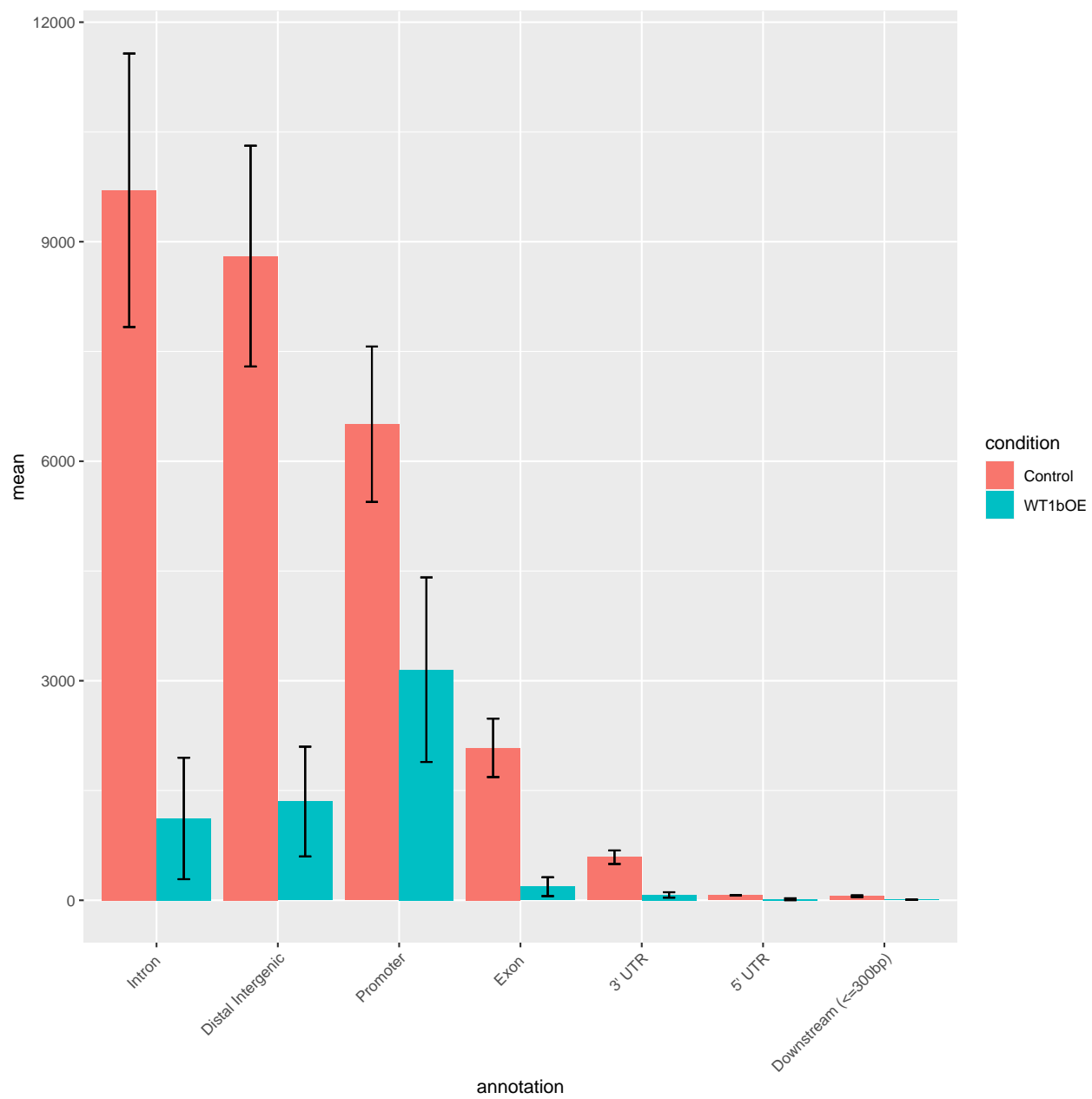
head(combine_peaks_df_counts)
```

```
## # A tibble: 6 x 6
## # Groups:   condition, annotation [3]
##   condition sample_name annotation counts    sd  mean
##   <chr>      <chr>      <fct>    <int> <dbl> <dbl>
## 1 Control  URV_1      "Intron "      11317 1868. 9703.
## 2 Control  URV_3      "Intron "      10135 1868. 9703.
## 3 Control  URV_1      "Distal Intergenic" 10084 1509. 8803.
## 4 Control  URV_3      "Distal Intergenic"  9184 1509. 8803.
## 5 Control  URV_4      "Intron "       7656 1868. 9703.
## 6 Control  URV_3      "Promoter"      7589 1062. 6505
```

write csv with total peak counts

```
write.csv(combine_peaks_df_counts, "./results/total_peak_counts.csv")
```

```
q <- ggplot(data= combine_peaks_df_counts, aes(x=annotation, y= mean, fill= condition))+
  geom_bar(stat = "identity", position = "dodge")+
  geom_errorbar(aes(ymin= mean-sd, ymax= mean+sd),width=.2,
               position=position_dodge(.9))+
  theme(axis.text.x = element_text(angle= 45, hjust=1))
print(q)
```



```
dev.copy(
  svg,
  file = paste0("./results/Total_peak_counts_annotations.svg"),
  width = 10,
  height = 8
)
```

```
## svg
## 3
```

```
dev.off ()
```

```
## pdf
## 2
```

Annotation of the differential peaks and representation in graphs (promoters, exons, introns, etc)

```
peakfiles_granges_diff <- "./results/diff_peaks_downsample_generich_3_samples_pval_norm.bed"
read_peaks_diff <- readPeakFile(peakfiles_granges_diff)
read_peaks_diff
```

```
## GRanges object with 1453 ranges and 13 metadata columns:
##      seqnames      ranges strand |      V4      V5      V6
##      <Rle>      <IRanges> <Rle> | <numeric> <numeric> <numeric>
##      [1]          2 14006003-14006402   * |  4.25687  2.66918  4.99438
##      [2]         13 50200080-50200479   * |  3.94954  2.40860  4.67754
##      [3]         17 53383460-53383859   * |  9.14931  8.82691  9.41266
##      [4]         13 52114476-52114875   * |  4.57885  3.27336  5.25278
##      [5]         21  3470697-3471096   * |  4.31855  2.89841  5.02017
##      ...      ...      ...      ... |  ...      ...      ...
## [1449] KN149878.1      2091-2490   * |  2.53777  1.53624  3.12324
## [1450]          3 16556864-16557263   * |  2.56912  1.56049  3.15695
## [1451]          6  9282025-9282424   * |  2.73867  1.89651  3.26693
## [1452]         22  8476049-8476705   * |  2.92763  2.09783  3.45109
## [1453]          5 11290583-11290982   * |  2.48841  1.39384  3.10357
```

```
##          V7          V8          V9          V10          V11          V12
##      <numeric> <numeric> <numeric> <character> <character> <character>
##      [1] -1.174633  1.04e-05 0.0569925          +          -          +
##      [2] -1.135109  1.82e-05 0.0569925          -          +          -
##      [3] -0.538733  2.16e-05 0.0569925          +          +          +
##      [4] -1.090197  2.99e-05 0.0592566          +          +          -
##      [5] -1.077603  3.96e-05 0.0602033          +          -          -
##      ...      ...      ...      ...      ...      ...
## [1449] -0.420671  0.0497316  1.000000          -          -          -
## [1450] -0.416365  0.0498947  1.000000          -          -          -
## [1451] -0.441763  0.0499529  1.000000          -          -          -
## [1452] -0.445844  0.0499544  0.421683          -          -          -
## [1453] -0.404680  0.0499974  1.000000          -          -          -
##          V13          V14          V15          V16
##      <character> <character> <character> <character>
##      [1]          +          +          +      peak_1
##      [2]          +          +          +      peak_2
##      [3]          +          +          +      peak_3
##      [4]          +          +          +      peak_4
##      [5]          +          +          +      peak_5
##      ...      ...      ...      ...
## [1449]          +          +          -      peak_1449
## [1450]          -          +          +      peak_1450
## [1451]          +          -          +      peak_1451
## [1452]          +          +          -      peak_1452
## [1453]          -          +          +      peak_1453
## -----
##      seqinfo: 49 sequences from an unspecified genome; no seqlengths
```

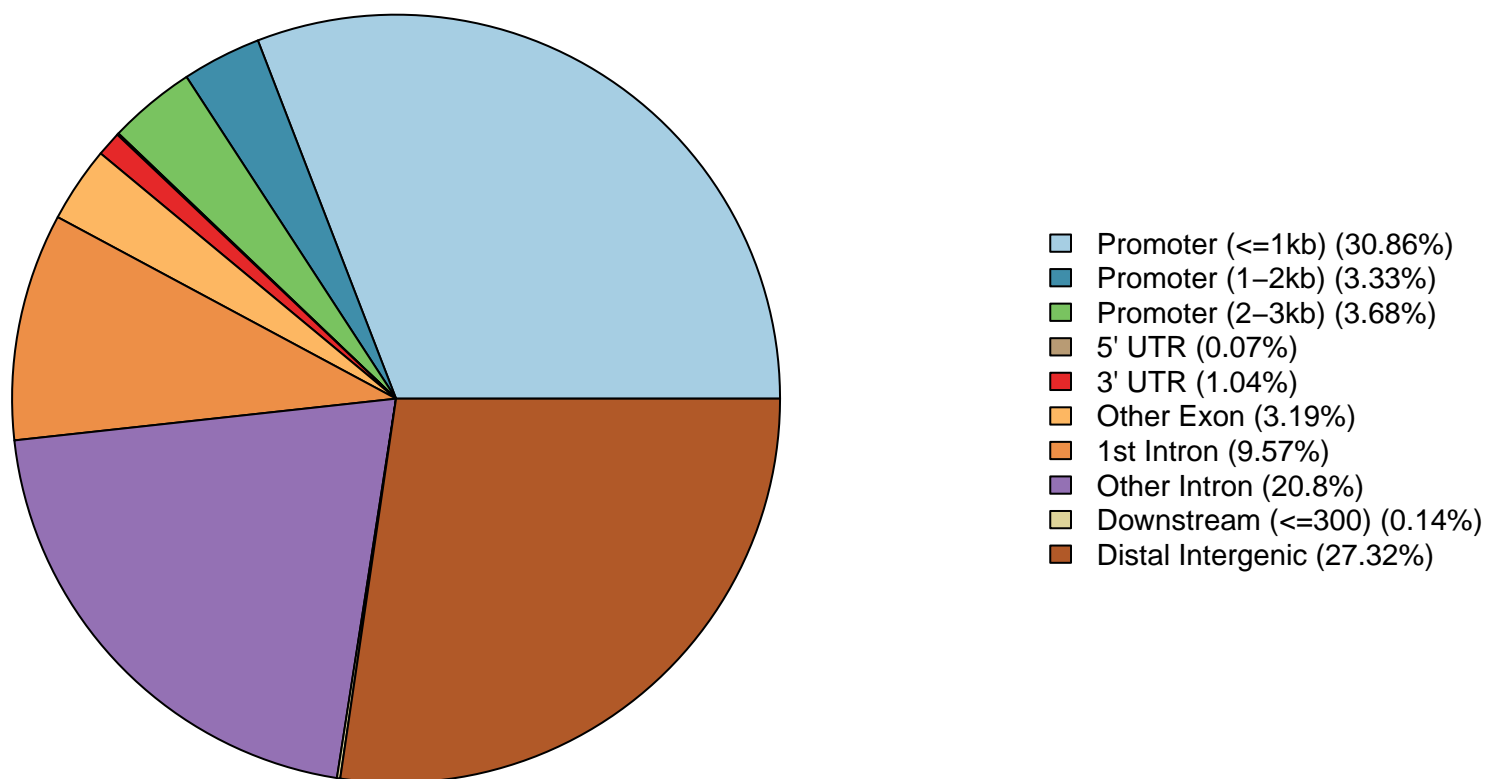
```
peakAnno_diff <- annotatePeak(read_peaks_diff, TxDb = txdb)
```

```
## >> preparing features information...      2021-07-06 10:00:38
## >> identifying nearest features...      2021-07-06 10:00:38
## >> calculating distance from peak to TSS... 2021-07-06 10:00:38
## >> assigning genomic annotation...      2021-07-06 10:00:38
## >> assigning chromosome lengths      2021-07-06 10:00:39
## >> done...      2021-07-06 10:00:39
```

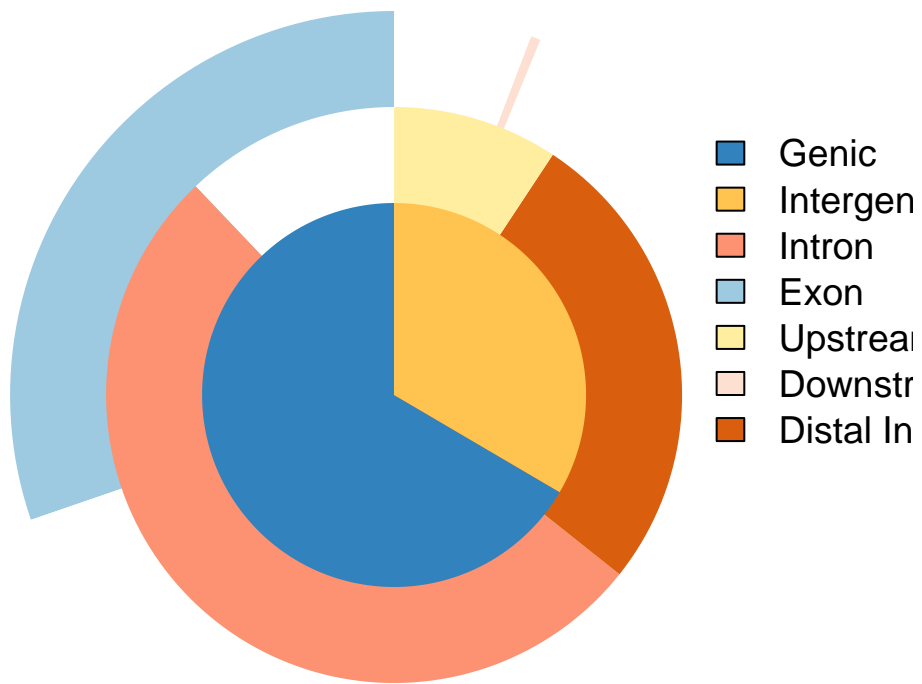
```
peakAnno_diff
```

```
## Annotated peaks generated by ChIPseeker
## 1442/1453 peaks were annotated
## Genomic Annotation Summary:
##           Feature    Frequency
## 8 Promoter (<=1kb) 30.85991678
## 9 Promoter (1-2kb) 3.32871012
## 10 Promoter (2-3kb) 3.67545076
## 3          5' UTR 0.06934813
## 2          3' UTR 1.04022191
## 6 Other Exon 3.19001387
## 1      1st Intron 9.57004161
## 7 Other Intron 20.80443828
## 5 Downstream (<=300) 0.13869626
## 4 Distal Intergenic 27.32316227
```

```
#plot the pie chart with all the info on peak distribution
plotAnnoPie(peakAnno_diff)
```

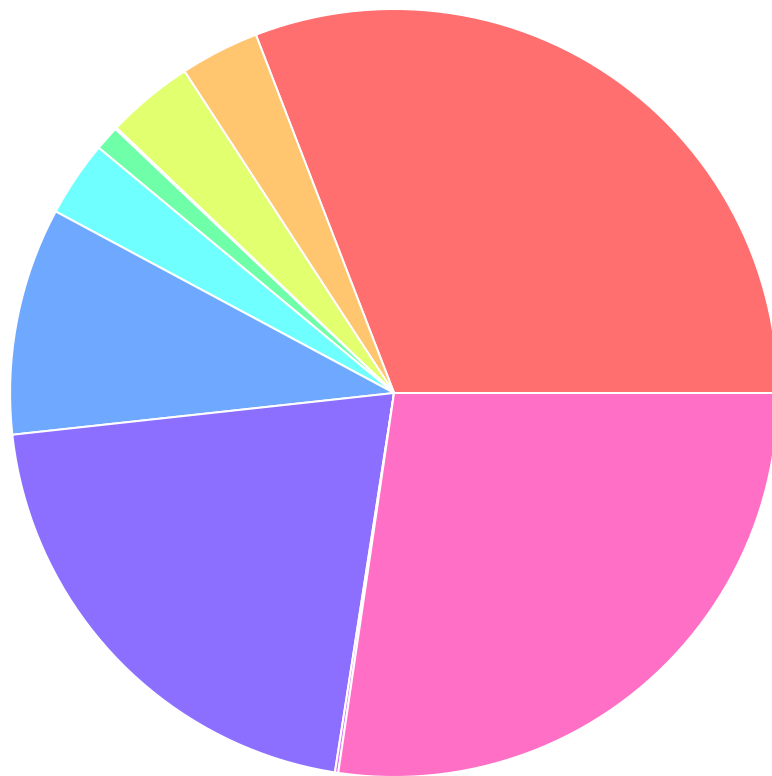



```
vennpie(peakAnno_diff)
```



```
plotAnnoPie(peakAnno_diff, col = rainbow(10, alpha = 0.7, s = 0.8), border = "white")
```

```
## Warning in if (is.na(col)) {: the condition has length > 1 and only the first
## element will be used
```



- Promoter (<=1kb) (30.86%)
- Promoter (1-2kb) (3.33%)
- Promoter (2-3kb) (3.68%)
- 5' UTR (0.07%)
- 3' UTR (1.04%)
- Other Exon (3.19%)
- 1st Intron (9.57%)
- Other Intron (20.8%)
- Downstream (<=300) (0.14%)
- Distal Intergenic (27.32%)

```
# display.brewer.all()
dev.copy(
  svg,
  file = paste0("./results/Piechart_rainbow.svg"),
  width = 10,
  height = 8
)
```

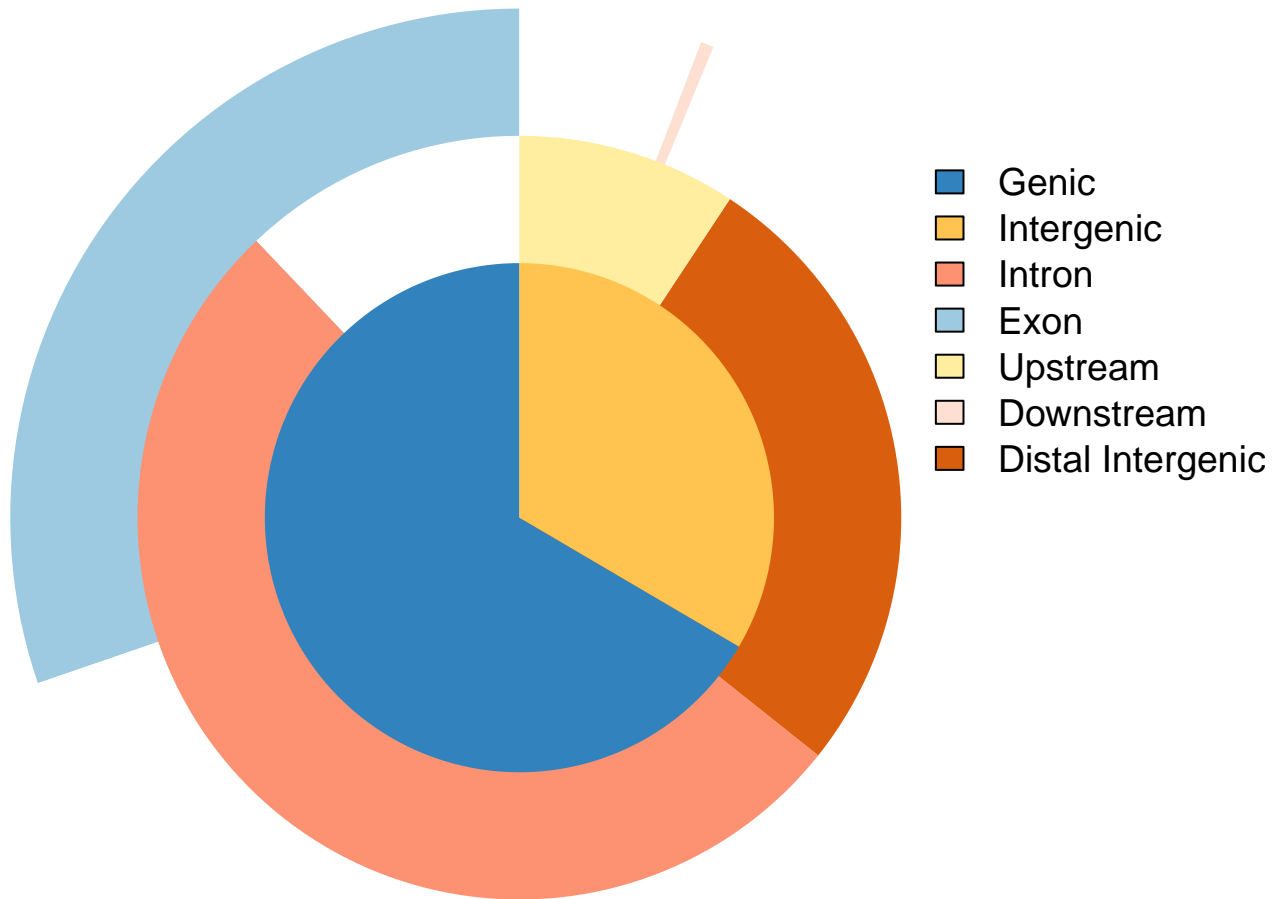
```
## svg
## 3
```

```
dev.off ()
```

```
## pdf
```

```
## 2
```

```
vennpie(peakAnno_diff)
```



Correct the names of the annotated dataframe of peaks

```
peakAnno_diff@anno
```

```
## GRanges object with 1442 ranges and 22 metadata columns:
```

##	seqnames	ranges	strand	V4	V5	V6	
##	<Rle>	<IRanges>	<Rle>	<numeric>	<numeric>	<numeric>	
##	[1]	2 14006003-14006402	*	4.25687	2.66918	4.99438	
##	[2]	13 50200080-50200479	*	3.94954	2.40860	4.67754	
##	[3]	17 53383460-53383859	*	9.14931	8.82691	9.41266	
##	[4]	13 52114476-52114875	*	4.57885	3.27336	5.25278	
##	[5]	21 3470697-3471096	*	4.31855	2.89841	5.02017	
##	
##	[1438]	7 30770152-30770551	*	2.50757	1.54597	3.07950	
##	[1439]	3 16556864-16557263	*	2.56912	1.56049	3.15695	
##	[1440]	6 9282025-9282424	*	2.73867	1.89651	3.26693	
##	[1441]	22 8476049-8476705	*	2.92763	2.09783	3.45109	
##	[1442]	5 11290583-11290982	*	2.48841	1.39384	3.10357	
##	V7	V8	V9	V10	V11	V12	
##	<numeric>	<numeric>	<numeric>	<character>	<character>	<character>	
##	[1]	-1.174633	1.04e-05	0.0569925	+	-	+
##	[2]	-1.135109	1.82e-05	0.0569925	-	+	-
##	[3]	-0.538733	2.16e-05	0.0569925	+	+	+
##	[4]	-1.090197	2.99e-05	0.0592566	+	+	-
##	[5]	-1.077603	3.96e-05	0.0602033	+	-	-
##	
##	[1438]	-0.424604	0.0497078	1.000000	-	-	-
##	[1439]	-0.416365	0.0498947	1.000000	-	-	-
##	[1440]	-0.441763	0.0499529	1.000000	-	-	-
##	[1441]	-0.445844	0.0499544	0.421683	-	-	-
##	[1442]	-0.404680	0.0499974	1.000000	-	-	-
##	V13	V14	V15	V16	annotation		
##	<character>	<character>	<character>	<character>	<character>		
##	[1]	+	+	+	peak_1 Distal Intergenic		
##	[2]	+	+	+	peak_2 Promoter (<=1kb)		
##	[3]	+	+	+	peak_3 Distal Intergenic		
##	[4]	+	+	+	peak_4 Promoter (2-3kb)		
##	[5]	+	+	+	peak_5 Distal Intergenic		
##		
##	[1438]	+	-	+	peak_1448 3' UTR		

```

## [1439] - + + peak_1450 Distal Intergenic
## [1440] + - + peak_1451 Promoter (<=1kb)
## [1441] + + - peak_1452 Distal Intergenic
## [1442] - + + peak_1453 Promoter (<=1kb)
##      geneChr geneStart geneEnd geneLength geneStrand geneId
##      <integer> <integer> <integer> <integer> <integer> <character>
## [1]      932  14007938  14009528      1591      2 ENSDARG00000098037
## [2]      943  50190664  50200348      9685      2 ENSDARG00000052396
## [3]      947  53388239  53393653      5415      2 ENSDARG000000117519
## [4]      943  52107945  52111988      4044      2 ENSDARG000000117319
## [5]      951   3419459   3452683     33225      2 ENSDARG00000016858
## ...      ...      ...      ...      ...      ...
## [1438]     937   30772991   30779575      6585      2 ENSDARG00000012495
## [1439]     933   16566361   16590254     23894      1 ENSDARG00000007129
## [1440]     936    9268857    9282080     13224      2 ENSDARG00000098734
## [1441]     952    8489400    8491260      1861      1 ENSDARG000000112069
## [1442]     935   11290851   11361593     70743      1 ENSDARG000000114825
##      transcriptId distanceToTSS
##      <character> <numeric>
## [1] ENSDART00000155813      3126
## [2] ENSDART00000038391        0
## [3] ENSDART00000194534      9794
## [4] ENSDART00000194638     -2488
## [5] ENSDART00000009740    -18014
## ...      ...      ...
## [1438] ENSDART00000004782      9024
## [1439] ENSDART00000132869    -9098
## [1440] ENSDART00000159506        0
## [1441] ENSDART00000129119    -12695
## [1442] ENSDART00000180408        0
## -----
## seqinfo: 49 sequences from an unspecified genome

```

```

peakAnno_diff_df <- as.data.frame(peakAnno_diff@anno)
head(peakAnno_diff_df)

```

```

##  seqnames      start      end width strand      V4      V5      V6      V7
## 1         2 14006003 14006402   400      * 4.256871 2.669184 4.994381 -1.1746329

```

```
## 2      13 50200080 50200479 400      * 3.949543 2.408604 4.677540 -1.1351088
## 3      17 53383460 53383859 400      * 9.149312 8.826909 9.412664 -0.5387331
## 4      13 52114476 52114875 400      * 4.578848 3.273357 5.252782 -1.0901966
## 5      21 3470697 3471096 400      * 4.318552 2.898409 5.020167 -1.0776029
## 6       4 77761293 77762031 739      * 3.578587 1.695711 4.368405 -1.0402813
##          V8          V9 V10 V11 V12 V13 V14 V15      V16          annotation geneChr
## 1 1.04e-05 0.05699248 + - + + + + peak_1 Distal Intergenic 932
## 2 1.82e-05 0.05699248 - + - + + + + peak_2 Promoter (<=1kb) 943
## 3 2.16e-05 0.05699248 + + + + + + + peak_3 Distal Intergenic 947
## 4 2.99e-05 0.05925655 + + - + + + + peak_4 Promoter (2-3kb) 943
## 5 3.96e-05 0.06020326 + - - + + + + peak_5 Distal Intergenic 951
## 6 4.56e-05 0.06020326 - - - + + + + peak_6 Promoter (<=1kb) 934
## geneStart geneEnd geneLength geneStrand          geneId
## 1 14007938 14009528      1591          2 ENSDARG00000098037
## 2 50190664 50200348      9685          2 ENSDARG00000052396
## 3 53388239 53393653      5415          2 ENSDARG00000117519
## 4 52107945 52111988      4044          2 ENSDARG00000117319
## 5 3419459 3452683      33225          2 ENSDARG00000016858
## 6 77761901 77795612      33712          1 ENSDARG00000104380
## transcriptId distanceToTSS
## 1 ENSDART00000155813      3126
## 2 ENSDART00000038391          0
## 3 ENSDART00000194534      9794
## 4 ENSDART00000194638     -2488
## 5 ENSDART00000009740    -18014
## 6 ENSDART00000159502          0
```

```
colnames(peakAnno_diff_df)[7:12] <- colnames(diff_peaks_pval_norm)[5:10]
peakAnno_diff_df %>% head()
```

```
## seqnames      start      end width strand      V4 Conc_WT1bOE Conc_Ctrl
## 1      2 14006003 14006402 400      * 4.256871 2.669184 4.994381
## 2     13 50200080 50200479 400      * 3.949543 2.408604 4.677540
## 3     17 53383460 53383859 400      * 9.149312 8.826909 9.412664
## 4     13 52114476 52114875 400      * 4.578848 3.273357 5.252782
## 5     21 3470697 3471096 400      * 4.318552 2.898409 5.020167
## 6      4 77761293 77762031 739      * 3.578587 1.695711 4.368405
##      Fold p.value      FDR wt1bOE_1 V11 V12 V13 V14 V15      V16
```

```
## 1 -1.1746329 1.04e-05 0.05699248      + - + + + + peak_1
## 2 -1.1351088 1.82e-05 0.05699248      - + - + + + peak_2
## 3 -0.5387331 2.16e-05 0.05699248      + + + + + + peak_3
## 4 -1.0901966 2.99e-05 0.05925655      + + - + + + peak_4
## 5 -1.0776029 3.96e-05 0.06020326      + - - + + + peak_5
## 6 -1.0402813 4.56e-05 0.06020326      - - - + + + peak_6
##          annotation geneChr geneStart geneEnd geneLength geneStrand
## 1 Distal Intergenic      932  14007938 14009528      1591      2
## 2 Promoter (<=1kb)      943  50190664 50200348      9685      2
## 3 Distal Intergenic      947  53388239 53393653      5415      2
## 4 Promoter (2-3kb)      943  52107945 52111988      4044      2
## 5 Distal Intergenic      951   3419459 3452683      33225      2
## 6 Promoter (<=1kb)      934  77761901 77795612      33712      1
##          geneId      transcriptId distanceToTSS
## 1 ENSDARG00000098037 ENSDART00000155813      3126
## 2 ENSDARG00000052396 ENSDART00000038391         0
## 3 ENSDARG00000117519 ENSDART00000194534      9794
## 4 ENSDARG00000117319 ENSDART00000194638     -2488
## 5 ENSDARG00000016858 ENSDART00000009740    -18014
## 6 ENSDARG00000104380 ENSDART00000159502         0
```

```
peakAnno_diff_df %>% names
```

```
## [1] "seqnames"      "start"          "end"            "width"
## [5] "strand"        "V4"             "Conc_WT1bOE"    "Conc_Ctrl"
## [9] "Fold"          "p.value"        "FDR"            "wt1bOE_1"
## [13] "V11"           "V12"            "V13"            "V14"
## [17] "V15"           "V16"            "annotation"      "geneChr"
## [21] "geneStart"     "geneEnd"        "geneLength"      "geneStrand"
## [25] "geneId"        "transcriptId"   "distanceToTSS"
```

```
as.data.frame(db_deseq2_df) %>% head
```

```
##   Chr   Start   End   Conc Conc_WT1bOE Conc_Ctrl   Fold p.value
## 1   2 14006002 14006402 4.256871   2.669184  4.994381 -1.1746329 1.04e-05
## 2  13 50200079 50200479 3.949543   2.408604  4.677540 -1.1351088 1.82e-05
## 3  17 53383459 53383859 9.149312   8.826909  9.412664 -0.5387331 2.16e-05
## 4  13 52114475 52114875 4.578848   3.273357  5.252782 -1.0901966 2.99e-05
```



```
## 5 21 3470696 3471096 4.318552 2.898409 5.020167 -1.0776029 3.96e-05
## 6 4 77761292 77762031 3.578587 1.695711 4.368405 -1.0402813 4.56e-05
##      FDR wt1bOE_1 wt1bOE_2 wt1bOE_3 URV_1 URV_3 URV_4 peak_name
## 1 0.05699248      +      -      +      +      +      + peak_1
## 2 0.05699248      -      +      -      +      +      + peak_2
## 3 0.05699248      +      +      +      +      +      + peak_3
## 4 0.05925655      +      +      -      +      +      + peak_4
## 5 0.06020326      +      -      -      +      +      + peak_5
## 6 0.06020326      -      -      -      +      +      + peak_6
```

```
diff_peaks_pval_norm %>% head()
```

```
##      Chr      Start      End      Conc Conc_WT1bOE Conc_Ctrl      Fold p.value
## 1 2 14006002 14006402 4.256871 2.669184 4.994381 -1.1746329 1.04e-05
## 2 13 50200079 50200479 3.949543 2.408604 4.677540 -1.1351088 1.82e-05
## 3 17 53383459 53383859 9.149312 8.826909 9.412664 -0.5387331 2.16e-05
## 4 13 52114475 52114875 4.578848 3.273357 5.252782 -1.0901966 2.99e-05
## 5 21 3470696 3471096 4.318552 2.898409 5.020167 -1.0776029 3.96e-05
## 6 4 77761292 77762031 3.578587 1.695711 4.368405 -1.0402813 4.56e-05
##      FDR wt1bOE_1 wt1bOE_2 wt1bOE_3 URV_1 URV_3 URV_4 peak_name
## 1 0.05699248      +      -      +      +      +      + peak_1
## 2 0.05699248      -      +      -      +      +      + peak_2
## 3 0.05699248      +      +      +      +      +      + peak_3
## 4 0.05925655      +      +      -      +      +      + peak_4
## 5 0.06020326      +      -      -      +      +      + peak_5
## 6 0.06020326      -      -      -      +      +      + peak_6
```

Convert genes names to symbol and entrezid

```
Genes=peakAnno_diff_df$geneId

ensembl_to_symbol_danio <- function(x){
  require(biomaRt)
  mart <- useMart(biomart = "ensembl", dataset = "drerio_gene_ensembl")
  # query biomaRt
  results <- getBM(attributes = c("ensembl_gene_id", "zfin_id_symbol",
```

```

      "entrezgene_id", "description"),
      filters = "ensembl_gene_id", values = x,
      mart = mart)

results
    }
zebrafish_genes <- ensembl_to_symbol_danio(Genes)

```

```
## Loading required package: biomaRt
```

```
head(zebrafish_genes)
```

```
##      ensembl_gene_id zfin_id_symbol entrezgene_id
## 1 ENSDARG00000000002      ccdc80      368419
## 2 ENSDARG00000000151      thraa      30670
## 3 ENSDARG00000000729      daxx      561006
## 4 ENSDARG00000000853      dstyk      402922
## 5 ENSDARG00000000861      scrib      368473
## 6 ENSDARG00000000966      ncor2      561503
##
##                                     description
## 1          coiled-coil domain containing 80 [Source:NCBI gene;Acc:368419]
## 2          thyroid hormone receptor alpha a [Source:NCBI gene;Acc:30670]
## 3          death-domain associated protein [Source:ZFIN;Acc:ZDB-GENE-010110-3]
## 4 dual serine/threonine and tyrosine protein kinase [Source:NCBI gene;Acc:402922]
## 5          scribble planar cell polarity protein [Source:NCBI gene;Acc:368473]
## 6          nuclear receptor corepressor 2 [Source:ZFIN;Acc:ZDB-GENE-030616-81]
```

```
head(peakAnno_diff_df)
```

```
##      seqnames      start      end width strand      V4 Conc_WT1bOE Conc_Ctrl
## 1          2 14006003 14006402   400      * 4.256871   2.669184  4.994381
## 2         13 50200080 50200479   400      * 3.949543   2.408604  4.677540
## 3         17 53383460 53383859   400      * 9.149312   8.826909  9.412664
## 4         13 52114476 52114875   400      * 4.578848   3.273357  5.252782
## 5         21 3470697 3471096   400      * 4.318552   2.898409  5.020167
## 6          4 77761293 77762031   739      * 3.578587   1.695711  4.368405
##      Fold p.value      FDR wt1bOE_1 V11 V12 V13 V14 V15      V16
```

```
## 1 -1.1746329 1.04e-05 0.05699248      + - + + + + peak_1
## 2 -1.1351088 1.82e-05 0.05699248      - + - + + + peak_2
## 3 -0.5387331 2.16e-05 0.05699248      + + + + + + peak_3
## 4 -1.0901966 2.99e-05 0.05925655      + + - + + + peak_4
## 5 -1.0776029 3.96e-05 0.06020326      + - - + + + peak_5
## 6 -1.0402813 4.56e-05 0.06020326      - - - + + + peak_6
##          annotation geneChr geneStart geneEnd geneLength geneStrand
## 1 Distal Intergenic      932  14007938 14009528      1591      2
## 2 Promoter (<=1kb)      943  50190664 50200348      9685      2
## 3 Distal Intergenic      947  53388239 53393653      5415      2
## 4 Promoter (2-3kb)      943  52107945 52111988      4044      2
## 5 Distal Intergenic      951   3419459 3452683      33225      2
## 6 Promoter (<=1kb)      934  77761901 77795612      33712      1
##          geneId      transcriptId distanceToTSS
## 1 ENSDARG00000098037 ENSDART00000155813      3126
## 2 ENSDARG00000052396 ENSDART00000038391         0
## 3 ENSDARG00000117519 ENSDART00000194534      9794
## 4 ENSDARG00000117319 ENSDART00000194638     -2488
## 5 ENSDARG00000016858 ENSDART00000009740    -18014
## 6 ENSDARG00000104380 ENSDART00000159502         0
```

```
peakAnno_diff_df_entrez <- merge(peakAnno_diff_df, zebrafish_genes, by.x= "geneId", by.y = "ensembl_gene_id", all.x=T)
(peakAnno_diff_df_entrez) %>% filter(geneId== "ENSDARG00000104380")
```

```
##          geneId seqnames      start      end width strand      V4
## 1 ENSDARG00000104380      4 77761293 77762031    739      * 3.578587
##   Conc_WT1b0E Conc_Ctrl      Fold p.value      FDR wt1b0E_1 V11 V12 V13 V14
## 1   1.695711  4.368405 -1.040281 4.56e-05 0.06020326      - - - + +
##   V15      V16      annotation geneChr geneStart geneEnd geneLength geneStrand
## 1   + peak_6 Promoter (<=1kb)      934  77761901 77795612      33712      1
##          transcriptId distanceToTSS      zfin_id_symbol entrezgene_id
## 1 ENSDART00000159502         0 si:dkey-238k10.2      NA
##          description
## 1 si:dkey-238k10.2 [Source:ZFIN;Acc:ZDB-GENE-070705-400]
```

```
dim(peakAnno_diff_df_entrez)
```

```
## [1] 1451 30
```

```
peakAnno_diff_df_entrez <- peakAnno_diff_df_entrez %>% arrange(p.value)
head(peakAnno_diff_df_entrez)
```

```
##          geneId seqnames      start      end width strand      V4
## 1 ENSDARG00000098037      2 14006003 14006402    400      * 4.256871
## 2 ENSDARG00000052396     13 50200080 50200479    400      * 3.949543
## 3 ENSDARG00000117519     17 53383460 53383859    400      * 9.149312
## 4 ENSDARG00000117319     13 52114476 52114875    400      * 4.578848
## 5 ENSDARG00000016858     21 3470697 3471096    400      * 4.318552
## 6 ENSDARG00000104380      4 77761293 77762031    739      * 3.578587
##   Conc_WT1bOE Conc_Ctrl      Fold  p.value      FDR wt1bOE_1 V11 V12 V13 V14
## 1    2.669184  4.994381 -1.1746329 1.04e-05 0.05699248      +  -  +  +  +
## 2    2.408604  4.677540 -1.1351088 1.82e-05 0.05699248      -  +  -  +  +
## 3    8.826909  9.412664 -0.5387331 2.16e-05 0.05699248      +  +  +  +  +
## 4    3.273357  5.252782 -1.0901966 2.99e-05 0.05925655      +  +  -  +  +
## 5    2.898409  5.020167 -1.0776029 3.96e-05 0.06020326      +  -  -  +  +
## 6    1.695711  4.368405 -1.0402813 4.56e-05 0.06020326      -  -  -  +  +
##   V15      V16      annotation geneChr geneStart  geneEnd geneLength geneStrand
## 1  + peak_1 Distal Intergenic    932  14007938 14009528    1591         2
## 2  + peak_2 Promoter (<=1kb)    943  50190664 50200348    9685         2
## 3  + peak_3 Distal Intergenic    947  53388239 53393653    5415         2
## 4  + peak_4 Promoter (2-3kb)    943  52107945 52111988    4044         2
## 5  + peak_5 Distal Intergenic    951   3419459 3452683    33225         2
## 6  + peak_6 Promoter (<=1kb)    934  77761901 77795612    33712         1
##   transcriptId distanceToTSS  zfin_id_symbol entrezgene_id
## 1 ENSDART00000155813      3126                    100331996
## 2 ENSDART00000038391         0                      pkz      503703
## 3 ENSDART00000194534      9794                    101884319
## 4 ENSDART00000194638     -2488                      NA
## 5 ENSDART00000009740    -18014                  smad7      326282
## 6 ENSDART00000159502         0 si:dkey-238k10.2          NA
##                                     description
## 1                uncharacterized LOC100331996 [Source:NCBI gene;Acc:100331996]
## 2 protein kinase containing Z-DNA binding domains [Source:ZFIN;Acc:ZDB-GENE-050301-2]
## 3                protein unc-79 homolog [Source:NCBI gene;Acc:101884319]
## 4
## 5                SMAD family member 7 [Source:ZFIN;Acc:ZDB-GENE-030128-3]
## 6                si:dkey-238k10.2 [Source:ZFIN;Acc:ZDB-GENE-070705-400]
```

Write the csv for annotated differential peaks

```
write.csv(peakAnno_diff_df_entrez, file= "./results/differential_peaks_with_annotations.csv",row.names=F,col.names=T)
```

```
## Warning in write.csv(peakAnno_diff_df_entrez, file = "./results/  
## differential_peaks_with_annotations.csv", : attempt to set 'col.names' ignored
```

Pathway analysis

Convert zebrafish genes to mouse genes and merge to original dataframe

```
zgGenes <- peakAnno_diff_df_entrez$geneId  
# Basic function to convert zebrafish to human gene names  
  
convertDanioGeneList_Mouse <- function(x){  
  require("biomaRt")  
  mouse = useMart("ensembl", dataset = "mmusculus_gene_ensembl")  
  danio = useMart("ensembl", dataset = "drerio_gene_ensembl")  
  genesV2 = getLDS(attributes = c("ensembl_gene_id", "zfin_id_symbol"),  
                   filters = "ensembl_gene_id", values = x ,  
                   mart = danio, attributesL = c("mgi_symbol", "ensembl_gene_id","entrezgene_id"),  
                   martL = mouse, uniqueRows=T)  
  
  colnames(genesV2)[colnames(genesV2)== "Gene.stable.ID"] <- "EnsemblID_Zebrafish"  
  colnames(genesV2)[colnames(genesV2)== "NCBI.gene..formerly.Entrezgene..ID"] <- "EntrezID_Mouse"  
  colnames(genesV2)[colnames(genesV2)== "Gene.stable.ID.1"] <- "EnsemblID_Mouse"  
  
  # Print the first 6 genes found to the screen  
  return(genesV2)  
}  
  
Mouse_Genes <- convertDanioGeneList_Mouse(zgGenes)  
Mouse_Genes %>% head
```

```
##   EnsemblID_Zebrafish ZFIN.symbol MGI.symbol   EnsemblID_Mouse EntrezID_Mouse
```

```
## 1 ENSDARG00000036630      plp2      Gm6169 ENSMUSG00000057762      NA
## 2 ENSDARG00000039576      fstl1b     Fstl1  ENSMUSG00000022816      14314
## 3 ENSDARG00000103069                Lsmp  ENSMUSG00000061080      268890
## 4 ENSDARG00000104573      slc12a2    Slc12a2 ENSMUSG00000024597      20496
## 5 ENSDARG00000061222      uba7       Uba1y  ENSMUSG00000069053      22202
## 6 ENSDARG00000105972                Shisa8 ENSMUSG00000096883      435145
```

```
peakAnno_diff_mouse <- merge(peakAnno_diff_df_entrez, Mouse_Genes, by.x= "geneId", by.y = "EnsemblID_Zebrafish", all.x= TRUE)
head(peakAnno_diff_mouse)
```

```
##          geneId seqnames      start      end width strand      V4
## 1 ENSDARG00000000002      9 34088122 34088521   400      * 2.569013
## 2 ENSDARG00000000002      9 34094262 34094661   400      * 2.726597
## 3 ENSDARG000000000151     3 34736807 34737206   400      * 1.922126
## 4 ENSDARG000000000151     3 34753178 34753577   400      * 4.249876
## 5 ENSDARG000000000729    19 7033489 7033888   400      * 2.886382
## 6 ENSDARG000000000853    22 506412 506811   400      * 3.275082
##   Conc_WT1bOE Conc_Ctrl      Fold   p.value      FDR wt1bOE_1 V11 V12 V13
## 1 1.5435457 3.162365 -0.4443180 0.03939099 1.0000000 - - - +
## 2 1.6935232 3.322418 -0.4866469 0.02833425 1.0000000 - - - +
## 3 0.3856869 2.649188 -0.3581551 0.04437472 1.0000000 - - - -
## 4 3.6419175 4.676275 -0.5482421 0.02209098 0.3451784 + + + +
## 5 1.8685860 3.477229 -0.5002292 0.02686056 0.3670568 - - - -
## 6 2.3367745 3.838880 -0.5610089 0.01762784 0.3230954 - - - -
##   V14 V15      V16
## 1 - + peak_1179
## 2 + + peak_892
## 3 + + peak_1297
## 4 + + peak_728
## 5 + + peak_853
## 6 + + peak_608
##
##          annotation geneChr
## 1 Promoter (<=1kb) 939
## 2 Intron (ENSDART000000000005/ENSDARG000000000002, intron 2 of 6) 939
## 3 Intron (ENSDART000000000160/ENSDARG000000000151, intron 1 of 9) 933
## 4 Promoter (<=1kb) 933
## 5 Promoter (<=1kb) 949
## 6 Promoter (<=1kb) 952
```

```
##      geneStart  geneEnd  geneLength  geneStrand      transcriptId  distanceToTSS
## 1   34089156  34113209      24054          1  ENSDART000000000005      -635
## 2   34089156  34113209      24054          1  ENSDART000000000005       5106
## 3   34697821  34724879      27059          2  ENSDART00000177021      -11928
## 4   34743579  34753541       9963          2  ENSDART00000151633         0
## 5    7014455  7033636      19182          2  ENSDART00000122815         0
## 6    472327   506522      34196          2  ENSDART00000106645         0
##      zfin_id_symbol  entrezgene_id
## 1             ccdc80      368419
## 2             ccdc80      368419
## 3             thraa       30670
## 4             thraa       30670
## 5             daxx       561006
## 6             dstyk       402922
##                                     description
## 1             coiled-coil domain containing 80 [Source:NCBI gene;Acc:368419]
## 2             coiled-coil domain containing 80 [Source:NCBI gene;Acc:368419]
## 3             thyroid hormone receptor alpha a [Source:NCBI gene;Acc:30670]
## 4             thyroid hormone receptor alpha a [Source:NCBI gene;Acc:30670]
## 5             death-domain associated protein [Source:ZFIN;Acc:ZDB-GENE-010110-3]
## 6 dual serine/threonine and tyrosine protein kinase [Source:NCBI gene;Acc:402922]
##      ZFIN.symbol  MGI.symbol      EnsemblID_Mouse  EntrezID_Mouse
## 1             <NA>      <NA>          <NA>          NA
## 2             <NA>      <NA>          <NA>          NA
## 3             thraa      Thra  ENSMUSG00000058756      21833
## 4             thraa      Thra  ENSMUSG00000058756      21833
## 5             daxx      Daxx  ENSMUSG00000002307      13163
## 6             dstyk      Dstyk  ENSMUSG00000042046      213452
```

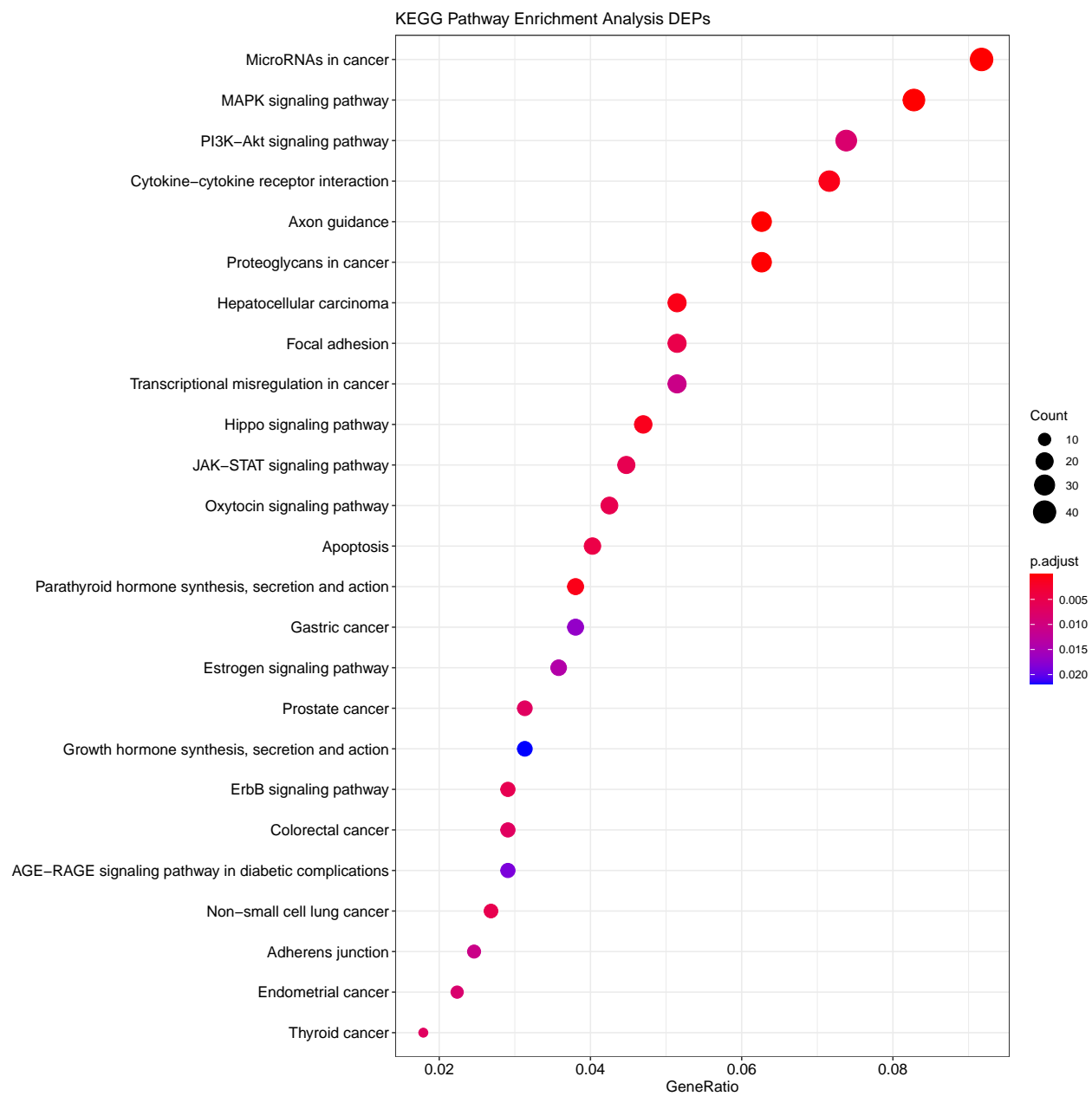
```
dim(peakAnno_diff_mouse)
```

```
## [1] 1633   34
```

```
write.csv(peakAnno_diff_mouse, "./results/mouse_differential_peaks_annotations.csv")
```

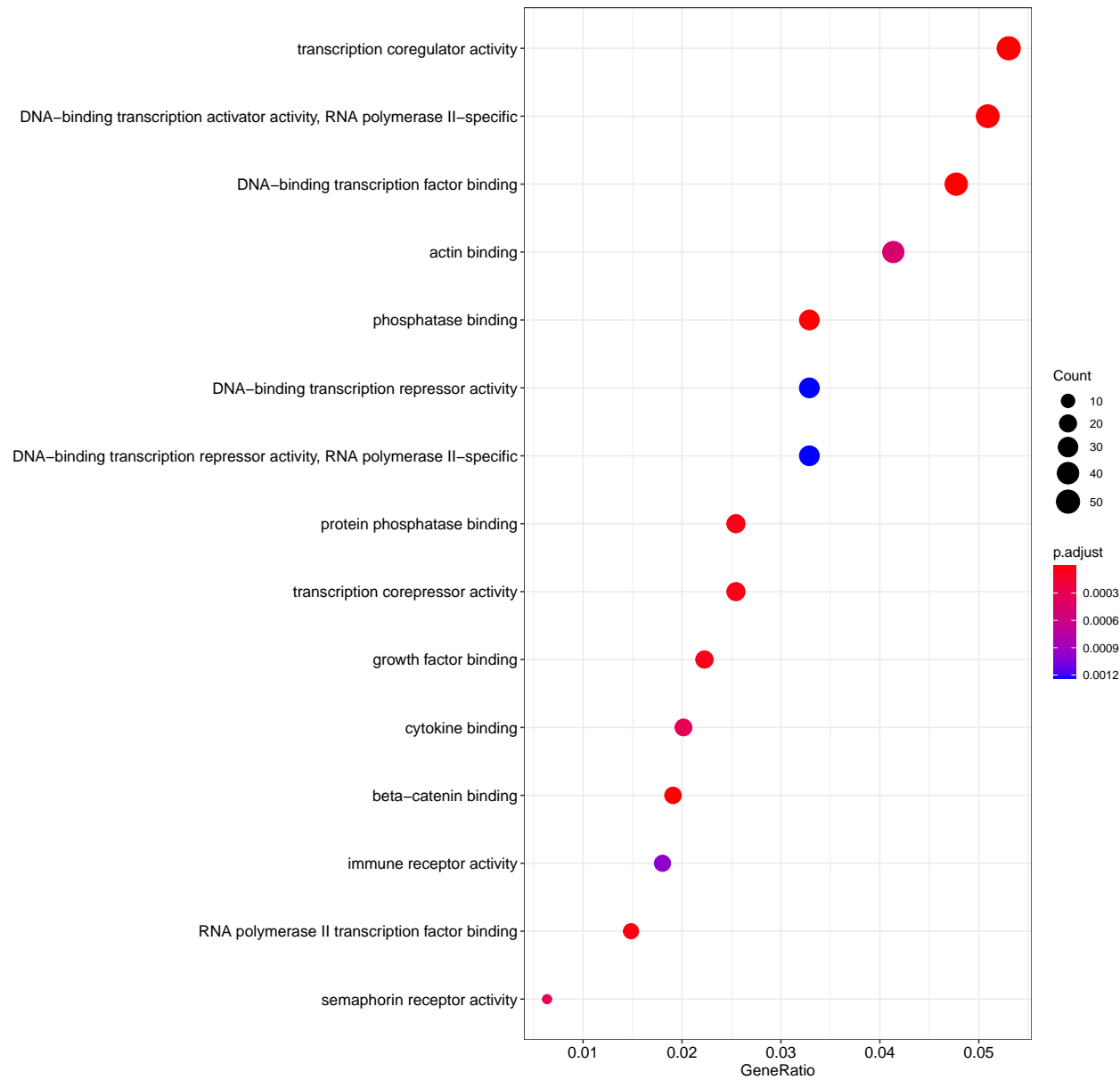
Perfroming enrichment analysis on differential peaks/genes.

```
compKEGG_diff <- enrichKEGG(peakAnno_diff_mouse$EntrezID_Mouse, pvalueCutoff = 0.05,  
                             pAdjustMethod = "BH",organism = "mmu")  
  
## Reading KEGG annotation online:  
##  
## Reading KEGG annotation online:  
  
dotplot(compKEGG_diff, showCategory = 25, title = "KEGG Pathway Enrichment Analysis DEPs")
```

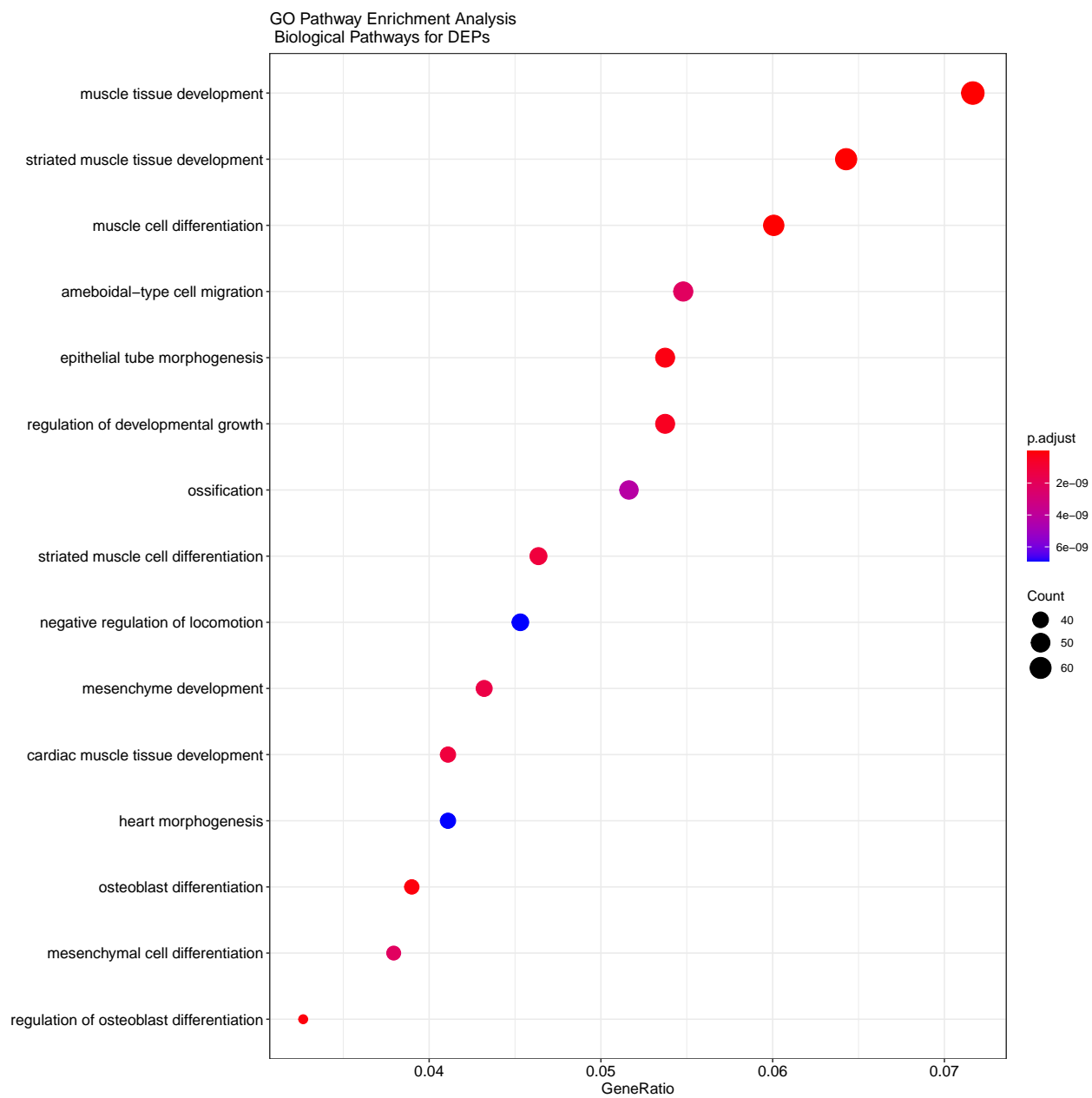



```
compGO_MF_diff <- enrichGO(gene = peakAnno_diff_mouse$EntrezID_Mouse, pvalueCutoff = 0.05,  
                             pAdjustMethod = "BH", OrgDb = "org.Mm.eg.db", ont = "MF", readable = TRUE)  
  
dotplot(compGO_MF_diff, showCategory = 15, title = "GO Pathway Enrichment Analysis \n Molecular Functions for DEPs",  
         font.size = 12)
```

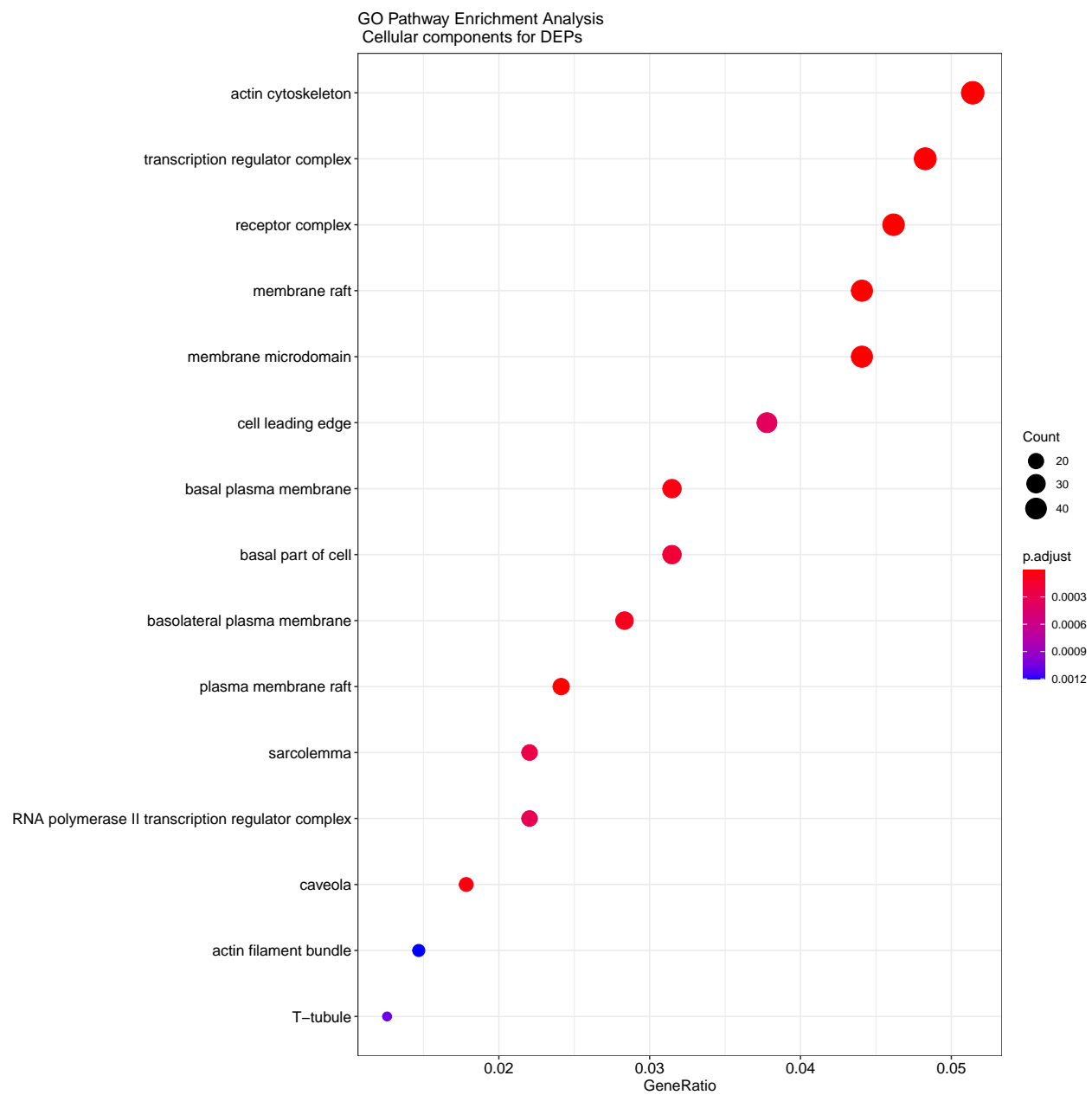
GO Pathway Enrichment Analysis
Molecular Functions for DEPs



```
compGO_BP_diff <- enrichGO(gene = peakAnno_diff_mouse$EntrezID_Mouse, pvalueCutoff = 0.05,  
                             pAdjustMethod = "BH", OrgDb = "org.Mm.eg.db", ont = "BP", readable = TRUE)  
  
dotplot(compGO_BP_diff, showCategory = 15, title = "GO Pathway Enrichment Analysis \n Biological Pathways for DEPs",  
         font.size = 12)
```



```
compGO_CC_diff <- enrichGO(gene = peakAnno_diff_mouse$EntrezID_Mouse, pvalueCutoff = 0.05,  
                             pAdjustMethod = "BH", OrgDb = "org.Mm.eg.db", ont = "CC", readable = TRUE)  
  
dotplot(compGO_CC_diff, showCategory = 15, title = "GO Pathway Enrichment Analysis \n Cellular components for DEPs",  
         font.size = 12)
```



```

reactome_pathways_diff <- enrichPathway(peakAnno_diff_mouse$EntrezID_Mouse, organism = "mouse", readable = T,
                                         pvalueCutoff = 0.05, qvalueCutoff = 0.2)
head(summary(reactome_pathways_diff))

```

```

## Warning in summary(reactome_pathways_diff): summary method to convert the object
## to data.frame is deprecated, please use as.data.frame instead.

```

```

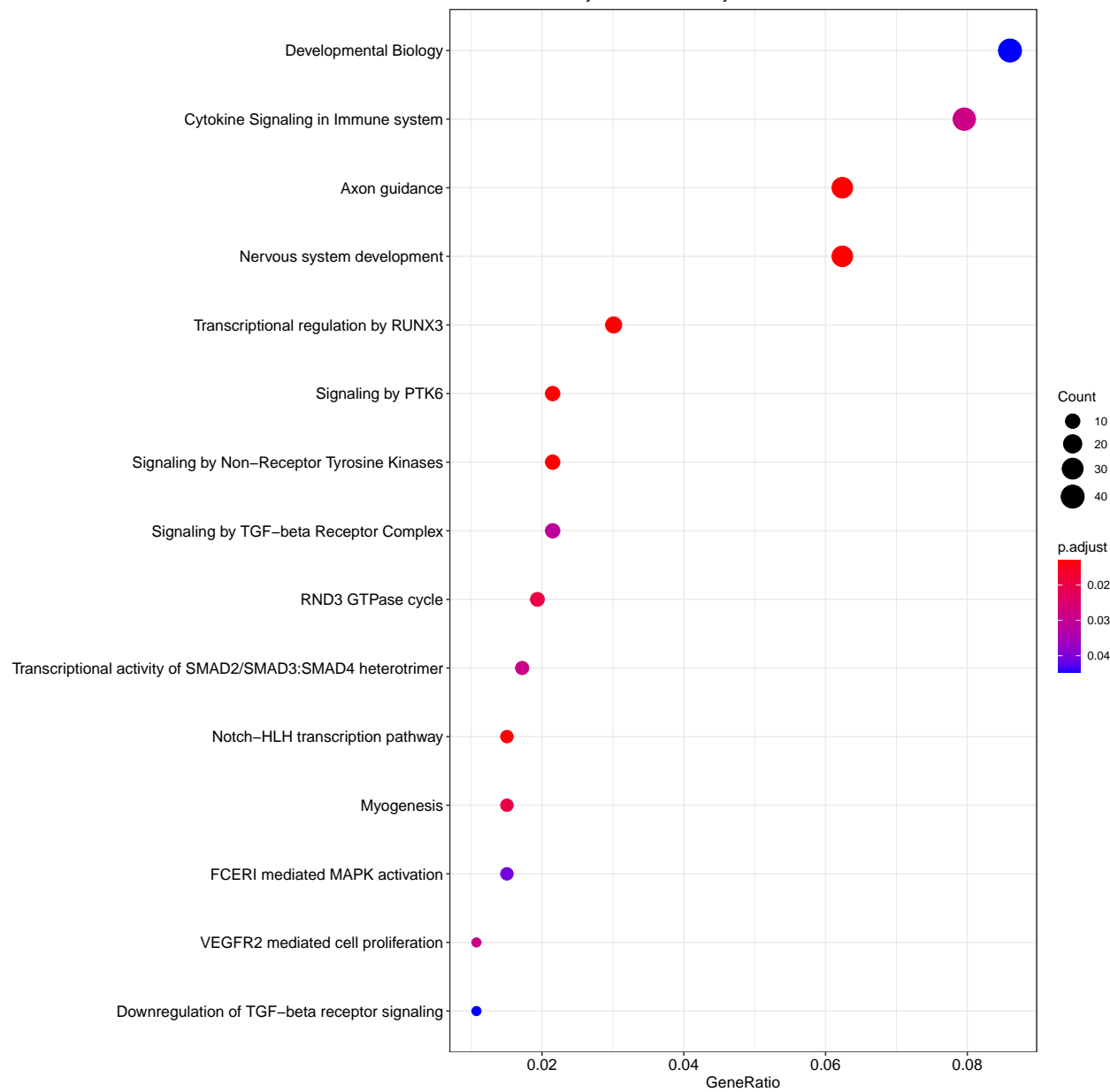
##              ID              Description
## R-MMU-8878159 R-MMU-8878159 Transcriptional regulation by RUNX3
## R-MMU-422475  R-MMU-422475  Axon guidance
## R-MMU-8848021 R-MMU-8848021  Signaling by PTK6
## R-MMU-9006927 R-MMU-9006927 Signaling by Non-Receptor Tyrosine Kinases
## R-MMU-9675108 R-MMU-9675108  Nervous system development
## R-MMU-350054  R-MMU-350054  Notch-HLH transcription pathway
##      GeneRatio  BgRatio      pvalue  p.adjust    qvalue
## R-MMU-8878159   14/465   78/8880 4.465274e-05 0.01300994 0.01206728
## R-MMU-422475    29/465  256/8880 6.794825e-05 0.01300994 0.01206728
## R-MMU-8848021   10/465   44/8880 6.979601e-05 0.01300994 0.01206728
## R-MMU-9006927   10/465   44/8880 6.979601e-05 0.01300994 0.01206728
## R-MMU-9675108   29/465  257/8880 7.292567e-05 0.01300994 0.01206728
## R-MMU-350054     7/465   22/8880 8.847670e-05 0.01315354 0.01220047
##
## R-MMU-8878159                                     Psme1/Tcf712/Brd2/Psmb7/Smurf2/Tcf711/Cb
## R-MMU-422475  Ptk2/Hras/Itga5/Plxna3/Fgfr1/Plxnc1/Trio/Sh3gl2/Ezr/Numb/Efna1/Enah/Abl1/Grb10/Rasa1/Plcg1/Uba52/Tuba1a/Ncam1/Egfr/Msn/Pl
## R-MMU-8848021                                     Hras/Soc
## R-MMU-9006927                                     Hras/Soc
## R-MMU-9675108 Ptk2/Hras/Itga5/Plxna3/Fgfr1/Plxnc1/Trio/Sh3gl2/Ezr/Numb/Efna1/Enah/Abl1/Grb10/Rasa1/Plcg1/Uba52/Tuba1a/Ncam1/Egfr/Msn/Pl
## R-MMU-350054
##      Count
## R-MMU-8878159   14
## R-MMU-422475    29
## R-MMU-8848021   10
## R-MMU-9006927   10
## R-MMU-9675108   29
## R-MMU-350054     7

```



```
dotplot(reactome_pathways_diff, showCategory = 25, title = "Reactome Pathway Enrichment Analysis",  
        font.size = 12)
```

Reactome Pathway Enrichment Analysis



Write the table for differential pathways

```
compGO_CC_diff_df <- as.data.frame(compGO_CC_diff)

compGO_CC_diff_df$GeneRatio_decimal <- compGO_CC_diff_df$GeneRatio
compGO_CC_diff_df$GeneRatio_decimal <- sapply(compGO_CC_diff_df$GeneRatio_decimal,
  function(x) (eval(parse(text = as.character(x)))))
compGO_CC_diff_df$BgRatio_decimal <- compGO_CC_diff_df$BgRatio
compGO_CC_diff_df$BgRatio_decimal <- sapply(compGO_CC_diff_df$BgRatio_decimal,
  function(x) (eval(parse(text = as.character(x)))))
compGO_CC_diff_df <- compGO_CC_diff_df %>% tidyr::separate_rows(geneID, sep = "/", convert = FALSE) %>%
  arrange(desc(GeneRatio_decimal))
compGO_CC_diff_df %>% head
```

```
## # A tibble: 6 x 11
##   ID      Description  GeneRatio BgRatio    pvalue p.adjust    qvalue geneID Count
##   <chr>   <chr>        <chr>    <chr>      <dbl>   <dbl>    <dbl> <chr>  <int>
## 1 G0:00~ actin cytosk~ 49/953    500/231~ 1.93e-8  2.25e-6  1.90e-6 Myh9      49
## 2 G0:00~ actin cytosk~ 49/953    500/231~ 1.93e-8  2.25e-6  1.90e-6 Dctn5     49
## 3 G0:00~ actin cytosk~ 49/953    500/231~ 1.93e-8  2.25e-6  1.90e-6 Ptk2      49
## 4 G0:00~ actin cytosk~ 49/953    500/231~ 1.93e-8  2.25e-6  1.90e-6 Ppp1r~    49
## 5 G0:00~ actin cytosk~ 49/953    500/231~ 1.93e-8  2.25e-6  1.90e-6 Ints6     49
## 6 G0:00~ actin cytosk~ 49/953    500/231~ 1.93e-8  2.25e-6  1.90e-6 Flnc      49
## # ... with 2 more variables: GeneRatio_decimal <dbl>, BgRatio_decimal <dbl>
```

```
write.csv(compGO_CC_diff_df, "results/differential_pathways_G0_Cell_components_norm.csv")
```

```
compGO_MF_diff_df <- as.data.frame(compGO_MF_diff)

compGO_MF_diff_df$GeneRatio_decimal <- compGO_MF_diff_df$GeneRatio
compGO_MF_diff_df$GeneRatio_decimal <- sapply(compGO_MF_diff_df$GeneRatio_decimal,
  function(x) (eval(parse(text = as.character(x)))))
compGO_MF_diff_df$BgRatio_decimal <- compGO_MF_diff_df$BgRatio
compGO_MF_diff_df$BgRatio_decimal <- sapply(compGO_MF_diff_df$BgRatio_decimal,
  function(x) (eval(parse(text = as.character(x)))))
compGO_MF_diff_df <- compGO_MF_diff_df %>% tidyr::separate_rows(geneID, sep = "/", convert = FALSE) %>%
```

```

  arrange(desc(GeneRatio_decimal))
  compGO_MF_diff_df %>% head

```

```

## # A tibble: 6 x 11
##   ID      Description  GeneRatio BgRatio  pvalue p.adjust  qvalue geneID Count
##   <chr>   <chr>        <chr>    <chr>    <dbl>  <dbl>    <dbl> <chr>  <int>
## 1 G0:00~ transcription~ 50/943    466/22~ 9.54e-10 7.99e-7 6.93e-7 Daxx    50
## 2 G0:00~ transcription~ 50/943    466/22~ 9.54e-10 7.99e-7 6.93e-7 Ncor2    50
## 3 G0:00~ transcription~ 50/943    466/22~ 9.54e-10 7.99e-7 6.93e-7 Runx1~   50
## 4 G0:00~ transcription~ 50/943    466/22~ 9.54e-10 7.99e-7 6.93e-7 Irf2b~   50
## 5 G0:00~ transcription~ 50/943    466/22~ 9.54e-10 7.99e-7 6.93e-7 Irf2b~   50
## 6 G0:00~ transcription~ 50/943    466/22~ 9.54e-10 7.99e-7 6.93e-7 Smarc~   50
## # ... with 2 more variables: GeneRatio_decimal <dbl>, BgRatio_decimal <dbl>

```

```

write.csv(compGO_MF_diff_df, "results/differential_pathways_G0_Molecular_Functions_norm.csv")

```

```

compGO_BP_diff_df <- as.data.frame(compGO_BP_diff)

compGO_BP_diff_df$GeneRatio_decimal <- compGO_BP_diff_df$GeneRatio
compGO_BP_diff_df$GeneRatio_decimal <- sapply(compGO_BP_diff_df$GeneRatio_decimal,
  function(x) (eval(parse(text = as.character(x)))))
compGO_BP_diff_df$BgRatio_decimal <- compGO_BP_diff_df$BgRatio
compGO_BP_diff_df$BgRatio_decimal <- sapply(compGO_BP_diff_df$BgRatio_decimal,
  function(x) (eval(parse(text = as.character(x)))))
compGO_BP_diff_df <- compGO_BP_diff_df %>% tidyr::separate_rows(geneID, sep = "/", convert = FALSE) %>%
  arrange(desc(GeneRatio_decimal))
compGO_BP_diff_df %>% head

```

```

## # A tibble: 6 x 11
##   ID      Description  GeneRatio BgRatio  pvalue p.adjust  qvalue geneID Count
##   <chr>   <chr>        <chr>    <chr>    <dbl>  <dbl>    <dbl> <chr>  <int>
## 1 G0:00~ muscle tissu~ 68/949    477/23~ 1.41e-19 7.27e-16 5.17e-16 Tbx3    68
## 2 G0:00~ muscle tissu~ 68/949    477/23~ 1.41e-19 7.27e-16 5.17e-16 Rcan1    68
## 3 G0:00~ muscle tissu~ 68/949    477/23~ 1.41e-19 7.27e-16 5.17e-16 Sox9     68
## 4 G0:00~ muscle tissu~ 68/949    477/23~ 1.41e-19 7.27e-16 5.17e-16 Usp19    68
## 5 G0:00~ muscle tissu~ 68/949    477/23~ 1.41e-19 7.27e-16 5.17e-16 Tcf7l2   68

```

```
## 6 GO:00~ muscle tissu~ 68/949    477/23~ 1.41e-19 7.27e-16 5.17e-16 SrpK3    68
## # ... with 2 more variables: GeneRatio_decimal <dbl>, BgRatio_decimal <dbl>
```

```
write.csv(compGO_BP_diff_df, "results/differential_pathways_GO_Biological_Pathways_norm.csv")
```

```
compGO_MF_diff_df_plot <- as.data.frame(compGO_MF_diff)
compGO_MF_diff_df_plot$GeneRatio_decimal <- compGO_MF_diff_df_plot$GeneRatio
compGO_MF_diff_df_plot$GeneRatio_decimal <- sapply(compGO_MF_diff_df_plot$GeneRatio_decimal,
                                                    function(x) (eval(parse(text = as.character(x)))))
compGO_MF_diff_df_plot <- compGO_MF_diff_df_plot %>% arrange(desc(GeneRatio_decimal))
head(compGO_MF_diff_df_plot, 25)
```

```
##          ID
## GO:0003712 GO:0003712
## GO:0001228 GO:0001228
## GO:0140297 GO:0140297
## GO:0003779 GO:0003779
## GO:0060589 GO:0060589
## GO:0008047 GO:0008047
## GO:0030695 GO:0030695
## GO:0019902 GO:0019902
## GO:0001217 GO:0001217
## GO:0001227 GO:0001227
## GO:0050839 GO:0050839
## GO:0051020 GO:0051020
## GO:0060090 GO:0060090
## GO:0003924 GO:0003924
## GO:0019903 GO:0019903
## GO:0003714 GO:0003714
## GO:0003713 GO:0003713
## GO:0008022 GO:0008022
## GO:0031267 GO:0031267
## GO:0019838 GO:0019838
## GO:0051015 GO:0051015
## GO:0030674 GO:0030674
## GO:0019955 GO:0019955
## GO:0005085 GO:0005085
## GO:0008013 GO:0008013
```

##	Description				
## G0:0003712	transcription coregulator activity				
## G0:0001228	DNA-binding transcription activator activity, RNA polymerase II-specific				
## G0:0140297	DNA-binding transcription factor binding				
## G0:0003779	actin binding				
## G0:0060589	nucleoside-triphosphatase regulator activity				
## G0:0008047	enzyme activator activity				
## G0:0030695	GTPase regulator activity				
## G0:0019902	phosphatase binding				
## G0:0001217	DNA-binding transcription repressor activity				
## G0:0001227	DNA-binding transcription repressor activity, RNA polymerase II-specific				
## G0:0050839	cell adhesion molecule binding				
## G0:0051020	GTPase binding				
## G0:0060090	molecular adaptor activity				
## G0:0003924	GTPase activity				
## G0:0019903	protein phosphatase binding				
## G0:0003714	transcription corepressor activity				
## G0:0003713	transcription coactivator activity				
## G0:0008022	protein C-terminus binding				
## G0:0031267	small GTPase binding				
## G0:0019838	growth factor binding				
## G0:0051015	actin filament binding				
## G0:0030674	protein-macromolecule adaptor activity				
## G0:0019955	cytokine binding				
## G0:0005085	guanyl-nucleotide exchange factor activity				
## G0:0008013	beta-catenin binding				
##	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
## G0:0003712	50/943	466/22682	9.539619e-10	7.994201e-07	6.928776e-07
## G0:0001228	48/943	496/22682	5.526813e-08	9.262938e-06	8.028423e-06
## G0:0140297	45/943	410/22682	3.277608e-09	1.358905e-06	1.177797e-06
## G0:0003779	39/943	437/22682	6.892845e-06	4.813504e-04	4.171985e-04
## G0:0060589	37/943	463/22682	1.227956e-04	4.549962e-03	3.943567e-03
## G0:0008047	35/943	470/22682	6.808507e-04	1.901843e-02	1.648375e-02
## G0:0030695	33/943	422/22682	4.117341e-04	1.327051e-02	1.150188e-02
## G0:0019902	31/943	225/22682	4.864815e-09	1.358905e-06	1.177797e-06
## G0:0001217	31/943	330/22682	2.211942e-05	1.235738e-03	1.071046e-03
## G0:0001227	31/943	330/22682	2.211942e-05	1.235738e-03	1.071046e-03
## G0:0050839	27/943	287/22682	7.147304e-05	3.152337e-03	2.732211e-03
## G0:0051020	26/943	307/22682	5.050214e-04	1.567437e-02	1.358537e-02

## G0:0060090	26/943	337/22682	1.946337e-03	4.077577e-02	3.534139e-02
## G0:0003924	25/943	298/22682	7.355778e-04	1.988433e-02	1.723425e-02
## G0:0019903	24/943	178/22682	3.909953e-07	4.680772e-05	4.056944e-05
## G0:0003714	24/943	180/22682	4.814726e-07	5.043425e-05	4.371264e-05
## G0:0003713	24/943	245/22682	9.471566e-05	3.779606e-03	3.275880e-03
## G0:0008022	23/943	234/22682	1.248796e-04	4.549962e-03	3.943567e-03
## G0:0031267	23/943	270/22682	9.581194e-04	2.433043e-02	2.108780e-02
## G0:0019838	21/943	147/22682	7.869379e-07	7.327266e-05	6.350727e-05
## G0:0051015	21/943	212/22682	2.170885e-04	7.580006e-03	6.569783e-03
## G0:0030674	21/943	239/22682	1.059738e-03	2.611943e-02	2.263837e-02
## G0:0019955	19/943	138/22682	4.552078e-06	3.467856e-04	3.005679e-04
## G0:0005085	19/943	197/22682	5.919400e-04	1.771592e-02	1.535483e-02
## G0:0008013	18/943	92/22682	3.770407e-08	7.899002e-06	6.846265e-06
##					
## G0:0003712	Daxx/Ncor2/Runx1t1/Irf2bpl/Irf2bp2/Smarca2/Tbl1xr1/Rbm39/Setd3/Lpin1/Btg2/Tob1/Brd8/Per2/Cited4/Kdm2b/Calcoco1/Ezh1/Smarcd3/				
## G0:0001228	Sox9/Tcf7l2/Foxf2/Foxi1/Esrrg/Irf4/Mycn/Myc/Atf3/Mef2c/Batf/Zfp523/Creb3l1/Elk3/Nr1i3/Nfatc1/Egr4/Osr2/Ts				
## G0:0140297	Daxx/Ncor2/Tbx3/Sox9/Tcf7l2/Hes1/Myc/Mef2c/Setd3/Taf7/Tal1/Usf2/Lpin1/				
## G0:0003779	Myh9/Ptk2/Setd3/Flnc/Mical2/Tns1/Myo1b/Ezr/Lsp1/				
## G0:0060589	Arhgef37/Gdi2/Katnb1/Srgap1/Wdr41/Atp1b1/Arhgef3/Rapgef1/Iqsec1/Denn				
## G0:0008047	Daxx/Psmel/Tnfrsf10b/Gdi2/Srgap1/Gas6/Atp1b1/Asap				
## G0:0030695	Arhgef37/Gdi2/Srgap1/Wdr41/Arhgef3/Rapgef1				
## G0:0019902	Hsp90b1/P				
## G0:0001217					
## G0:0001227					
## G0:0050839					
## G0:0051020					
## G0:0060090					
## G0:0003924					
## G0:0019903					
## G0:0003714					
## G0:0003713					
## G0:0008022					
## G0:0031267					
## G0:0019838					
## G0:0051015					
## G0:0030674					
## G0:0019955					
## G0:0005085					
## G0:0008013					

##	Count	GeneRatio_decimal
## G0:0003712	50	0.05302227
## G0:0001228	48	0.05090138
## G0:0140297	45	0.04772004
## G0:0003779	39	0.04135737
## G0:0060589	37	0.03923648
## G0:0008047	35	0.03711559
## G0:0030695	33	0.03499470
## G0:0019902	31	0.03287381
## G0:0001217	31	0.03287381
## G0:0001227	31	0.03287381
## G0:0050839	27	0.02863203
## G0:0051020	26	0.02757158
## G0:0060090	26	0.02757158
## G0:0003924	25	0.02651113
## G0:0019903	24	0.02545069
## G0:0003714	24	0.02545069
## G0:0003713	24	0.02545069
## G0:0008022	23	0.02439024
## G0:0031267	23	0.02439024
## G0:0019838	21	0.02226935
## G0:0051015	21	0.02226935
## G0:0030674	21	0.02226935
## G0:0019955	19	0.02014846
## G0:0005085	19	0.02014846
## G0:0008013	18	0.01908802

```
compG0_MF_diff_df_plot[1:25,]
```

##	ID
## G0:0003712	G0:0003712
## G0:0001228	G0:0001228
## G0:0140297	G0:0140297
## G0:0003779	G0:0003779
## G0:0060589	G0:0060589
## G0:0008047	G0:0008047
## G0:0030695	G0:0030695
## G0:0019902	G0:0019902

GO:0001217 GO:0001217
 ## GO:0001227 GO:0001227
 ## GO:0050839 GO:0050839
 ## GO:0051020 GO:0051020
 ## GO:0060090 GO:0060090
 ## GO:0003924 GO:0003924
 ## GO:0019903 GO:0019903
 ## GO:0003714 GO:0003714
 ## GO:0003713 GO:0003713
 ## GO:0008022 GO:0008022
 ## GO:0031267 GO:0031267
 ## GO:0019838 GO:0019838
 ## GO:0051015 GO:0051015
 ## GO:0030674 GO:0030674
 ## GO:0019955 GO:0019955
 ## GO:0005085 GO:0005085
 ## GO:0008013 GO:0008013

	Description
## GO:0003712	transcription coregulator activity
## GO:0001228	DNA-binding transcription activator activity, RNA polymerase II-specific
## GO:0140297	DNA-binding transcription factor binding
## GO:0003779	actin binding
## GO:0060589	nucleoside-triphosphatase regulator activity
## GO:0008047	enzyme activator activity
## GO:0030695	GTPase regulator activity
## GO:0019902	phosphatase binding
## GO:0001217	DNA-binding transcription repressor activity
## GO:0001227	DNA-binding transcription repressor activity, RNA polymerase II-specific
## GO:0050839	cell adhesion molecule binding
## GO:0051020	GTPase binding
## GO:0060090	molecular adaptor activity
## GO:0003924	GTPase activity
## GO:0019903	protein phosphatase binding
## GO:0003714	transcription corepressor activity
## GO:0003713	transcription coactivator activity
## GO:0008022	protein C-terminus binding
## GO:0031267	small GTPase binding
## GO:0019838	growth factor binding
## GO:0051015	actin filament binding

G0:0030674 protein-macromolecule adaptor activity
 ## G0:0019955 cytokine binding
 ## G0:0005085 guanyl-nucleotide exchange factor activity
 ## G0:0008013 beta-catenin binding

##	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
## G0:0003712	50/943	466/22682	9.539619e-10	7.994201e-07	6.928776e-07
## G0:0001228	48/943	496/22682	5.526813e-08	9.262938e-06	8.028423e-06
## G0:0140297	45/943	410/22682	3.277608e-09	1.358905e-06	1.177797e-06
## G0:0003779	39/943	437/22682	6.892845e-06	4.813504e-04	4.171985e-04
## G0:0060589	37/943	463/22682	1.227956e-04	4.549962e-03	3.943567e-03
## G0:0008047	35/943	470/22682	6.808507e-04	1.901843e-02	1.648375e-02
## G0:0030695	33/943	422/22682	4.117341e-04	1.327051e-02	1.150188e-02
## G0:0019902	31/943	225/22682	4.864815e-09	1.358905e-06	1.177797e-06
## G0:0001217	31/943	330/22682	2.211942e-05	1.235738e-03	1.071046e-03
## G0:0001227	31/943	330/22682	2.211942e-05	1.235738e-03	1.071046e-03
## G0:0050839	27/943	287/22682	7.147304e-05	3.152337e-03	2.732211e-03
## G0:0051020	26/943	307/22682	5.050214e-04	1.567437e-02	1.358537e-02
## G0:0060090	26/943	337/22682	1.946337e-03	4.077577e-02	3.534139e-02
## G0:0003924	25/943	298/22682	7.355778e-04	1.988433e-02	1.723425e-02
## G0:0019903	24/943	178/22682	3.909953e-07	4.680772e-05	4.056944e-05
## G0:0003714	24/943	180/22682	4.814726e-07	5.043425e-05	4.371264e-05
## G0:0003713	24/943	245/22682	9.471566e-05	3.779606e-03	3.275880e-03
## G0:0008022	23/943	234/22682	1.248796e-04	4.549962e-03	3.943567e-03
## G0:0031267	23/943	270/22682	9.581194e-04	2.433043e-02	2.108780e-02
## G0:0019838	21/943	147/22682	7.869379e-07	7.327266e-05	6.350727e-05
## G0:0051015	21/943	212/22682	2.170885e-04	7.580006e-03	6.569783e-03
## G0:0030674	21/943	239/22682	1.059738e-03	2.611943e-02	2.263837e-02
## G0:0019955	19/943	138/22682	4.552078e-06	3.467856e-04	3.005679e-04
## G0:0005085	19/943	197/22682	5.919400e-04	1.771592e-02	1.535483e-02
## G0:0008013	18/943	92/22682	3.770407e-08	7.899002e-06	6.846265e-06

##

G0:0003712 Daxx/Ncor2/Runx1t1/Irf2bpl/Irf2bp2/Smarca2/Tb1l1xr1/Rbm39/Setd3/Lpin1/Btg2/Tob1/Brd8/Per2/Cited4/Kdm2b/Calcoco1/Ezh1/Smarcd3/
 ## G0:0001228 Sox9/Tcf7l2/Foxf2/Foxi1/Esrrg/Irf4/Mycn/Myc/Atf3/Mef2c/Batf/Zfp523/Creb3l1/Elk3/Nr1i3/Nfatc1/Egr4/Osr2/Ts
 ## G0:0140297 Daxx/Ncor2/Tbx3/Sox9/Tcf7l2/Hes1/Myc/Mef2c/Setd3/Taf7/Tal1/Usf2/Lpin1/
 ## G0:0003779 Myh9/Ptk2/Setd3/Flnc/Mical2/Tns1/Myo1b/Ezr/Lsp1/
 ## G0:0060589 Arhgef37/Gdi2/Katnb1/Srgap1/Wdr41/Atp1b1/Arhgef3/Rapgef1/Iqsec1/Denn
 ## G0:0008047 Daxx/Psme1/Tnfrsf10b/Gdi2/Srgap1/Gas6/Atp1b1/Asap
 ## G0:0030695 Arhgef37/Gdi2/Srgap1/Wdr41/Arhgef3/Rapgef1
 ## G0:0019902 Hsp90b1/F

```

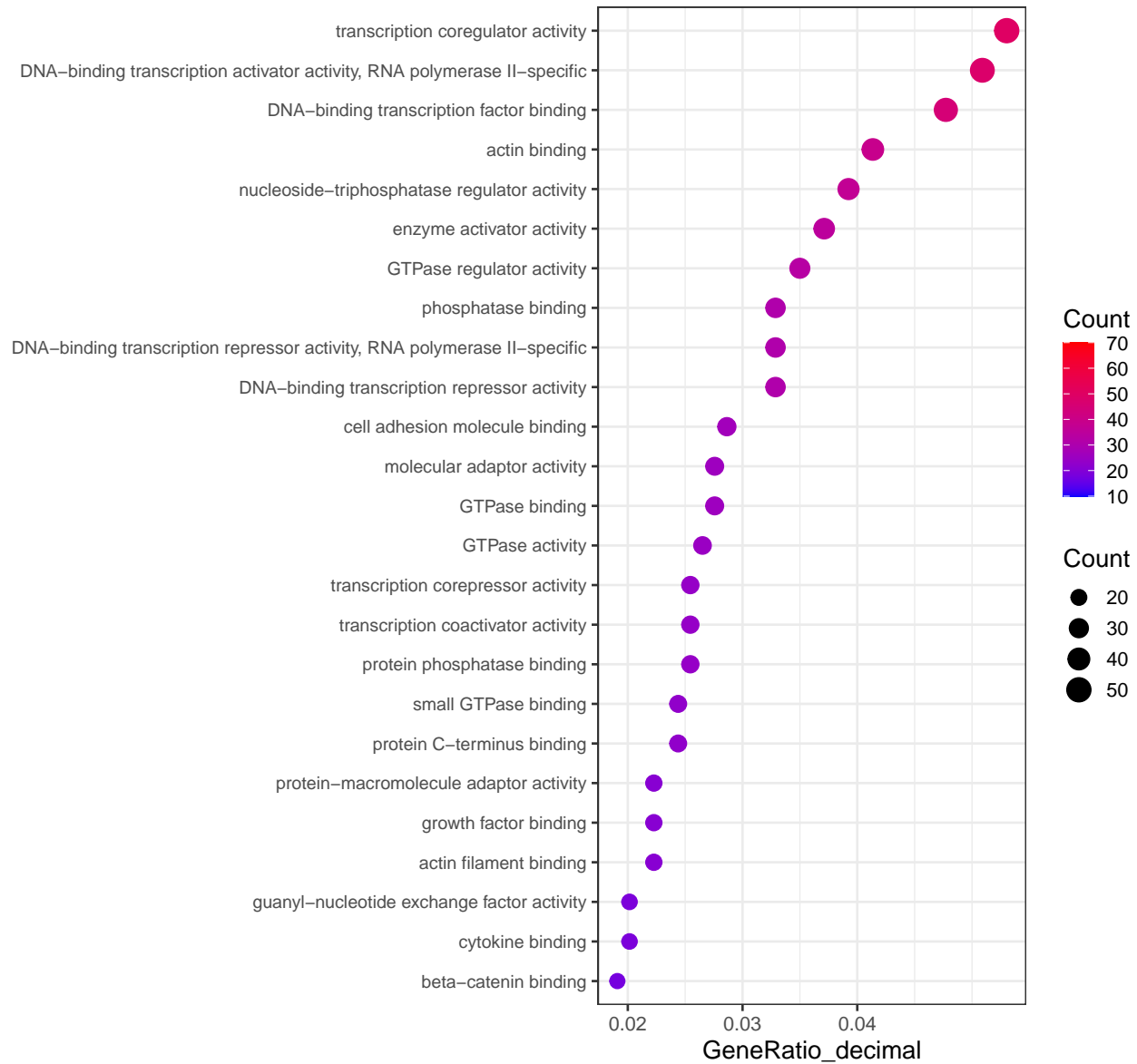
## G0:0001217
## G0:0001227
## G0:0050839
## G0:0051020
## G0:0060090
## G0:0003924
## G0:0019903
## G0:0003714
## G0:0003713
## G0:0008022
## G0:0031267
## G0:0019838
## G0:0051015
## G0:0030674
## G0:0019955
## G0:0005085
## G0:0008013
##
## Count GeneRatio_decimal
## G0:0003712 50 0.05302227
## G0:0001228 48 0.05090138
## G0:0140297 45 0.04772004
## G0:0003779 39 0.04135737
## G0:0060589 37 0.03923648
## G0:0008047 35 0.03711559
## G0:0030695 33 0.03499470
## G0:0019902 31 0.03287381
## G0:0001217 31 0.03287381
## G0:0001227 31 0.03287381
## G0:0050839 27 0.02863203
## G0:0051020 26 0.02757158
## G0:0060090 26 0.02757158
## G0:0003924 25 0.02651113
## G0:0019903 24 0.02545069
## G0:0003714 24 0.02545069
## G0:0003713 24 0.02545069
## G0:0008022 23 0.02439024
## G0:0031267 23 0.02439024
## G0:0019838 21 0.02226935
## G0:0051015 21 0.02226935

```

```
## G0:0030674    21      0.02226935
## G0:0019955    19      0.02014846
## G0:0005085    19      0.02014846
## G0:0008013    18      0.01908802
```

```
ggplot(compG0_MF_diff_df_plot[1:25,], # you can replace the numbers to the
      # row number of pathway of your interest
      aes(x = GeneRatio_decimal, y = reorder(Description, GeneRatio_decimal))) +
  geom_point(aes(size = Count, color = Count)) +
  theme_bw(base_size = 14) +
  scale_size_area()+
  scale_colour_gradient( low="blue", high = "red", limits= c(10,70)) +
  theme(axis.text.y = element_text(size=9.5))+
  scale_x_continuous(breaks=seq(0.02, 0.04, 0.01))+
  ylab(NULL) +
  ggtitle("G0 Molecular Functions \n pathway enrichment")
```

GO Molecular Functions pathway enrichment



```
dev.copy(
  svg,
  file = paste0("./results/GO_MF_pathways.svg"),
  width = 10,
  height = 8
)
```

```
## svg
## 3
```

```
dev.off ()
```

```
## pdf
## 2
```

```
compGO_CC_diff_df_plot <- as.data.frame(compGO_CC_diff)
compGO_CC_diff_df_plot$GeneRatio_decimal <- compGO_CC_diff_df_plot$GeneRatio
compGO_CC_diff_df_plot$GeneRatio_decimal <- sapply(compGO_CC_diff_df_plot$GeneRatio_decimal,
  function(x) (eval(parse(text = as.character(x)))))
compGO_CC_diff_df_plot <- compGO_CC_diff_df_plot %>% arrange(desc(GeneRatio_decimal))
head(compGO_CC_diff_df_plot, 25)
```

```
## ID
## GO:0015629 GO:0015629
## GO:0005667 GO:0005667
## GO:0043235 GO:0043235
## GO:0045121 GO:0045121
## GO:0098857 GO:0098857
## GO:0031252 GO:0031252
## GO:0099572 GO:0099572
## GO:0045177 GO:0045177
## GO:0098984 GO:0098984
## GO:0014069 GO:0014069
## GO:0032279 GO:0032279
## GO:0009925 GO:0009925
## GO:0045178 GO:0045178
## GO:0016324 GO:0016324
```

##	G0:0016323	G0:0016323			
##	G0:0019898	G0:0019898			
##	G0:0044853	G0:0044853			
##	G0:0005938	G0:0005938			
##	G0:0042383	G0:0042383			
##	G0:0090575	G0:0090575			
##	G0:0043209	G0:0043209			
##	G0:0061695	G0:0061695			
##	G0:0098802	G0:0098802			
##	G0:0030426	G0:0030426			
##	G0:0030427	G0:0030427			
##				Description	
##	G0:0015629			actin cytoskeleton	
##	G0:0005667			transcription regulator complex	
##	G0:0043235			receptor complex	
##	G0:0045121			membrane raft	
##	G0:0098857			membrane microdomain	
##	G0:0031252			cell leading edge	
##	G0:0099572			postsynaptic specialization	
##	G0:0045177			apical part of cell	
##	G0:0098984			neuron to neuron synapse	
##	G0:0014069			postsynaptic density	
##	G0:0032279			asymmetric synapse	
##	G0:0009925			basal plasma membrane	
##	G0:0045178			basal part of cell	
##	G0:0016324			apical plasma membrane	
##	G0:0016323			basolateral plasma membrane	
##	G0:0019898			extrinsic component of membrane	
##	G0:0044853			plasma membrane raft	
##	G0:0005938			cell cortex	
##	G0:0042383			sarcolemma	
##	G0:0090575			RNA polymerase II transcription regulator complex	
##	G0:0043209			myelin sheath	
##	G0:0061695			transferase complex, transferring phosphorus-containing groups	
##	G0:0098802			plasma membrane signaling receptor complex	
##	G0:0030426			growth cone	
##	G0:0030427			site of polarized growth	
##		GeneRatio	BgRatio	pvalue	p.adjust
##	G0:0015629	49/953	500/23177	1.932900e-08	2.249895e-06
				qvalue	
				1.900346e-06	

##	G0:0005667	46/953	469/23177	5.135432e-08	4.981369e-06	4.207451e-06
##	G0:0043235	44/953	414/23177	9.466190e-09	1.377331e-06	1.163345e-06
##	G0:0045121	42/953	381/23177	7.086608e-09	1.377331e-06	1.163345e-06
##	G0:0098857	42/953	382/23177	7.654643e-09	1.377331e-06	1.163345e-06
##	G0:0031252	36/953	396/23177	8.133115e-06	3.641133e-04	3.075437e-04
##	G0:0099572	36/953	439/23177	7.265770e-05	2.013656e-03	1.700809e-03
##	G0:0045177	36/953	460/23177	1.839239e-04	4.583612e-03	3.871490e-03
##	G0:0098984	35/953	433/23177	1.207904e-04	3.195455e-03	2.699001e-03
##	G0:0014069	32/953	400/23177	2.789173e-04	5.236448e-03	4.422900e-03
##	G0:0032279	32/953	405/23177	3.470414e-04	6.311816e-03	5.331195e-03
##	G0:0009925	30/953	265/23177	5.626294e-07	4.093129e-05	3.457209e-05
##	G0:0045178	30/953	288/23177	3.196205e-06	1.860191e-04	1.571187e-04
##	G0:0016324	28/953	354/23177	7.781376e-04	1.161221e-02	9.808104e-03
##	G0:0016323	27/953	232/23177	1.195206e-06	7.728999e-05	6.528201e-05
##	G0:0019898	25/953	323/23177	1.948780e-03	2.577705e-02	2.177226e-02
##	G0:0044853	23/953	131/23177	3.692566e-09	1.377331e-06	1.163345e-06
##	G0:0005938	23/953	306/23177	4.143842e-03	4.637915e-02	3.917357e-02
##	G0:0042383	21/953	166/23177	4.852758e-06	2.567550e-04	2.168649e-04
##	G0:0090575	21/953	169/23177	6.448158e-06	3.127357e-04	2.641482e-04
##	G0:0043209	21/953	213/23177	1.999198e-04	4.654133e-03	3.931055e-03
##	G0:0061695	20/953	241/23177	2.386119e-03	3.086047e-02	2.606591e-02
##	G0:0098802	19/953	187/23177	2.693050e-04	5.236448e-03	4.422900e-03
##	G0:0030426	18/953	209/23177	2.581792e-03	3.266528e-02	2.759031e-02
##	G0:0030427	18/953	216/23177	3.670951e-03	4.332813e-02	3.659656e-02
##						
##	G0:0015629	Myh9/Dctn5/Ptk2/Ppp1r12a/Ints6/Flnc/Cdh2/Dgkh/Micalc1/Akap13/Myo1b/Ezr/Asap1/Enah/My16b/My16/Abl1/Coro1c/My19/Epha3/Ehbp1/Ab				
##	G0:0005667	Ncor2/Tbx3/Sox9/Tcf712/Foxf2/Gatad2b/Zbtb16/Tb11xr1/E2f7/Smad7/Taf7/Tal1/Lpin1/Crem/Nfatc1/Tcf71				
##	G0:0043235	Scrib/Plxnb2/Itga5/Plxna3/Ryk/Syk/Fgfr1/Itpr2/Insr/Plxnc1/Tgfbr1/Bmpr1a/Ghr/Il27ra/Crlf2/Il6ra/Il11ra				
##	G0:0045121	Kirrel/Plpp2/Tnfrsf23/Tnfrsf10b/Insr/Atp1b1/Cblb/Tgfbr1/Cdh2/Bmpr1a/Tnr/Ezr/Hmox1/Ppp2ca/Prkar2a/Tgf				
##	G0:0098857	Kirrel/Plpp2/Tnfrsf23/Tnfrsf10b/Insr/Atp1b1/Cblb/Tgfbr1/Cdh2/Bmpr1a/Tnr/Ezr/Hmox1/Ppp2ca/Prkar2a/Tgf				
##	G0:0031252	Scrib/Myh9/Nradd/Ptk2/Itga5/Aak1/Insr/Dg				
##	G0:0099572	Scrib/Arrb2/Picalm/Dgkz/Iqsec1/Cdh2/Numb/Synj2bp/Ppp2ca/Erc1				
##	G0:0045177	Dstyk/Scrib/Pard6b/Ptk2/Atp1b1/Tgfbr1/Cdh2/Slc7a13/				
##	G0:0098984	Scrib/Arrb2/Picalm/Dgkz/Iqsec1/Cdh2/Tnr/Sh3gl2/Numb				
##	G0:0014069	Scrib/Arrb2/Picalm/Dgkz/Iqsec1/Cd				
##	G0:0032279	Scrib/Arrb2/Picalm/Dgkz/Iqsec1/Cd				
##	G0:0009925	Dstyk/Scrib/				
##	G0:0045178	Dstyk/Scrib/				
##	G0:0016324	Dstyk/				


```

## G0:0016323
## G0:0019898
## G0:0044853
## G0:0005938
## G0:0042383
## G0:0090575
## G0:0043209
## G0:0061695
## G0:0098802
## G0:0030426
## G0:0030427
##
## Count GeneRatio_decimal
## G0:0015629 49 0.05141658
## G0:0005667 46 0.04826863
## G0:0043235 44 0.04616999
## G0:0045121 42 0.04407135
## G0:0098857 42 0.04407135
## G0:0031252 36 0.03777545
## G0:0099572 36 0.03777545
## G0:0045177 36 0.03777545
## G0:0098984 35 0.03672613
## G0:0014069 32 0.03357817
## G0:0032279 32 0.03357817
## G0:0009925 30 0.03147954
## G0:0045178 30 0.03147954
## G0:0016324 28 0.02938090
## G0:0016323 27 0.02833158
## G0:0019898 25 0.02623295
## G0:0044853 23 0.02413431
## G0:0005938 23 0.02413431
## G0:0042383 21 0.02203568
## G0:0090575 21 0.02203568
## G0:0043209 21 0.02203568
## G0:0061695 20 0.02098636
## G0:0098802 19 0.01993704
## G0:0030426 18 0.01888772
## G0:0030427 18 0.01888772

```

```
compG0_CC_diff_df_plot[1:25,]
```

##	ID	
##	G0:0015629	G0:0015629
##	G0:0005667	G0:0005667
##	G0:0043235	G0:0043235
##	G0:0045121	G0:0045121
##	G0:0098857	G0:0098857
##	G0:0031252	G0:0031252
##	G0:0099572	G0:0099572
##	G0:0045177	G0:0045177
##	G0:0098984	G0:0098984
##	G0:0014069	G0:0014069
##	G0:0032279	G0:0032279
##	G0:0009925	G0:0009925
##	G0:0045178	G0:0045178
##	G0:0016324	G0:0016324
##	G0:0016323	G0:0016323
##	G0:0019898	G0:0019898
##	G0:0044853	G0:0044853
##	G0:0005938	G0:0005938
##	G0:0042383	G0:0042383
##	G0:0090575	G0:0090575
##	G0:0043209	G0:0043209
##	G0:0061695	G0:0061695
##	G0:0098802	G0:0098802
##	G0:0030426	G0:0030426
##	G0:0030427	G0:0030427
##		Description
##	G0:0015629	actin cytoskeleton
##	G0:0005667	transcription regulator complex
##	G0:0043235	receptor complex
##	G0:0045121	membrane raft
##	G0:0098857	membrane microdomain
##	G0:0031252	cell leading edge
##	G0:0099572	postsynaptic specialization
##	G0:0045177	apical part of cell
##	G0:0098984	neuron to neuron synapse

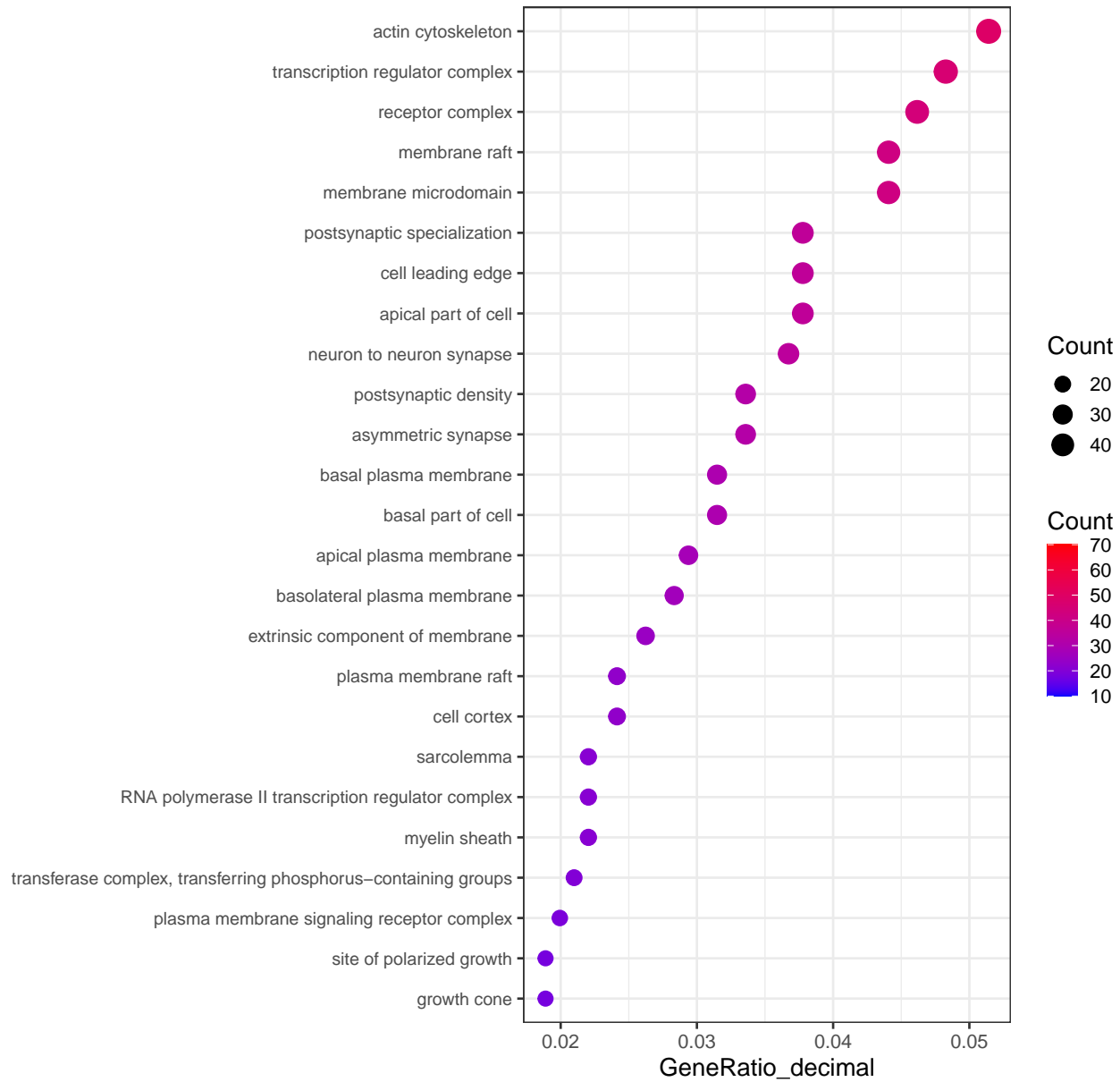
## G0:0014069						postsynaptic density
## G0:0032279						asymmetric synapse
## G0:0009925						basal plasma membrane
## G0:0045178						basal part of cell
## G0:0016324						apical plasma membrane
## G0:0016323						basolateral plasma membrane
## G0:0019898						extrinsic component of membrane
## G0:0044853						plasma membrane raft
## G0:0005938						cell cortex
## G0:0042383						sarcolemma
## G0:0090575						RNA polymerase II transcription regulator complex
## G0:0043209						myelin sheath
## G0:0061695						transferase complex, transferring phosphorus-containing groups
## G0:0098802						plasma membrane signaling receptor complex
## G0:0030426						growth cone
## G0:0030427						site of polarized growth
##	GeneRatio	BgRatio	pvalue	p.adjust	qvalue	
## G0:0015629	49/953	500/23177	1.932900e-08	2.249895e-06	1.900346e-06	
## G0:0005667	46/953	469/23177	5.135432e-08	4.981369e-06	4.207451e-06	
## G0:0043235	44/953	414/23177	9.466190e-09	1.377331e-06	1.163345e-06	
## G0:0045121	42/953	381/23177	7.086608e-09	1.377331e-06	1.163345e-06	
## G0:0098857	42/953	382/23177	7.654643e-09	1.377331e-06	1.163345e-06	
## G0:0031252	36/953	396/23177	8.133115e-06	3.641133e-04	3.075437e-04	
## G0:0099572	36/953	439/23177	7.265770e-05	2.013656e-03	1.700809e-03	
## G0:0045177	36/953	460/23177	1.839239e-04	4.583612e-03	3.871490e-03	
## G0:0098984	35/953	433/23177	1.207904e-04	3.195455e-03	2.699001e-03	
## G0:0014069	32/953	400/23177	2.789173e-04	5.236448e-03	4.422900e-03	
## G0:0032279	32/953	405/23177	3.470414e-04	6.311816e-03	5.331195e-03	
## G0:0009925	30/953	265/23177	5.626294e-07	4.093129e-05	3.457209e-05	
## G0:0045178	30/953	288/23177	3.196205e-06	1.860191e-04	1.571187e-04	
## G0:0016324	28/953	354/23177	7.781376e-04	1.161221e-02	9.808104e-03	
## G0:0016323	27/953	232/23177	1.195206e-06	7.728999e-05	6.528201e-05	
## G0:0019898	25/953	323/23177	1.948780e-03	2.577705e-02	2.177226e-02	
## G0:0044853	23/953	131/23177	3.692566e-09	1.377331e-06	1.163345e-06	
## G0:0005938	23/953	306/23177	4.143842e-03	4.637915e-02	3.917357e-02	
## G0:0042383	21/953	166/23177	4.852758e-06	2.567550e-04	2.168649e-04	
## G0:0090575	21/953	169/23177	6.448158e-06	3.127357e-04	2.641482e-04	
## G0:0043209	21/953	213/23177	1.999198e-04	4.654133e-03	3.931055e-03	
## G0:0061695	20/953	241/23177	2.386119e-03	3.086047e-02	2.606591e-02	

## G0:0098802	19/953	187/23177	2.693050e-04	5.236448e-03	4.422900e-03
## G0:0030426	18/953	209/23177	2.581792e-03	3.266528e-02	2.759031e-02
## G0:0030427	18/953	216/23177	3.670951e-03	4.332813e-02	3.659656e-02
##					
## G0:0015629	Myh9/Dctn5/Ptk2/Ppp1r12a/Ints6/Flnc/Cdh2/Dgkh/Micalcl/Akap13/Myo1b/Ezr/Asap1/Enah/Myl6b/Myl6/Abl1/Coro1c/Myl9/Epha3/Ehbp1/Ab				
## G0:0005667	Ncor2/Tbx3/Sox9/Tcf7l2/Foxf2/Gatad2b/Zbtb16/Tbl1xr1/E2f7/Smad7/Taf7/Tal1/Lpin1/Crem/Nfatc1/Tcf7l				
## G0:0043235	Scrib/Plxnb2/Itga5/Plxna3/Ryk/Syk/Fgfr1/Itpr2/Insr/Plxnc1/Tgfbr1/Bmpr1a/Ghr/Il27ra/Crlf2/Il6ra/Il11ra				
## G0:0045121	Kirrel/Plpp2/Tnfrrsf23/Tnfrrsf10b/Insr/Atp1b1/Cblb/Tgfbr1/Cdh2/Bmpr1a/Tnr/Ezr/Hmox1/Ppp2ca/Prkar2a/Tgf				
## G0:0098857	Kirrel/Plpp2/Tnfrrsf23/Tnfrrsf10b/Insr/Atp1b1/Cblb/Tgfbr1/Cdh2/Bmpr1a/Tnr/Ezr/Hmox1/Ppp2ca/Prkar2a/Tgf				
## G0:0031252	Scrib/Myh9/Nradd/Ptk2/Itga5/Aak1/Insr/Dg				
## G0:0099572	Scrib/Arrb2/Picalm/Dgkz/Iqsec1/Cdh2/Numb/Synj2bp/Ppp2ca/Erc1				
## G0:0045177	Dstyky/Scrib/Pard6b/Ptk2/Atp1b1/Tgfbr1/Cdh2/Slc7a13/				
## G0:0098984	Scrib/Arrb2/Picalm/Dgkz/Iqsec1/Cdh2/Tnr/Sh3gl2/Numb				
## G0:0014069	Scrib/Arrb2/Picalm/Dgkz/Iqsec1/Cd				
## G0:0032279	Scrib/Arrb2/Picalm/Dgkz/Iqsec1/Cd				
## G0:0009925	Dstyky/Scrib/				
## G0:0045178	Dstyky/Scrib/				
## G0:0016324	Dstyky/				
## G0:0016323					
## G0:0019898					
## G0:0044853					
## G0:0005938					
## G0:0042383					
## G0:0090575					
## G0:0043209					
## G0:0061695					
## G0:0098802					
## G0:0030426					
## G0:0030427					
##	Count	GeneRatio_decimal			
## G0:0015629	49	0.05141658			
## G0:0005667	46	0.04826863			
## G0:0043235	44	0.04616999			
## G0:0045121	42	0.04407135			
## G0:0098857	42	0.04407135			
## G0:0031252	36	0.03777545			
## G0:0099572	36	0.03777545			
## G0:0045177	36	0.03777545			
## G0:0098984	35	0.03672613			

## G0:0014069	32	0.03357817
## G0:0032279	32	0.03357817
## G0:0009925	30	0.03147954
## G0:0045178	30	0.03147954
## G0:0016324	28	0.02938090
## G0:0016323	27	0.02833158
## G0:0019898	25	0.02623295
## G0:0044853	23	0.02413431
## G0:0005938	23	0.02413431
## G0:0042383	21	0.02203568
## G0:0090575	21	0.02203568
## G0:0043209	21	0.02203568
## G0:0061695	20	0.02098636
## G0:0098802	19	0.01993704
## G0:0030426	18	0.01888772
## G0:0030427	18	0.01888772

```
ggplot(compG0_CC_diff_df_plot[1:25,], # you can replace the numbers to the row number of pathway of your interest
  aes(x = GeneRatio_decimal, y = reorder(Description, GeneRatio_decimal))) +
  geom_point(aes(size = Count, color = Count)) +
  scale_size_area()+
  theme_bw(base_size = 14) +
  scale_colour_gradient( low="blue", high = "red", limits= c(10,70)) +
  theme(axis.text.y = element_text(size=9.5))+
  ylab(NULL) +
  ggtitle("G0 Cellular Component pathways enrichment")
```

GO Cellular Component pathways enrichment



```
dev.copy(
  svg,
  file = paste0("./results/GO_CC_pathways.svg"),
  width = 10,
  height = 8
)
```

```
## svg
## 3
```

```
dev.off ()
```

```
## pdf
## 2
```

```
compGO_BP_diff_df_plot <- as.data.frame(compGO_BP_diff)
compGO_BP_diff_df_plot$GeneRatio_decimal <- compGO_BP_diff_df_plot$GeneRatio
compGO_BP_diff_df_plot$GeneRatio_decimal <- sapply(compGO_BP_diff_df_plot$GeneRatio_decimal,
  function(x) (eval(parse(text = as.character(x)))))
compGO_BP_diff_df_plot <- compGO_BP_diff_df_plot %>% arrange(desc(GeneRatio_decimal))
head(compGO_BP_diff_df_plot, 25)
```

```
## ID
## GO:0060537 GO:0060537
## GO:0014706 GO:0014706
## GO:0042692 GO:0042692
## GO:0001667 GO:0001667
## GO:0060562 GO:0060562
## GO:0048638 GO:0048638
## GO:0045786 GO:0045786
## GO:0001503 GO:0001503
## GO:0048608 GO:0048608
## GO:0061458 GO:0061458
## GO:0007346 GO:0007346
## GO:0007264 GO:0007264
## GO:0007409 GO:0007409
## GO:0051146 GO:0051146
```

##	G0:0040013	G0:0040013
##	G0:0003012	G0:0003012
##	G0:0001558	G0:0001558
##	G0:0050767	G0:0050767
##	G0:0001655	G0:0001655
##	G0:0016055	G0:0016055
##	G0:0045860	G0:0045860
##	G0:0198738	G0:0198738
##	G0:0060485	G0:0060485
##	G0:0007517	G0:0007517
##	G0:0007178	G0:0007178
##		Description
##	G0:0060537	muscle tissue development
##	G0:0014706	striated muscle tissue development
##	G0:0042692	muscle cell differentiation
##	G0:0001667	ameboidal-type cell migration
##	G0:0060562	epithelial tube morphogenesis
##	G0:0048638	regulation of developmental growth
##	G0:0045786	negative regulation of cell cycle
##	G0:0001503	ossification
##	G0:0048608	reproductive structure development
##	G0:0061458	reproductive system development
##	G0:0007346	regulation of mitotic cell cycle
##	G0:0007264	small GTPase mediated signal transduction
##	G0:0007409	axonogenesis
##	G0:0051146	striated muscle cell differentiation
##	G0:0040013	negative regulation of locomotion
##	G0:0003012	muscle system process
##	G0:0001558	regulation of cell growth
##	G0:0050767	regulation of neurogenesis
##	G0:0001655	urogenital system development
##	G0:0016055	Wnt signaling pathway
##	G0:0045860	positive regulation of protein kinase activity
##	G0:0198738	cell-cell signaling by wnt
##	G0:0060485	mesenchyme development
##	G0:0007517	muscle organ development
##	G0:0007178	transmembrane receptor protein serine/threonine kinase signaling pathway
##	GeneRatio	BgRatio pvalue p.adjust qvalue
##	G0:0060537	68/949 477/23355 1.407660e-19 7.273378e-16 5.169816e-16

## G0:0014706	61/949	453/23355	1.783595e-16	4.607917e-13	3.275243e-13
## G0:0042692	57/949	410/23355	4.247137e-16	7.314986e-13	5.199390e-13
## G0:0001667	52/949	439/23355	5.032217e-12	2.166789e-09	1.540123e-09
## G0:0060562	51/949	393/23355	2.432551e-13	2.094832e-10	1.488977e-10
## G0:0048638	51/949	403/23355	6.489102e-13	4.789884e-10	3.404583e-10
## G0:0045786	51/949	452/23355	4.884593e-11	1.577418e-08	1.121207e-08
## G0:0001503	49/949	406/23355	1.061953e-11	4.220857e-09	3.000126e-09
## G0:0048608	47/949	478/23355	2.387319e-08	2.739814e-06	1.947422e-06
## G0:0061458	47/949	482/23355	3.084292e-08	3.201275e-06	2.275422e-06
## G0:0007346	46/949	453/23355	1.260073e-08	1.770242e-06	1.258263e-06
## G0:0007264	45/949	439/23355	1.364902e-08	1.831699e-06	1.301946e-06
## G0:0007409	45/949	483/23355	2.259431e-07	1.477782e-05	1.050387e-05
## G0:0051146	44/949	322/23355	1.852401e-12	1.176329e-09	8.361180e-10
## G0:0040013	43/949	332/23355	1.994414e-11	6.870090e-09	4.883164e-09
## G0:0003012	43/949	412/23355	1.694173e-08	2.084236e-06	1.481446e-06
## G0:0001558	43/949	439/23355	1.049754e-07	7.860982e-06	5.587477e-06
## G0:0050767	43/949	443/23355	1.353766e-07	9.470742e-06	6.731671e-06
## G0:0001655	42/949	379/23355	4.317256e-09	8.261949e-07	5.872478e-07
## G0:0016055	42/949	438/23355	2.633732e-07	1.680061e-05	1.194164e-05
## G0:0045860	42/949	439/23355	2.801307e-07	1.723137e-05	1.224782e-05
## G0:0198738	42/949	440/23355	2.978828e-07	1.789722e-05	1.272109e-05
## G0:0060485	41/949	288/23355	2.784093e-12	1.438541e-09	1.022495e-09
## G0:0007517	41/949	357/23355	2.317436e-09	5.206171e-07	3.700473e-07
## G0:0007178	41/949	381/23355	1.534828e-08	1.982613e-06	1.409214e-06
##					
## G0:0060537	Tbx3/Rcan1/Sox9/Usp19/Tcf712/Srp3/Myc/Tcap/Atf3/Mef2c/Nr1d2/Fgfr1/Arrb2/Lmna/Smad7/Tgfbr1/Bmpr1a/Bmp4/Btg2/Dsp/Akap13/Igfbp				
## G0:0014706	Tbx3/Rcan1/Usp19/Tcf712/Srp3/Myc/Tcap/Atf3/Mef2c/Nr1d2/Fgfr1/Arrb2/Lmna/Smad7/				
## G0:0042692	Thra/Daxx/Myh9/Tbx3/Rcan1/Sox9/Hes1/Tcap/Smad7/Mef2c/Nr1d2/Fgfr1/Arrb2/Lmna/Smad7/				
## G0:0001667	Myh9/Sox9/Ptk2/Mef2c/Fgfr1/Iqsec1/				
## G0:0060562	Scrib/Tbx3/Sox9/				
## G0:0048638	Daxx/Tcf712/Ptk2/Cel1/				
## G0:0045786	Scrib/Tcf712/Rassf1/Celf1/Hes1/				
## G0:0001503	Thra/Sox9/Tcf712/Tnfrsf1/				
## G0:0048608	Tb				
## G0:0061458	Tb				
## G0:0007346					
## G0:0007264	Ar				
## G0:0007409					
## G0:0051146					

```

## G0:0040013
## G0:0003012
## G0:0001558
## G0:0050767
## G0:0001655
## G0:0016055
## G0:0045860
## G0:0198738
## G0:0060485
## G0:0007517
## G0:0007178
##      Count GeneRatio_decimal
## G0:0060537    68      0.07165437
## G0:0014706    61      0.06427819
## G0:0042692    57      0.06006322
## G0:0001667    52      0.05479452
## G0:0060562    51      0.05374078
## G0:0048638    51      0.05374078
## G0:0045786    51      0.05374078
## G0:0001503    49      0.05163330
## G0:0048608    47      0.04952582
## G0:0061458    47      0.04952582
## G0:0007346    46      0.04847208
## G0:0007264    45      0.04741834
## G0:0007409    45      0.04741834
## G0:0051146    44      0.04636459
## G0:0040013    43      0.04531085
## G0:0003012    43      0.04531085
## G0:0001558    43      0.04531085
## G0:0050767    43      0.04531085
## G0:0001655    42      0.04425711
## G0:0016055    42      0.04425711
## G0:0045860    42      0.04425711
## G0:0198738    42      0.04425711
## G0:0060485    41      0.04320337
## G0:0007517    41      0.04320337
## G0:0007178    41      0.04320337

```

```
compG0_BP_diff_df_plot[1:25,]
```

##	ID	
##	G0:0060537	G0:0060537
##	G0:0014706	G0:0014706
##	G0:0042692	G0:0042692
##	G0:0001667	G0:0001667
##	G0:0060562	G0:0060562
##	G0:0048638	G0:0048638
##	G0:0045786	G0:0045786
##	G0:0001503	G0:0001503
##	G0:0048608	G0:0048608
##	G0:0061458	G0:0061458
##	G0:0007346	G0:0007346
##	G0:0007264	G0:0007264
##	G0:0007409	G0:0007409
##	G0:0051146	G0:0051146
##	G0:0040013	G0:0040013
##	G0:0003012	G0:0003012
##	G0:0001558	G0:0001558
##	G0:0050767	G0:0050767
##	G0:0001655	G0:0001655
##	G0:0016055	G0:0016055
##	G0:0045860	G0:0045860
##	G0:0198738	G0:0198738
##	G0:0060485	G0:0060485
##	G0:0007517	G0:0007517
##	G0:0007178	G0:0007178
##		Description
##	G0:0060537	muscle tissue development
##	G0:0014706	striated muscle tissue development
##	G0:0042692	muscle cell differentiation
##	G0:0001667	ameboidal-type cell migration
##	G0:0060562	epithelial tube morphogenesis
##	G0:0048638	regulation of developmental growth
##	G0:0045786	negative regulation of cell cycle
##	G0:0001503	ossification
##	G0:0048608	reproductive structure development

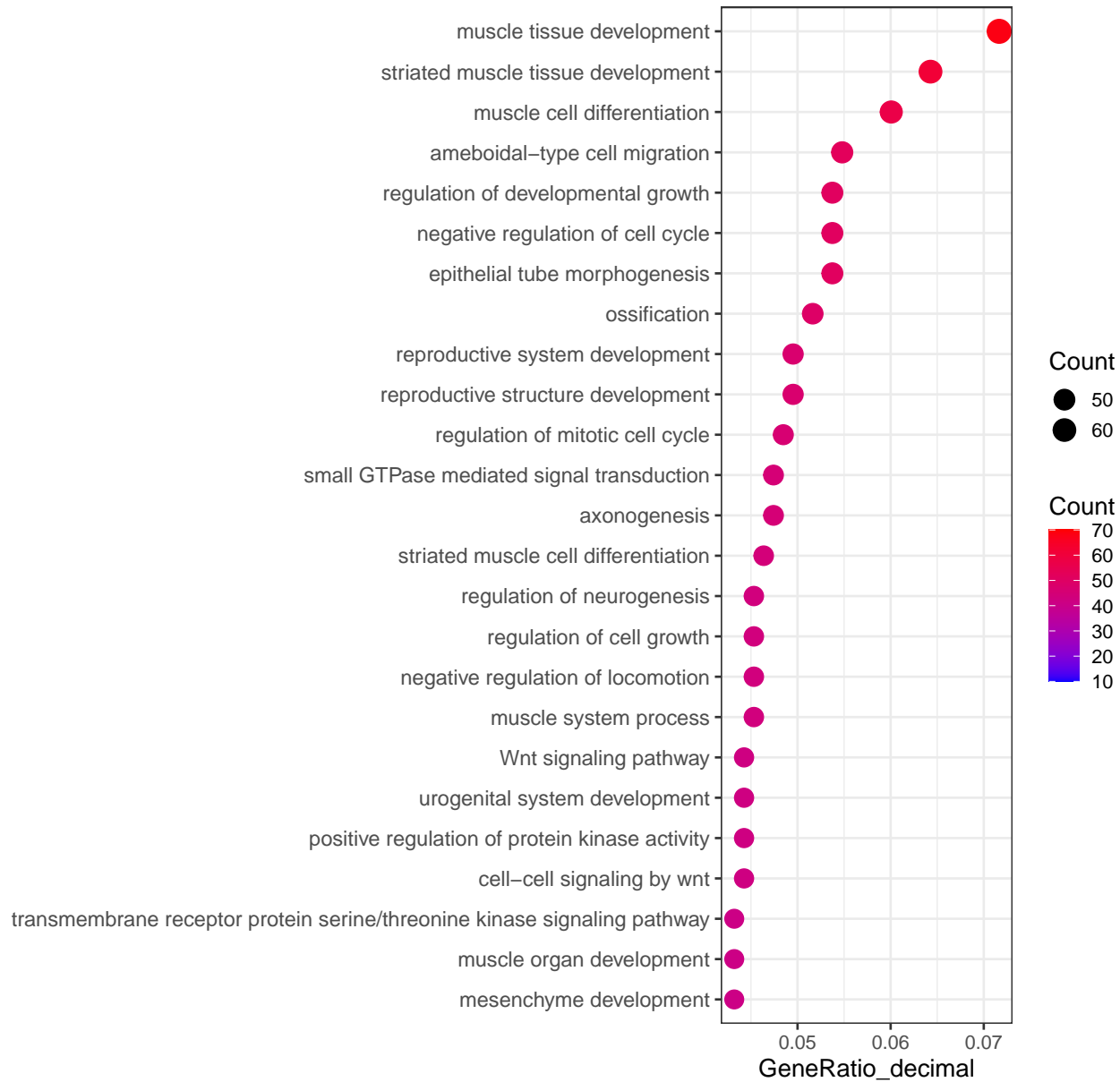
## G0:0061458	reproductive system development
## G0:0007346	regulation of mitotic cell cycle
## G0:0007264	small GTPase mediated signal transduction
## G0:0007409	axonogenesis
## G0:0051146	striated muscle cell differentiation
## G0:0040013	negative regulation of locomotion
## G0:0003012	muscle system process
## G0:0001558	regulation of cell growth
## G0:0050767	regulation of neurogenesis
## G0:0001655	urogenital system development
## G0:0016055	Wnt signaling pathway
## G0:0045860	positive regulation of protein kinase activity
## G0:0198738	cell-cell signaling by wnt
## G0:0060485	mesenchyme development
## G0:0007517	muscle organ development
## G0:0007178	transmembrane receptor protein serine/threonine kinase signaling pathway
##	GeneRatio BgRatio pvalue p.adjust qvalue
## G0:0060537	68/949 477/23355 1.407660e-19 7.273378e-16 5.169816e-16
## G0:0014706	61/949 453/23355 1.783595e-16 4.607917e-13 3.275243e-13
## G0:0042692	57/949 410/23355 4.247137e-16 7.314986e-13 5.199390e-13
## G0:0001667	52/949 439/23355 5.032217e-12 2.166789e-09 1.540123e-09
## G0:0060562	51/949 393/23355 2.432551e-13 2.094832e-10 1.488977e-10
## G0:0048638	51/949 403/23355 6.489102e-13 4.789884e-10 3.404583e-10
## G0:0045786	51/949 452/23355 4.884593e-11 1.577418e-08 1.121207e-08
## G0:0001503	49/949 406/23355 1.061953e-11 4.220857e-09 3.000126e-09
## G0:0048608	47/949 478/23355 2.387319e-08 2.739814e-06 1.947422e-06
## G0:0061458	47/949 482/23355 3.084292e-08 3.201275e-06 2.275422e-06
## G0:0007346	46/949 453/23355 1.260073e-08 1.770242e-06 1.258263e-06
## G0:0007264	45/949 439/23355 1.364902e-08 1.831699e-06 1.301946e-06
## G0:0007409	45/949 483/23355 2.259431e-07 1.477782e-05 1.050387e-05
## G0:0051146	44/949 322/23355 1.852401e-12 1.176329e-09 8.361180e-10
## G0:0040013	43/949 332/23355 1.994414e-11 6.870090e-09 4.883164e-09
## G0:0003012	43/949 412/23355 1.694173e-08 2.084236e-06 1.481446e-06
## G0:0001558	43/949 439/23355 1.049754e-07 7.860982e-06 5.587477e-06
## G0:0050767	43/949 443/23355 1.353766e-07 9.470742e-06 6.731671e-06
## G0:0001655	42/949 379/23355 4.317256e-09 8.261949e-07 5.872478e-07
## G0:0016055	42/949 438/23355 2.633732e-07 1.680061e-05 1.194164e-05
## G0:0045860	42/949 439/23355 2.801307e-07 1.723137e-05 1.224782e-05
## G0:0198738	42/949 440/23355 2.978828e-07 1.789722e-05 1.272109e-05

## G0:0060485	41/949	288/23355	2.784093e-12	1.438541e-09	1.022495e-09	
## G0:0007517	41/949	357/23355	2.317436e-09	5.206171e-07	3.700473e-07	
## G0:0007178	41/949	381/23355	1.534828e-08	1.982613e-06	1.409214e-06	
##						
## G0:0060537	Tbx3/Rcan1/Sox9/Usp19/Tcf7l2/Srpk3/Myc/Tcap/Atf3/Mef2c/Nr1d2/Fgfr1/Arrb2/Lmna/Smad7/Tgfbr1/Bmpr1a/Bmp4/Btg2/Dsp/Akap13/Igfbp					
## G0:0014706			Tbx3/Rcan1/Usp19/Tcf7l2/Srpk3/Myc/Tcap/Atf3/Mef2c/Nr1d2/Fgfr1/Arrb2/Lmna/Smad7/			
## G0:0042692				Thra/Daxx/Myh9/Tbx3/Rcan1/Sox9/Hes1/Tcap/Smad7/Mef		
## G0:0001667					Myh9/Sox9/Ptk2/Mef2c/Fgfr1/Iqsec1/	
## G0:0060562						Scrib/Tbx3/Sox
## G0:0048638						Daxx/Tcf7l2/Ptk2/Cel
## G0:0045786						Scrib/Tcf7l2/Rassf1/Celf1/H
## G0:0001503						Thra/Sox9/Tcf7l2/Tnfrsf1
## G0:0048608						Tb
## G0:0061458						Tb
## G0:0007346						
## G0:0007264						Ar
## G0:0007409						
## G0:0051146						
## G0:0040013						
## G0:0003012						
## G0:0001558						
## G0:0050767						
## G0:0001655						
## G0:0016055						
## G0:0045860						
## G0:0198738						
## G0:0060485						
## G0:0007517						
## G0:0007178						
##	Count	GeneRatio_decimal				
## G0:0060537	68	0.07165437				
## G0:0014706	61	0.06427819				
## G0:0042692	57	0.06006322				
## G0:0001667	52	0.05479452				
## G0:0060562	51	0.05374078				
## G0:0048638	51	0.05374078				
## G0:0045786	51	0.05374078				
## G0:0001503	49	0.05163330				
## G0:0048608	47	0.04952582				

## G0:0061458	47	0.04952582
## G0:0007346	46	0.04847208
## G0:0007264	45	0.04741834
## G0:0007409	45	0.04741834
## G0:0051146	44	0.04636459
## G0:0040013	43	0.04531085
## G0:0003012	43	0.04531085
## G0:0001558	43	0.04531085
## G0:0050767	43	0.04531085
## G0:0001655	42	0.04425711
## G0:0016055	42	0.04425711
## G0:0045860	42	0.04425711
## G0:0198738	42	0.04425711
## G0:0060485	41	0.04320337
## G0:0007517	41	0.04320337
## G0:0007178	41	0.04320337

```
ggplot(compG0_BP_diff_df_plot[1:25,], # you can replace the numbers to the row number of pathway of your interest
  aes(x = GeneRatio_decimal, y = reorder(Description, GeneRatio_decimal))) +
  geom_point(aes(size = Count, color = Count)) +
  scale_size_area()+
  theme_bw(base_size = 14) +
  scale_colour_gradient( low="blue", high = "red", limits= c(10,70)) +
  theme(axis.text.y = element_text(size=12))+
  ylab(NULL) +
  ggtitle("G0 Biological Pathways enrichment")
```

GO Biological Pathways enrichr



```
dev.copy(
  svg,
  file = paste0("./results/GO_BP_pathways.svg"),
  width = 10,
  height = 8
)
```

```
## svg
## 3
```

```
dev.off ()
```

```
## pdf
## 2
```

```
compKEGG_diff_readable <- setReadable(compKEGG_diff, OrgDb = org.Mm.eg.db, keyType="ENTREZID")
## The geneID column is translated to symbol
head(compKEGG_diff_readable, 3)
```

```
##           ID           Description GeneRatio  BgRatio      pvalue
## mmu05206 mmu05206   MicroRNAs in cancer   41/447 303/8924 4.402715e-09
## mmu04360 mmu04360       Axon guidance    28/447 181/8924 8.092458e-08
## mmu04010 mmu04010 MAPK signaling pathway  37/447 294/8924 1.756819e-07
##           p.adjust      qvalue
## mmu05206 1.289996e-06 9.176185e-07
## mmu04360 1.185545e-05 8.433193e-06
## mmu04010 1.715827e-05 1.220527e-05
##
## mmu05206 Stmn1/Rassf1/Dnmt3a/Hras/Itga5/Myc/Zeb1/Pak4/Tnr/Ezr/Hmox1/Efna1/Abl1/Plcg1/Zfpm2/Trp63/Notch1/Egfr/E2f3/Hdac5/Met/Trim71/Ccng
## mmu04360                                     Plxnb2/Pard6b/Ptk2/Hras/Plxna3/Ryk/Sr
## mmu04010                                     Daxx/Stmn1/Hras/Myc/Mef2c/Fgfr1/Arrb2/Insr/Tgfbr1/Flnc/Hspa2/Gna12/Gadd45b/Efna1/Tgfbr2/Rasa1/M
##           Count
## mmu05206     41
## mmu04360     28
## mmu04010     37
```



```
compKEGG_diff_readable_df <- as.data.frame(compKEGG_diff_readable)

compKEGG_diff_readable_df$GeneRatio_decimal <- compKEGG_diff_readable_df$GeneRatio
compKEGG_diff_readable_df$GeneRatio_decimal <- sapply(compKEGG_diff_readable_df$GeneRatio_decimal, function(x) (eval(parse(text = as.character(x)))))
compKEGG_diff_readable_df$BgRatio_decimal <- compKEGG_diff_readable_df$BgRatio
compKEGG_diff_readable_df$BgRatio_decimal <- sapply(compKEGG_diff_readable_df$BgRatio_decimal, function(x) (eval(parse(text = as.character(x)))))
compKEGG_diff_readable_df <- compKEGG_diff_readable_df %>% tidyr::separate_rows(geneID, sep = "/", convert = FALSE) %>% arrange(desc(GeneRatio_decimal))
compKEGG_diff_readable_df %>% head
```

```
## # A tibble: 6 x 11
##   ID      Description  GeneRatio BgRatio    pvalue p.adjust    qvalue geneID Count
##   <chr>   <chr>         <chr>     <chr>    <dbl>   <dbl>    <dbl> <chr>  <int>
## 1 mmu05~ MicroRNAs in~ 41/447    303/8924 4.40e-9 1.29e-6 9.18e-7 Stmn1    41
## 2 mmu05~ MicroRNAs in~ 41/447    303/8924 4.40e-9 1.29e-6 9.18e-7 Rassf1   41
## 3 mmu05~ MicroRNAs in~ 41/447    303/8924 4.40e-9 1.29e-6 9.18e-7 Dnmt3a   41
## 4 mmu05~ MicroRNAs in~ 41/447    303/8924 4.40e-9 1.29e-6 9.18e-7 Hras     41
## 5 mmu05~ MicroRNAs in~ 41/447    303/8924 4.40e-9 1.29e-6 9.18e-7 Itga5   41
## 6 mmu05~ MicroRNAs in~ 41/447    303/8924 4.40e-9 1.29e-6 9.18e-7 Myc     41
## # ... with 2 more variables: GeneRatio_decimal <dbl>, BgRatio_decimal <dbl>
```

```
write.csv(compKEGG_diff_readable_df, "./results/differential_pathways_KEGG_Pathways.csv")
```

Writing all pathways reactome pathways in excel files– suggested cutoff– qvalue<0.2

```
reactome_pathways_diff_df <- as.data.frame(reactome_pathways_diff@result)

reactome_pathways_diff_df$GeneRatio_decimal <- reactome_pathways_diff_df$GeneRatio
reactome_pathways_diff_df$GeneRatio_decimal <- sapply(reactome_pathways_diff_df$GeneRatio_decimal, function(x) (eval(parse(text = as.character(x)))))
reactome_pathways_diff_df$BgRatio_decimal <- reactome_pathways_diff_df$BgRatio
reactome_pathways_diff_df$BgRatio_decimal <- sapply(reactome_pathways_diff_df$BgRatio_decimal, function(x) (eval(parse(text = as.character(x)))))
reactome_pathways_diff_df <- reactome_pathways_diff_df %>% tidyr::separate_rows(geneID, sep = "/", convert = FALSE) %>% arrange(desc(qvalue))
reactome_pathways_diff_df %>% head
```

```
## # A tibble: 6 x 11
##   ID      Description  GeneRatio BgRatio    pvalue p.adjust    qvalue geneID Count
```

```
##   <chr>      <chr>      <chr>      <chr>      <dbl>      <dbl> <dbl> <chr> <int>
## 1 R-MMU-42~ Axon guidan~ 29/465    256/88~ 6.79e-5    0.0130 0.0121 Ptk2      29
## 2 R-MMU-42~ Axon guidan~ 29/465    256/88~ 6.79e-5    0.0130 0.0121 Hras       29
## 3 R-MMU-42~ Axon guidan~ 29/465    256/88~ 6.79e-5    0.0130 0.0121 Itga5      29
## 4 R-MMU-42~ Axon guidan~ 29/465    256/88~ 6.79e-5    0.0130 0.0121 Plxna3     29
## 5 R-MMU-42~ Axon guidan~ 29/465    256/88~ 6.79e-5    0.0130 0.0121 Fgfr1      29
## 6 R-MMU-42~ Axon guidan~ 29/465    256/88~ 6.79e-5    0.0130 0.0121 Plxnc1     29
## # ... with 2 more variables: GeneRatio_decimal <dbl>, BgRatio_decimal <dbl>
```

```
write.csv(reactome_pathways_diff_df, "./results/differential_pathways_Reactome_Pathways_norm.csv")
```

Perform GO analysis on genes common in CHIPT-ATLAS and current ATACseq analysis

```
chip_atac_common <- openxlsx::read.xlsx("CHIP_ATLAS/chipatlas_genrich_common.xlsx")
head(chip_atac_common)
```

```
##   426.common.elements.in."chip-atlas".and."ATAC":
## 1                                     Rxra
## 2                                     Bmpr1a
## 3                                     Brd2
## 4                                     Pard6b
## 5                                     Cct7
## 6                                     Zmym2
```

```
Genes <- chip_atac_common$`426.common.elements.in."chip-atlas".and."ATAC":`

symbol_to_entrez_mouse <- function(x){
  require(biomaRt)
  mart <- useMart(biomaRt = "ensembl", dataset = "mmusculus_gene_ensembl")
  # query biomaRt
  results <- getBM(attributes = c("entrezgene_id", "mgi_symbol"),
                    filters = "mgi_symbol", values = x,
                    mart = mart)

  results
}
```

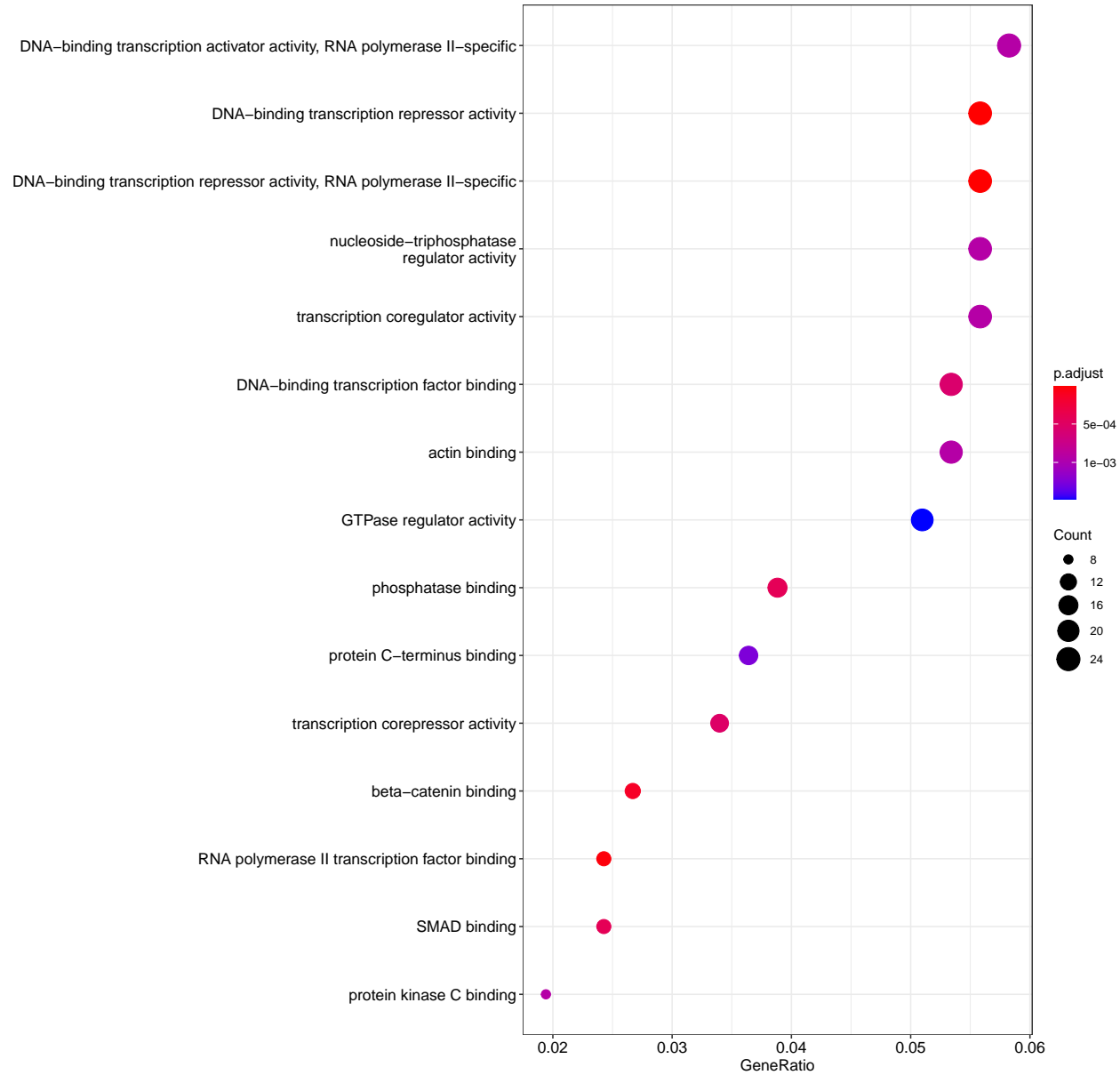
```
common_mouse_entrez <- symbol_to_entrez_mouse(Genes)
head(common_mouse_entrez)
```

```
##   entrezgene_id   mgi_symbol
## 1      72190 2510009E07Rik
## 2      97159 A430005L14Rik
## 3     269774      Aak1
## 4      67618   Aasdhppt
## 5      68904   Abhd13
## 6     226016   Abhd17b
```

```
compGO_common_MF_diff <- enrichGO(gene = common_mouse_entrez$entrezgene_id, pvalueCutoff = 0.05,
                                   pAdjustMethod = "BH", OrgDb = "org.Mm.eg.db", ont = "MF", readable = TRUE)

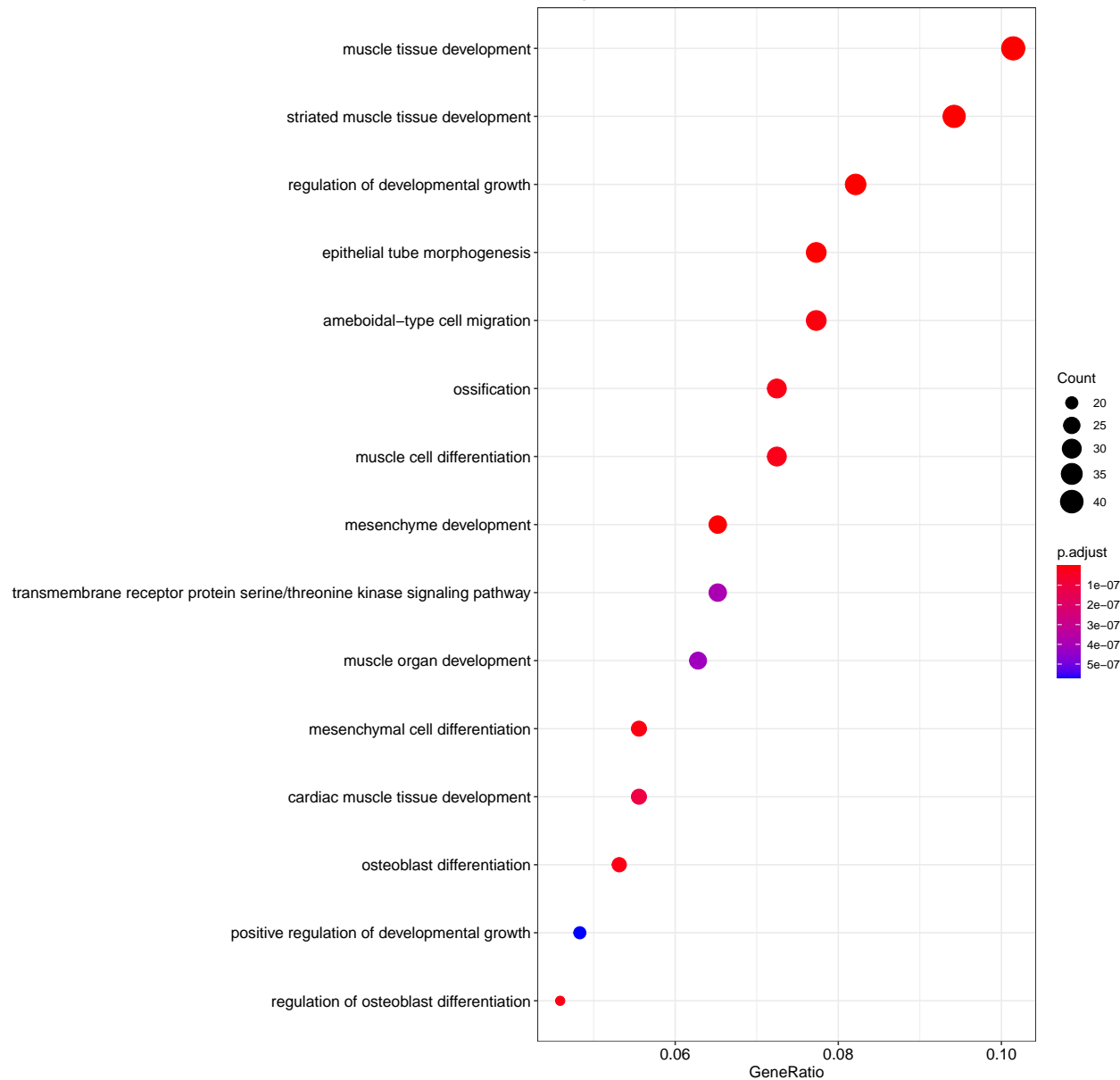
dotplot(compGO_common_MF_diff, showCategory = 15,
        title = "CHIPATLAS common peaks GO Pathway Enrichment Analysis \n Molecular Functions for DARs",
        font.size = 12)
```

CHIPATLAS common peaks GO Pathway Enrichment Analysis
Molecular Functions for DARs



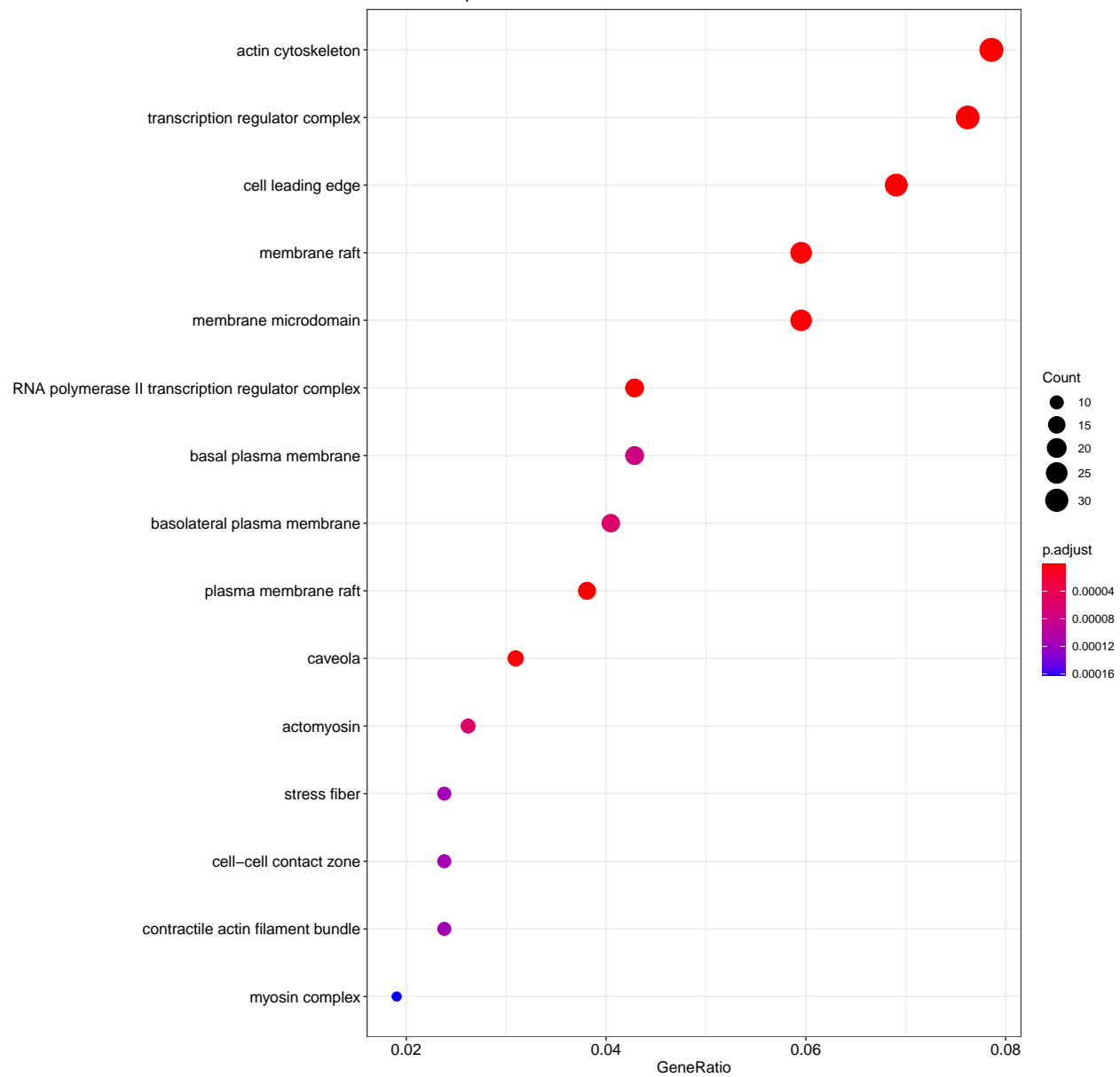
```
compGO_common_BP_diff <- enrichGO(gene = common_mouse_entrez$entrezgene_id, pvalueCutoff = 0.05,  
                                   pAdjustMethod = "BH", OrgDb = "org.Mm.eg.db", ont = "BP", readable = TRUE)  
  
dotplot(compGO_common_BP_diff, showCategory = 15,  
        title = "CHIPATLAS common peaks GO Pathway Enrichment Analysis \n Biological Pathways for DARs",  
        font.size = 12)
```

CHIPATLAS common peaks GO Pathway Enrichment Analysis
Biological Pathways for DARs



```
compGO_common_CC_diff <- enrichGO(gene = common_mouse_entrez$entrezgene_id, pvalueCutoff = 0.05,  
                                   pAdjustMethod = "BH", OrgDb = "org.Mm.eg.db", ont = "CC", readable = TRUE)  
  
dotplot(compGO_common_CC_diff, showCategory = 15,  
        title = "CHIPATLAS common peaks GO Pathway Enrichment Analysis \n Cellular components for DARs",  
        font.size = 12)
```

CHIPATLAS common peaks GO Pathway Enrichment Analysis
Cellular components for DARs




```

compGO_common_CC_diff_df <- as.data.frame(compGO_common_CC_diff)

compGO_common_CC_diff_df$GeneRatio_decimal <- compGO_common_CC_diff_df$GeneRatio
compGO_common_CC_diff_df$GeneRatio_decimal <- sapply(compGO_common_CC_diff_df$GeneRatio_decimal,
  function(x) (eval(parse(text = as.character(x)))))
compGO_common_CC_diff_df$BgRatio_decimal <- compGO_common_CC_diff_df$BgRatio
compGO_common_CC_diff_df$BgRatio_decimal <- sapply(compGO_common_CC_diff_df$BgRatio_decimal,
  function(x) (eval(parse(text = as.character(x)))))
compGO_common_CC_diff_df <- compGO_common_CC_diff_df %>%
  tidyr::separate_rows(geneID, sep = "/", convert = FALSE) %>%
  arrange(desc(GeneRatio_decimal))
compGO_common_CC_diff_df %>% head

```

```

## # A tibble: 6 x 11
##   ID      Description  GeneRatio BgRatio    pvalue p.adjust    qvalue geneID Count
##   <chr>   <chr>         <chr>     <chr>    <dbl>   <dbl>    <dbl> <chr>  <int>
## 1 G0:00~ actin cytosk~ 33/420    500/231~ 1.95e-10 3.42e-8 2.57e-8 Abl1    33
## 2 G0:00~ actin cytosk~ 33/420    500/231~ 1.95e-10 3.42e-8 2.57e-8 Arhge~   33
## 3 G0:00~ actin cytosk~ 33/420    500/231~ 1.95e-10 3.42e-8 2.57e-8 Asap1    33
## 4 G0:00~ actin cytosk~ 33/420    500/231~ 1.95e-10 3.42e-8 2.57e-8 Aut2    33
## 5 G0:00~ actin cytosk~ 33/420    500/231~ 1.95e-10 3.42e-8 2.57e-8 Cdh2     33
## 6 G0:00~ actin cytosk~ 33/420    500/231~ 1.95e-10 3.42e-8 2.57e-8 Coro1c  33
## # ... with 2 more variables: GeneRatio_decimal <dbl>, BgRatio_decimal <dbl>

```

```

write.csv(compGO_common_CC_diff_df,
  "results/chip_common_differential_pathways_G0_Cell_components_norm.csv")

```

```

compGO_common_MF_diff_df <- as.data.frame(compGO_common_MF_diff)

compGO_common_MF_diff_df$GeneRatio_decimal <- compGO_common_MF_diff_df$GeneRatio
compGO_common_MF_diff_df$GeneRatio_decimal <- sapply(compGO_common_MF_diff_df$GeneRatio_decimal,
  function(x) (eval(parse(text = as.character(x)))))
compGO_common_MF_diff_df$BgRatio_decimal <- compGO_common_MF_diff_df$BgRatio
compGO_common_MF_diff_df$BgRatio_decimal <- sapply(compGO_common_MF_diff_df$BgRatio_decimal,
  function(x) (eval(parse(text = as.character(x)))))

```

```
compGO_common_MF_diff_df <- compGO_common_MF_diff_df %>%
  tidyr::separate_rows(geneID, sep = "/", convert = FALSE) %>%
  arrange(desc(GeneRatio_decimal))
compGO_common_MF_diff_df %>% head
```

```
## # A tibble: 6 x 11
##   ID      Description      GeneRatio BgRatio  pvalue p.adjust  qvalue geneID Count
##   <chr>   <chr>          <chr>    <chr>    <dbl>  <dbl>    <dbl> <chr>  <int>
## 1 G0:00~ DNA-binding tr~ 24/412    496/22~ 1.51e-5 0.000930 7.58e-4 Atf3      24
## 2 G0:00~ DNA-binding tr~ 24/412    496/22~ 1.51e-5 0.000930 7.58e-4 Atf4      24
## 3 G0:00~ DNA-binding tr~ 24/412    496/22~ 1.51e-5 0.000930 7.58e-4 Ddit3     24
## 4 G0:00~ DNA-binding tr~ 24/412    496/22~ 1.51e-5 0.000930 7.58e-4 E2f3      24
## 5 G0:00~ DNA-binding tr~ 24/412    496/22~ 1.51e-5 0.000930 7.58e-4 Elk3      24
## 6 G0:00~ DNA-binding tr~ 24/412    496/22~ 1.51e-5 0.000930 7.58e-4 Foxc1     24
## # ... with 2 more variables: GeneRatio_decimal <dbl>, BgRatio_decimal <dbl>
```

```
write.csv(compGO_common_MF_diff_df, "results/chip_common_differential_pathways_G0_Molecular_Functions_norm.csv")
```

```
compGO_common_BP_diff_df <- as.data.frame(compGO_common_BP_diff)

compGO_common_BP_diff_df$GeneRatio_decimal <- compGO_common_BP_diff_df$GeneRatio
compGO_common_BP_diff_df$GeneRatio_decimal <- sapply(compGO_common_BP_diff_df$GeneRatio_decimal,
  function(x) (eval(parse(text = as.character(x)))))
compGO_common_BP_diff_df$BgRatio_decimal <- compGO_common_BP_diff_df$BgRatio
compGO_common_BP_diff_df$BgRatio_decimal <- sapply(compGO_common_BP_diff_df$BgRatio_decimal,
  function(x) (eval(parse(text = as.character(x)))))
compGO_common_BP_diff_df <- compGO_common_BP_diff_df %>%
  tidyr::separate_rows(geneID, sep = "/", convert = FALSE) %>%
  arrange(desc(GeneRatio_decimal))
compGO_common_BP_diff_df %>% head
```

```
## # A tibble: 6 x 11
##   ID      Description      GeneRatio BgRatio  pvalue p.adjust  qvalue geneID Count
##   <chr>   <chr>          <chr>    <chr>    <dbl>  <dbl>    <dbl> <chr>  <int>
## 1 G0:00~ muscle tissu~ 42/414    477/23~ 1.19e-17 4.86e-14 3.35e-14 Aldh1~    42
## 2 G0:00~ muscle tissu~ 42/414    477/23~ 1.19e-17 4.86e-14 3.35e-14 Arid1a    42
```

```
## 3 G0:00~ muscle tissu~ 42/414 477/23~ 1.19e-17 4.86e-14 3.35e-14 Arrb2 42
## 4 G0:00~ muscle tissu~ 42/414 477/23~ 1.19e-17 4.86e-14 3.35e-14 Atf3 42
## 5 G0:00~ muscle tissu~ 42/414 477/23~ 1.19e-17 4.86e-14 3.35e-14 Bcl2 42
## 6 G0:00~ muscle tissu~ 42/414 477/23~ 1.19e-17 4.86e-14 3.35e-14 Bmpr1a 42
## # ... with 2 more variables: GeneRatio_decimal <dbl>, BgRatio_decimal <dbl>
```

```
write.csv(compGO_common_BP_diff_df,
          "results/chip_common_differential_pathways_G0_Biological_Pathways_norm.csv")
```

```
compGO_common_MF_diff_df_plot <- as.data.frame(compGO_common_MF_diff)
compGO_common_MF_diff_df_plot$GeneRatio_decimal <- compGO_common_MF_diff_df_plot$GeneRatio
compGO_common_MF_diff_df_plot$GeneRatio_decimal <- sapply(compGO_common_MF_diff_df_plot$GeneRatio_decimal,
                  function(x) (eval(parse(text = as.character(x)))))
compGO_common_MF_diff_df_plot <- compGO_common_MF_diff_df_plot %>%
  arrange(desc(GeneRatio_decimal))
head(compGO_common_MF_diff_df_plot, 25)
```

```
## ID
## G0:0001228 G0:0001228
## G0:0001217 G0:0001217
## G0:0001227 G0:0001227
## G0:0060589 G0:0060589
## G0:0003712 G0:0003712
## G0:0140297 G0:0140297
## G0:0003779 G0:0003779
## G0:0030695 G0:0030695
## G0:0008047 G0:0008047
## G0:0005543 G0:0005543
## G0:0019902 G0:0019902
## G0:0008022 G0:0008022
## G0:0003714 G0:0003714
## G0:0050839 G0:0050839
## G0:0051015 G0:0051015
## G0:0005096 G0:0005096
## G0:0035091 G0:0035091
## G0:0008013 G0:0008013
## G0:0019838 G0:0019838
## G0:0019903 G0:0019903
```

```

## G0:0001085 G0:0001085
## G0:0046332 G0:0046332
## G0:0005085 G0:0005085
## G0:0003774 G0:0003774
## G0:0005080 G0:0005080
##
## Description
## G0:0001228 DNA-binding transcription activator activity, RNA polymerase II-specific
## G0:0001217 DNA-binding transcription repressor activity
## G0:0001227 DNA-binding transcription repressor activity, RNA polymerase II-specific
## G0:0060589 nucleoside-triphosphatase regulator activity
## G0:0003712 transcription coregulator activity
## G0:0140297 DNA-binding transcription factor binding
## G0:0003779 actin binding
## G0:0030695 GTPase regulator activity
## G0:0008047 enzyme activator activity
## G0:0005543 phospholipid binding
## G0:0019902 phosphatase binding
## G0:0008022 protein C-terminus binding
## G0:0003714 transcription corepressor activity
## G0:0050839 cell adhesion molecule binding
## G0:0051015 actin filament binding
## G0:0005096 GTPase activator activity
## G0:0035091 phosphatidylinositol binding
## G0:0008013 beta-catenin binding
## G0:0019838 growth factor binding
## G0:0019903 protein phosphatase binding
## G0:0001085 RNA polymerase II transcription factor binding
## G0:0046332 SMAD binding
## G0:0005085 guanyl-nucleotide exchange factor activity
## G0:0003774 motor activity
## G0:0005080 protein kinase C binding
##
## GeneRatio BgRatio pvalue p.adjust qvalue
## G0:0001228 24/412 496/22682 1.507787e-05 9.299404e-04 7.576506e-04
## G0:0001217 23/412 330/22682 4.497247e-08 1.452611e-05 1.183486e-05
## G0:0001227 23/412 330/22682 4.497247e-08 1.452611e-05 1.183486e-05
## G0:0060589 23/412 463/22682 1.498188e-05 9.299404e-04 7.576506e-04
## G0:0003712 23/412 466/22682 1.659711e-05 9.299404e-04 7.576506e-04
## G0:0140297 22/412 410/22682 6.945398e-06 5.608409e-04 4.569341e-04
## G0:0003779 22/412 437/22682 1.871397e-05 9.299404e-04 7.576506e-04

```

##	G0:0030695	21/412	422/22682	3.422523e-05	1.473967e-03	1.200885e-03
##	G0:0008047	20/412	470/22682	4.180681e-04	1.264130e-02	1.029925e-02
##	G0:0005543	17/412	450/22682	3.820543e-03	4.570502e-02	3.723726e-02
##	G0:0019902	16/412	225/22682	3.960790e-06	4.264451e-04	3.474377e-04
##	G0:0008022	15/412	234/22682	2.705522e-05	1.248405e-03	1.017113e-03
##	G0:0003714	14/412	180/22682	5.687251e-06	5.248520e-04	4.276129e-04
##	G0:0050839	13/412	287/22682	2.414943e-03	3.563863e-02	2.903587e-02
##	G0:0051015	12/412	212/22682	5.295278e-04	1.368300e-02	1.114795e-02
##	G0:0005096	12/412	238/22682	1.445666e-03	2.524054e-02	2.056424e-02
##	G0:0035091	12/412	264/22682	3.398042e-03	4.221414e-02	3.439314e-02
##	G0:0008013	11/412	92/22682	8.937405e-07	1.443391e-04	1.175974e-04
##	G0:0019838	11/412	147/22682	8.073335e-05	3.259609e-03	2.655702e-03
##	G0:0019903	11/412	178/22682	4.305086e-04	1.264130e-02	1.029925e-02
##	G0:0001085	10/412	63/22682	1.899848e-07	4.091006e-05	3.333067e-05
##	G0:0046332	10/412	85/22682	3.261370e-06	4.213691e-04	3.433022e-04
##	G0:0005085	10/412	197/22682	3.344778e-03	4.221414e-02	3.439314e-02
##	G0:0003774	9/412	126/22682	5.005043e-04	1.347191e-02	1.097597e-02
##	G0:0005080	8/412	63/22682	1.790802e-05	9.299404e-04	7.576506e-04
##						
##	G0:0001228		Atf3/Atf4/Ddit3/E2f3/Elk3/Foxc1/Hivep2/Jun/Klf4/Klf6/Mycn/Nfatc1/Nfic/Nfyb/Osr2/Plagl1/Plagl2/Sox11/Tcf12/Tcf712/Te			
##	G0:0001217		Atf3/Atf7/Crem/E2f7/Elk3/Erf/Hdac5/Jun/Klf16/Mxi1/Nfatc1/Nfil3/Nr1d2/Osr2/Sall1/Tbx3/Tgif1/Zbtb14/Zbtb16/Z			
##	G0:0001227		Atf3/Atf7/Crem/E2f7/Elk3/Erf/Hdac5/Jun/Klf16/Mxi1/Nfatc1/Nfil3/Nr1d2/Osr2/Sall1/Tbx3/Tgif1/Zbtb14/Zbtb16/Z			
##	G0:0060589	Agfg2/Agrn/Arhgap23/Arhgef12/Arhgef2/Asap1/Atp1b1/Dennd5a/Dock8/Farp1/Gdi2/Jun/Myo9a/Myo9b/Rab3ip/Rasa1/Rasgef1b/Rgs3/Ric1/S				
##	G0:0003712		Btg2/Calcoco1/Cbfb/Erf/Ezh1/Hdac4/Hdac5/Hdac7/Irf2bp2/Irf2bpl/Jmjd1c/Kat2a/Kat6b/Kdm2b/Rbfox2/Rbm39/Rybp/Tob1/Tox2			
##	G0:0140297		Arid1a/Atf4/Bcl2/Cry2/Ddit3/Hdac4/Hdac5/Hdac7/Id4/Jun/Klf4/Nfatc1/Nfyb/Rxra/Sall1/Tbx3/Tcf12/Tcf7			
##	G0:0003779		Abl1/Clmn/Coro1c/Ctnna1/Dst/Enah/Ezr/Fhod3/Gmfb/Limch1/Lsp1/Myh9/Myo1b/Myo1c/Myo9a/Myo9b/Pdlim2/Ppp1r9b			
##	G0:0030695		Agfg2/Arhgap23/Arhgef12/Arhgef2/Asap1/Dennd5a/Dock8/Farp1/Gdi2/Jun/Myo9a/Myo9b/Rab3ip/Rasa1/Rasgef1b/Rgs3/Ric1/S			
##	G0:0008047		Agfg2/Arhgap23/Arhgef12/Asap1/Atp1b1/Ccnt2/Cflar/Gdi2/Gnas/Hmgb1/Irgm2/Jun/Myo9a/Myo9b/Psme1/Rasa1/P			
##	G0:0005543		Akt1/Asap1/Atg2a/Eea1/Epn2/Fzd7/Hmgb1/Itpr2/Myo1b/Myo1c/Osbsp15/Pacsin2/Picalm/Ple			
##	G0:0019902		Akt1/Bcl2/Cdh2/Cry2/Dlg1/Eif4ebp1/Ghr/Hsp90b1/Insr/Kat2a/Lmna/Ncam1/Ppp1r			
##	G0:0008022		Abl1/Atf4/Atp1b1/Bcam/Calcoco1/Dgkz/Dlg1/Ezr/Hras/Myo1c/Ppp1r9b/Slc			
##	G0:0003714		Btg2/Erf/Ezh1/Hdac4/Hdac5/Hdac7/Irf2bp2/Irf2bpl/Rbfox2/Ryb			
##	G0:0050839		Bcam/Cdh2/Ctnna1/Dsp/Ezr/Fgfr1/Ilk/Itga5/Kirr			
##	G0:0051015		Abl1/Clmn/Coro1c/Ctnna1/Ezr/Fhod3/Myh9/Myo1b/My			
##	G0:0005096		Agfg2/Arhgap23/Arhgef12/Asap1/Gdi2/Jun/Myo9a/Myo9b/P			
##	G0:0035091		Akt1/Asap1/Atg2a/Eea1/Fzd7/Itpr2/Myo1b/Osbsp15/Pica			
##	G0:0008013		Calcoco1/Cdh2/Ctnna1/Kdm6b/Klf4/Pin1/Sall1/Slc9a			
##	G0:0019838		Acvr2b/Agrn/Fgfr1/Ghr/Il11ra1/Insr/Nradd/C			
##	G0:0019903		Akt1/Bcl2/Cdh2/Eif4ebp1/Ghr/Hsp90b1/Insr/			

```

## G0:0001085 Atf4/Hdac4/Hdac5/Id4/Klf4/Nfatc
## G0:0046332 Bmpr1a/Jun/Rnf111/Smad7/Smurf2/Tcf12/T
## G0:0005085 Arhgef12/Arhgef2/Dennd5a/Dock8/Farp1/Rab3ip/R
## G0:0003774 Dync1li2/Myh9/My16/Myo1b/Myo1c
## G0:0005080 Abl1/Akt1/Dsp/Hdac5/

```

```

## Count GeneRatio_decimal
## G0:0001228 24 0.05825243
## G0:0001217 23 0.05582524
## G0:0001227 23 0.05582524
## G0:0060589 23 0.05582524
## G0:0003712 23 0.05582524
## G0:0140297 22 0.05339806
## G0:0003779 22 0.05339806
## G0:0030695 21 0.05097087
## G0:0008047 20 0.04854369
## G0:0005543 17 0.04126214
## G0:0019902 16 0.03883495
## G0:0008022 15 0.03640777
## G0:0003714 14 0.03398058
## G0:0050839 13 0.03155340
## G0:0051015 12 0.02912621
## G0:0005096 12 0.02912621
## G0:0035091 12 0.02912621
## G0:0008013 11 0.02669903
## G0:0019838 11 0.02669903
## G0:0019903 11 0.02669903
## G0:0001085 10 0.02427184
## G0:0046332 10 0.02427184
## G0:0005085 10 0.02427184
## G0:0003774 9 0.02184466
## G0:0005080 8 0.01941748

```

```
compG0_common_MF_diff_df_plot[1:25,]
```

```

## ID
## G0:0001228 G0:0001228
## G0:0001217 G0:0001217
## G0:0001227 G0:0001227

```

G0:0060589 G0:0060589
 ## G0:0003712 G0:0003712
 ## G0:0140297 G0:0140297
 ## G0:0003779 G0:0003779
 ## G0:0030695 G0:0030695
 ## G0:0008047 G0:0008047
 ## G0:0005543 G0:0005543
 ## G0:0019902 G0:0019902
 ## G0:0008022 G0:0008022
 ## G0:0003714 G0:0003714
 ## G0:0050839 G0:0050839
 ## G0:0051015 G0:0051015
 ## G0:0005096 G0:0005096
 ## G0:0035091 G0:0035091
 ## G0:0008013 G0:0008013
 ## G0:0019838 G0:0019838
 ## G0:0019903 G0:0019903
 ## G0:0001085 G0:0001085
 ## G0:0046332 G0:0046332
 ## G0:0005085 G0:0005085
 ## G0:0003774 G0:0003774
 ## G0:0005080 G0:0005080

	Description
## G0:0001228	DNA-binding transcription activator activity, RNA polymerase II-specific
## G0:0001217	DNA-binding transcription repressor activity
## G0:0001227	DNA-binding transcription repressor activity, RNA polymerase II-specific
## G0:0060589	nucleoside-triphosphatase regulator activity
## G0:0003712	transcription coregulator activity
## G0:0140297	DNA-binding transcription factor binding
## G0:0003779	actin binding
## G0:0030695	GTPase regulator activity
## G0:0008047	enzyme activator activity
## G0:0005543	phospholipid binding
## G0:0019902	phosphatase binding
## G0:0008022	protein C-terminus binding
## G0:0003714	transcription corepressor activity
## G0:0050839	cell adhesion molecule binding
## G0:0051015	actin filament binding
## G0:0005096	GTPase activator activity

G0:0035091 phosphatidylinositol binding
 ## G0:0008013 beta-catenin binding
 ## G0:0019838 growth factor binding
 ## G0:0019903 protein phosphatase binding
 ## G0:0001085 RNA polymerase II transcription factor binding
 ## G0:0046332 SMAD binding
 ## G0:0005085 guanyl-nucleotide exchange factor activity
 ## G0:0003774 motor activity
 ## G0:0005080 protein kinase C binding

##	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
## G0:0001228	24/412	496/22682	1.507787e-05	9.299404e-04	7.576506e-04
## G0:0001217	23/412	330/22682	4.497247e-08	1.452611e-05	1.183486e-05
## G0:0001227	23/412	330/22682	4.497247e-08	1.452611e-05	1.183486e-05
## G0:0060589	23/412	463/22682	1.498188e-05	9.299404e-04	7.576506e-04
## G0:0003712	23/412	466/22682	1.659711e-05	9.299404e-04	7.576506e-04
## G0:0140297	22/412	410/22682	6.945398e-06	5.608409e-04	4.569341e-04
## G0:0003779	22/412	437/22682	1.871397e-05	9.299404e-04	7.576506e-04
## G0:0030695	21/412	422/22682	3.422523e-05	1.473967e-03	1.200885e-03
## G0:0008047	20/412	470/22682	4.180681e-04	1.264130e-02	1.029925e-02
## G0:0005543	17/412	450/22682	3.820543e-03	4.570502e-02	3.723726e-02
## G0:0019902	16/412	225/22682	3.960790e-06	4.264451e-04	3.474377e-04
## G0:0008022	15/412	234/22682	2.705522e-05	1.248405e-03	1.017113e-03
## G0:0003714	14/412	180/22682	5.687251e-06	5.248520e-04	4.276129e-04
## G0:0050839	13/412	287/22682	2.414943e-03	3.563863e-02	2.903587e-02
## G0:0051015	12/412	212/22682	5.295278e-04	1.368300e-02	1.114795e-02
## G0:0005096	12/412	238/22682	1.445666e-03	2.524054e-02	2.056424e-02
## G0:0035091	12/412	264/22682	3.398042e-03	4.221414e-02	3.439314e-02
## G0:0008013	11/412	92/22682	8.937405e-07	1.443391e-04	1.175974e-04
## G0:0019838	11/412	147/22682	8.073335e-05	3.259609e-03	2.655702e-03
## G0:0019903	11/412	178/22682	4.305086e-04	1.264130e-02	1.029925e-02
## G0:0001085	10/412	63/22682	1.899848e-07	4.091006e-05	3.333067e-05
## G0:0046332	10/412	85/22682	3.261370e-06	4.213691e-04	3.433022e-04
## G0:0005085	10/412	197/22682	3.344778e-03	4.221414e-02	3.439314e-02
## G0:0003774	9/412	126/22682	5.005043e-04	1.347191e-02	1.097597e-02
## G0:0005080	8/412	63/22682	1.790802e-05	9.299404e-04	7.576506e-04

##

G0:0001228 Atf3/Atf4/Ddit3/E2f3/Elk3/Foxc1/Hivep2/Jun/Klf4/Klf6/Mycn/Nfatc1/Nfic/Nfyb/Osr2/Plagl1/Plagl2/Sox11/Tcf12/Tcf712/Te
 ## G0:0001217 Atf3/Atf7/Crem/E2f7/Elk3/Erf/Hdac5/Jun/Klf16/Mxi1/Nfatc1/Nfil3/Nr1d2/Osr2/Sall1/Tbx3/Tgif1/Zbtb14/Zbtb16/Z
 ## G0:0001227 Atf3/Atf7/Crem/E2f7/Elk3/Erf/Hdac5/Jun/Klf16/Mxi1/Nfatc1/Nfil3/Nr1d2/Osr2/Sall1/Tbx3/Tgif1/Zbtb14/Zbtb16/Z


```

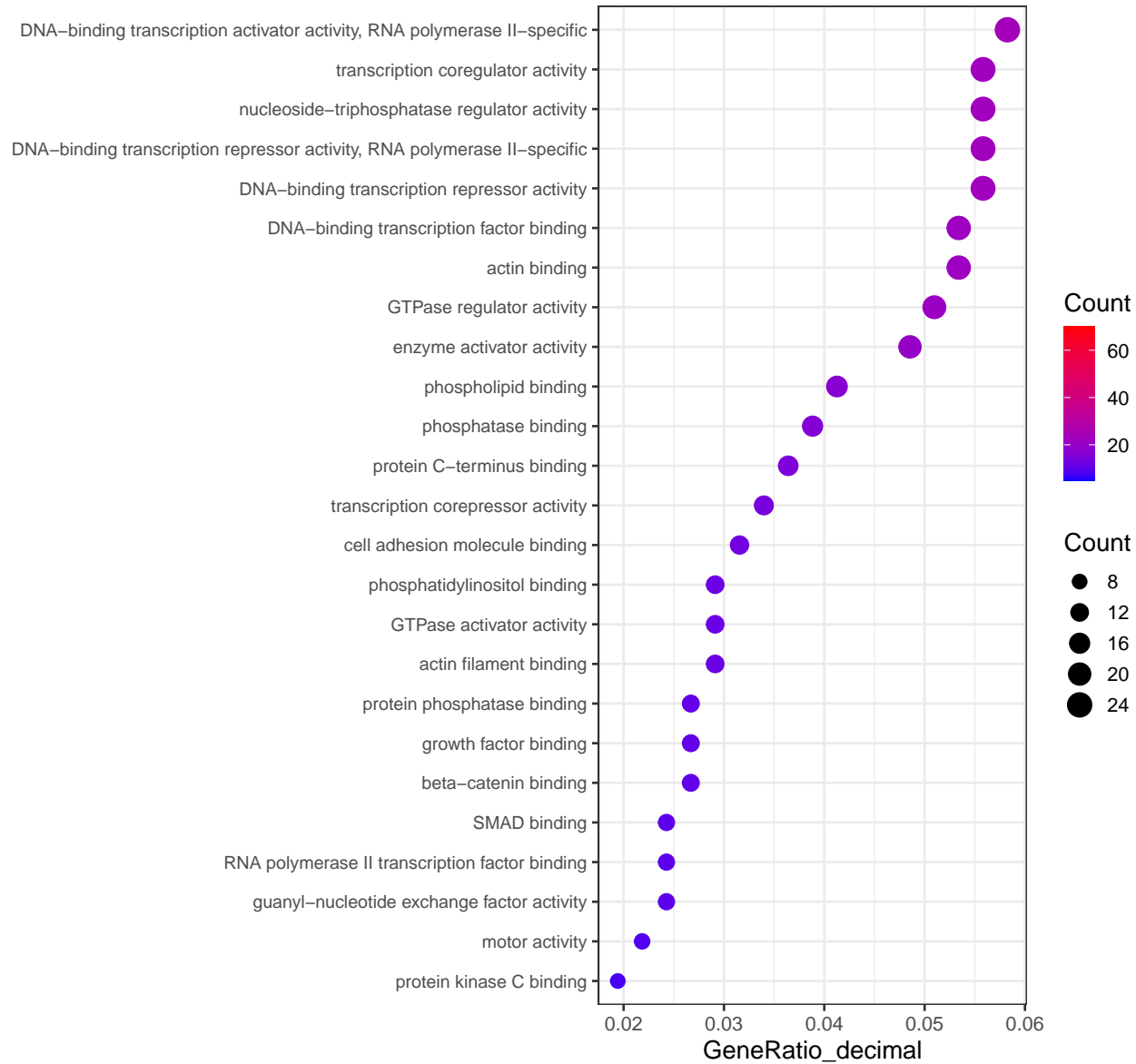
## G0:0060589 Agfg2/Agrn/Arhgap23/Arhgef12/Arhgef2/Asap1/Atp1b1/Dennd5a/Dock8/Farp1/Gdi2/Jun/Myo9a/Myo9b/Rab3ip/Rasa1/Rasgef1b/Rgs3/Ric1/S
## G0:0003712 Btg2/Calcoco1/Cbfb/Erf/Ezh1/Hdac4/Hdac5/Hdac7/Irf2bp2/Irf2bp1/Jmjd1c/Kat2a/Kat6b/Kdm2b/Rbfox2/Rbm39/Rybp/Tob1/Tox2
## G0:0140297 Arid1a/Atf4/Bcl2/Cry2/Ddit3/Hdac4/Hdac5/Hdac7/Id4/Jun/Klf4/Nfatc1/Nfyb/Rxra/Sall1/Tbx3/Tcf12/Tcf7
## G0:0003779 Abl1/Clmn/Coro1c/Ctnna1/Dst/Enah/Ezr/Fhod3/Gmfb/Limch1/Lsp1/Myh9/Myo1b/Myo1c/Myo9a/Myo9b/Pdlim2/Ppp1r9b
## G0:0030695 Agfg2/Arhgap23/Arhgef12/Arhgef2/Asap1/Dennd5a/Dock8/Farp1/Gdi2/Jun/Myo9a/Myo9b/Rab3ip/Rasa1/Rasgef1b/Rgs3/Ric1/S
## G0:0008047 Agfg2/Arhgap23/Arhgef12/Asap1/Atp1b1/Ccnt2/Cflar/Gdi2/Gnas/Hmgb1/Irgm2/Jun/Myo9a/Myo9b/Psme1/Rasa1/F
## G0:0005543 Akt1/Asap1/Atg2a/Eea1/Epn2/Fzd7/Hmgb1/Itpr2/Myo1b/Myo1c/Osbpl5/Pacsin2/Picalm/Ple
## G0:0019902 Akt1/Bcl2/Cdh2/Cry2/Dlg1/Eif4ebp1/Ghr/Hsp90b1/Insr/Kat2a/Lmna/Ncam1/Ppp1r
## G0:0008022 Abl1/Atf4/Atp1b1/Bcam/Calcoco1/Dgkz/Dlg1/Ezr/Hras/Myo1c/Ppp1r9b/Slc
## G0:0003714 Btg2/Erf/Ezh1/Hdac4/Hdac5/Hdac7/Irf2bp2/Irf2bp1/Rbfox2/Ryb
## G0:0050839 Bcam/Cdh2/Ctnna1/Dsp/Ezr/Fgfr1/Ilk/Itga5/Kirr
## G0:0051015 Abl1/Clmn/Coro1c/Ctnna1/Ezr/Fhod3/Myh9/Myo1b/My
## G0:0005096 Agfg2/Arhgap23/Arhgef12/Asap1/Gdi2/Jun/Myo9a/Myo9b/F
## G0:0035091 Akt1/Asap1/Atg2a/Eea1/Fzd7/Itpr2/Myo1b/Osbpl5/Pica
## G0:0008013 Calcoco1/Cdh2/Ctnna1/Kdm6b/Klf4/Pin1/Sall1/Slc9a
## G0:0019838 Acvr2b/Agrn/Fgfr1/Ghr/Il11ra1/Insr/Nradd/C
## G0:0019903 Akt1/Bcl2/Cdh2/Eif4ebp1/Ghr/Hsp90b1/Insr/
## G0:0001085 Atf4/Hdac4/Hdac5/Id4/Klf4/Nfatc
## G0:0046332 Bmpr1a/Jun/Rnf111/Smad7/Smurf2/Tcf12/T
## G0:0005085 Arhgef12/Arhgef2/Dennd5a/Dock8/Farp1/Rab3ip/R
## G0:0003774 Dync1li2/Myh9/My16/Myo1b/Myo1c
## G0:0005080 Abl1/Akt1/Dsp/Hdac5/
##
## Count GeneRatio_decimal
## G0:0001228 24 0.05825243
## G0:0001217 23 0.05582524
## G0:0001227 23 0.05582524
## G0:0060589 23 0.05582524
## G0:0003712 23 0.05582524
## G0:0140297 22 0.05339806
## G0:0003779 22 0.05339806
## G0:0030695 21 0.05097087
## G0:0008047 20 0.04854369
## G0:0005543 17 0.04126214
## G0:0019902 16 0.03883495
## G0:0008022 15 0.03640777
## G0:0003714 14 0.03398058
## G0:0050839 13 0.03155340
## G0:0051015 12 0.02912621
## G0:0005096 12 0.02912621

```

```
## G0:0035091    12      0.02912621
## G0:0008013    11      0.02669903
## G0:0019838    11      0.02669903
## G0:0019903    11      0.02669903
## G0:0001085    10      0.02427184
## G0:0046332    10      0.02427184
## G0:0005085    10      0.02427184
## G0:0003774     9      0.02184466
## G0:0005080     8      0.01941748
```

```
ggplot(compG0_common_MF_diff_df_plot[1:25,], # you can replace the numbers
       #to the row number of pathway of your interest
       aes(x = GeneRatio_decimal, y = reorder(Description, GeneRatio_decimal))) +
  geom_point(aes(size = Count, color = Count)) +
  theme_bw(base_size = 14) +
  scale_size_area()+
  scale_colour_gradient( low="blue", high = "red", limits= c(5,70)) +
  theme(axis.text.y = element_text(size=9.5))+
  # scale_x_continuous(breaks=seq(0.02, 0.08, 0.01))+
  ylab(NULL) +
  ggtitle("G0 Molecular Functions \n pathway enrichment")
```

GO Molecular Functions pathway enrichment



```
dev.copy(
  svg,
  file = paste0("./results/chip_common_GO_MF_pathways.svg"),
  width = 10,
  height = 8
)
```

```
## svg
## 3
```

```
dev.off ()
```

```
## pdf
## 2
```

```
compGO_common_CC_diff_df_plot <- as.data.frame(compGO_common_CC_diff)
compGO_common_CC_diff_df_plot$GeneRatio_decimal <- compGO_common_CC_diff_df_plot$GeneRatio
compGO_common_CC_diff_df_plot$GeneRatio_decimal <- sapply(compGO_common_CC_diff_df_plot$GeneRatio_decimal,
  function(x) (eval(parse(text = as.character(x)))))
compGO_common_CC_diff_df_plot <- compGO_common_CC_diff_df_plot %>% arrange(desc(GeneRatio_decimal))
head(compGO_common_CC_diff_df_plot, 25)
```

```
## ID
## GO:0015629 GO:0015629
## GO:0005667 GO:0005667
## GO:0031252 GO:0031252
## GO:0045121 GO:0045121
## GO:0098857 GO:0098857
## GO:0043235 GO:0043235
## GO:0090575 GO:0090575
## GO:0009925 GO:0009925
## GO:0045178 GO:0045178
## GO:0019898 GO:0019898
## GO:0045177 GO:0045177
## GO:0016323 GO:0016323
## GO:0016607 GO:0016607
## GO:0044853 GO:0044853
```

##	G0:0150034	G0:0150034			
##	G0:0043209	G0:0043209			
##	G0:0031253	G0:0031253			
##	G0:0016324	G0:0016324			
##	G0:0005901	G0:0005901			
##	G0:0030027	G0:0030027			
##	G0:0001726	G0:0001726			
##	G0:0030055	G0:0030055			
##	G0:0061695	G0:0061695			
##	G0:0042641	G0:0042641			
##	G0:0005912	G0:0005912			
##				Description	
##	G0:0015629			actin cytoskeleton	
##	G0:0005667			transcription regulator complex	
##	G0:0031252			cell leading edge	
##	G0:0045121			membrane raft	
##	G0:0098857			membrane microdomain	
##	G0:0043235			receptor complex	
##	G0:0090575			RNA polymerase II transcription regulator complex	
##	G0:0009925			basal plasma membrane	
##	G0:0045178			basal part of cell	
##	G0:0019898			extrinsic component of membrane	
##	G0:0045177			apical part of cell	
##	G0:0016323			basolateral plasma membrane	
##	G0:0016607			nuclear speck	
##	G0:0044853			plasma membrane raft	
##	G0:0150034			distal axon	
##	G0:0043209			myelin sheath	
##	G0:0031253			cell projection membrane	
##	G0:0016324			apical plasma membrane	
##	G0:0005901			caveola	
##	G0:0030027			lamellipodium	
##	G0:0001726			ruffle	
##	G0:0030055			cell-substrate junction	
##	G0:0061695			transferase complex, transferring phosphorus-containing groups	
##	G0:0042641			actomyosin	
##	G0:0005912			adherens junction	
##		GeneRatio	BgRatio	pvalue	p.adjust
##	G0:0015629	33/420	500/23177	1.948814e-10	3.420748e-08
				qvalue	
				2.567378e-08	

##	G0:0005667	32/420	469/23177	1.605469e-10	3.420748e-08	2.567378e-08
##	G0:0031252	29/420	396/23177	2.300952e-10	3.420748e-08	2.567378e-08
##	G0:0045121	25/420	381/23177	3.662936e-08	2.149117e-06	1.612979e-06
##	G0:0098857	25/420	382/23177	3.854919e-08	2.149117e-06	1.612979e-06
##	G0:0043235	22/420	414/23177	7.842415e-06	2.057481e-04	1.544203e-04
##	G0:0090575	18/420	169/23177	1.956171e-09	1.877316e-07	1.408984e-07
##	G0:0009925	18/420	265/23177	1.860476e-06	7.543384e-05	5.661544e-05
##	G0:0045178	18/420	288/23177	5.974623e-06	1.665426e-04	1.249954e-04
##	G0:0019898	18/420	323/23177	2.807155e-05	5.443440e-04	4.085470e-04
##	G0:0045177	18/420	460/23177	1.987631e-03	1.846841e-02	1.386111e-02
##	G0:0016323	17/420	232/23177	1.277825e-06	5.699100e-05	4.277351e-05
##	G0:0016607	17/420	331/23177	1.264914e-04	2.089451e-03	1.568198e-03
##	G0:0044853	16/420	131/23177	2.104614e-09	1.877316e-07	1.408984e-07
##	G0:0150034	16/420	385/23177	1.909875e-03	1.812349e-02	1.360224e-02
##	G0:0043209	15/420	213/23177	8.681300e-06	2.065142e-04	1.549953e-04
##	G0:0031253	15/420	301/23177	4.280832e-04	6.158875e-03	4.622427e-03
##	G0:0016324	14/420	354/23177	5.508881e-03	4.387431e-02	3.292903e-02
##	G0:0005901	13/420	97/23177	2.218810e-08	1.649315e-06	1.237862e-06
##	G0:0030027	13/420	172/23177	1.626542e-05	3.627188e-04	2.722317e-04
##	G0:0001726	12/420	149/23177	1.821335e-05	3.868168e-04	2.903180e-04
##	G0:0030055	12/420	180/23177	1.153809e-04	1.979226e-03	1.485471e-03
##	G0:0061695	12/420	241/23177	1.577657e-03	1.636360e-02	1.228139e-02
##	G0:0042641	11/420	95/23177	1.211121e-06	5.699100e-05	4.277351e-05
##	G0:0005912	11/420	143/23177	6.167692e-05	1.146163e-03	8.602307e-04
##						
##	G0:0015629	Abl1/Arhgef2/Asap1/Auts2/Cdh2/Coro1c/Ctnna1/Dgkh/Dst/Ehbp1/Enah/Ezr/Fhod3/Flot2/Hdac4/Ilk/Ints6/Keap1/Limch1/Myh9/Myl6/Myl9/				
##	G0:0005667	Arid4a/Atf4/Bdp1/Cbfb/Cdk4/Crem/Ddit3/E2f3/E2f7/Hdac4/Hdac5/Hivep2/Hnrnpab/Jun/Kat2a/Klf4/Mxi1/Nfatc1/Nfyb/Rxra				
##	G0:0031252	Aak1/Abl1/Arf6/Arhgef2/Atf4/Cacna1d/Cdh2/Coro1c/Ctnna1/Dgkz/Dock8/Enah/Ezr/Flot2/Gnas/Ilk/Insr/I				
##	G0:0045121	Atp1b1/Bmpr1a/Cavin1/Cblb/Cdh2/Cflar/Coro1c/Ctnna1/Dlg1/Ezr/Flot2/Gnas/Insr/Kcnd3/K				
##	G0:0098857	Atp1b1/Bmpr1a/Cavin1/Cblb/Cdh2/Cflar/Coro1c/Ctnna1/Dlg1/Ezr/Flot2/Gnas/Insr/Kcnd3/K				
##	G0:0043235	Acvr2b/Bmpr1a/Dlg1/Fgfr1/Ghr/I111ra1/I113ra1/Insr/Itga5/Itpr				
##	G0:0090575	Atf4/Cbfb/Crem/Ddit3/E2f3/E2f7				
##	G0:0009925	Arrb2/Atp1b1/Cdh2/Dlg1/Dsp/Dsty/E				
##	G0:0045178	Arrb2/Atp1b1/Cdh2/Dlg1/Dsp/Dsty/E				
##	G0:0019898	Atg2a/Cdh2/Ctnna1/Eea1/Ezr/Farp1				
##	G0:0045177	Atp1b1/Cacna1d/Cdh2/Clic4/Dsty/Ezr/Lrp1				
##	G0:0016323	Arrb2/Atp1b1/Cdh2/Dlg1/Dsp/D				
##	G0:0016607	Brd2/Cry2/Dgkz/E2f7/Erbin/Gac				
##	G0:0044853	Atp1b1/Bmpr1a/Cavin1/Cdh2/Cor				

G0:0150034
 ## G0:0043209
 ## G0:0031253
 ## G0:0016324
 ## G0:0005901
 ## G0:0030027
 ## G0:0001726
 ## G0:0030055
 ## G0:0061695
 ## G0:0042641
 ## G0:0005912

##	Count	GeneRatio_decimal
## G0:0015629	33	0.07857143
## G0:0005667	32	0.07619048
## G0:0031252	29	0.06904762
## G0:0045121	25	0.05952381
## G0:0098857	25	0.05952381
## G0:0043235	22	0.05238095
## G0:0090575	18	0.04285714
## G0:0009925	18	0.04285714
## G0:0045178	18	0.04285714
## G0:0019898	18	0.04285714
## G0:0045177	18	0.04285714
## G0:0016323	17	0.04047619
## G0:0016607	17	0.04047619
## G0:0044853	16	0.03809524
## G0:0150034	16	0.03809524
## G0:0043209	15	0.03571429
## G0:0031253	15	0.03571429
## G0:0016324	14	0.03333333
## G0:0005901	13	0.03095238
## G0:0030027	13	0.03095238
## G0:0001726	12	0.02857143
## G0:0030055	12	0.02857143
## G0:0061695	12	0.02857143
## G0:0042641	11	0.02619048
## G0:0005912	11	0.02619048

Aak1/Abl1/Agrn/Auts2
 Arf6/Atp1b1
 Arf6/Arhgef2/Asap1/
 Atp1b1/Cacna1d/Cd
 Atp1b1/Bmpr1a/
 Cdh2

```
compG0_common_CC_diff_df_plot[1:25,]
```

##	ID	Description
##	G0:0015629 G0:0015629	
##	G0:0005667 G0:0005667	
##	G0:0031252 G0:0031252	
##	G0:0045121 G0:0045121	
##	G0:0098857 G0:0098857	
##	G0:0043235 G0:0043235	
##	G0:0090575 G0:0090575	
##	G0:0009925 G0:0009925	
##	G0:0045178 G0:0045178	
##	G0:0019898 G0:0019898	
##	G0:0045177 G0:0045177	
##	G0:0016323 G0:0016323	
##	G0:0016607 G0:0016607	
##	G0:0044853 G0:0044853	
##	G0:0150034 G0:0150034	
##	G0:0043209 G0:0043209	
##	G0:0031253 G0:0031253	
##	G0:0016324 G0:0016324	
##	G0:0005901 G0:0005901	
##	G0:0030027 G0:0030027	
##	G0:0001726 G0:0001726	
##	G0:0030055 G0:0030055	
##	G0:0061695 G0:0061695	
##	G0:0042641 G0:0042641	
##	G0:0005912 G0:0005912	
##		
##	G0:0015629	actin cytoskeleton
##	G0:0005667	transcription regulator complex
##	G0:0031252	cell leading edge
##	G0:0045121	membrane raft
##	G0:0098857	membrane microdomain
##	G0:0043235	receptor complex
##	G0:0090575	RNA polymerase II transcription regulator complex
##	G0:0009925	basal plasma membrane
##	G0:0045178	basal part of cell

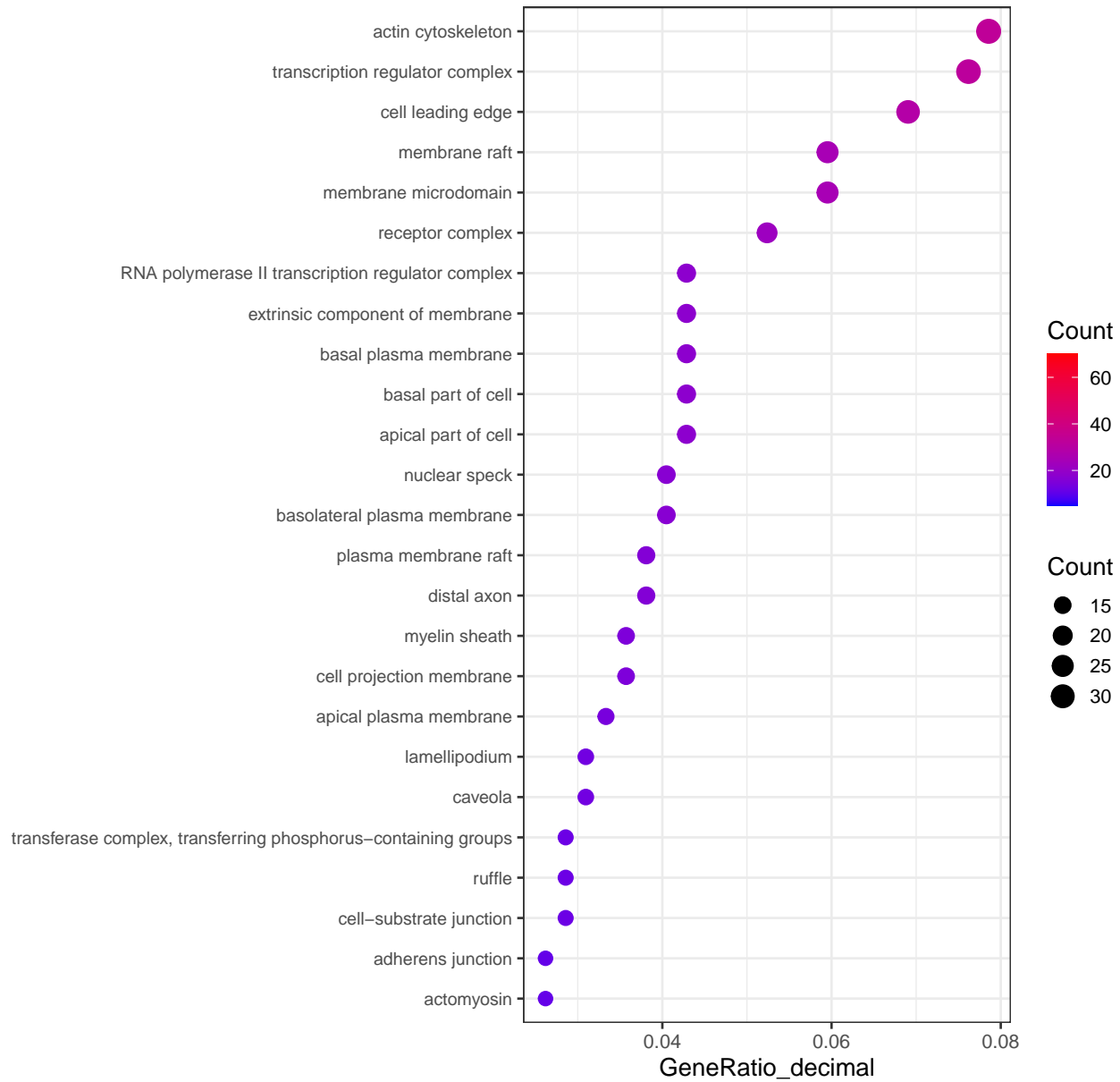
## G0:0019898					extrinsic component of membrane
## G0:0045177					apical part of cell
## G0:0016323					basolateral plasma membrane
## G0:0016607					nuclear speck
## G0:0044853					plasma membrane raft
## G0:0150034					distal axon
## G0:0043209					myelin sheath
## G0:0031253					cell projection membrane
## G0:0016324					apical plasma membrane
## G0:0005901					caveola
## G0:0030027					lamellipodium
## G0:0001726					ruffle
## G0:0030055					cell-substrate junction
## G0:0061695					transferase complex, transferring phosphorus-containing groups
## G0:0042641					actomyosin
## G0:0005912					adherens junction
##	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
## G0:0015629	33/420	500/23177	1.948814e-10	3.420748e-08	2.567378e-08
## G0:0005667	32/420	469/23177	1.605469e-10	3.420748e-08	2.567378e-08
## G0:0031252	29/420	396/23177	2.300952e-10	3.420748e-08	2.567378e-08
## G0:0045121	25/420	381/23177	3.662936e-08	2.149117e-06	1.612979e-06
## G0:0098857	25/420	382/23177	3.854919e-08	2.149117e-06	1.612979e-06
## G0:0043235	22/420	414/23177	7.842415e-06	2.057481e-04	1.544203e-04
## G0:0090575	18/420	169/23177	1.956171e-09	1.877316e-07	1.408984e-07
## G0:0009925	18/420	265/23177	1.860476e-06	7.543384e-05	5.661544e-05
## G0:0045178	18/420	288/23177	5.974623e-06	1.665426e-04	1.249954e-04
## G0:0019898	18/420	323/23177	2.807155e-05	5.443440e-04	4.085470e-04
## G0:0045177	18/420	460/23177	1.987631e-03	1.846841e-02	1.386111e-02
## G0:0016323	17/420	232/23177	1.277825e-06	5.699100e-05	4.277351e-05
## G0:0016607	17/420	331/23177	1.264914e-04	2.089451e-03	1.568198e-03
## G0:0044853	16/420	131/23177	2.104614e-09	1.877316e-07	1.408984e-07
## G0:0150034	16/420	385/23177	1.909875e-03	1.812349e-02	1.360224e-02
## G0:0043209	15/420	213/23177	8.681300e-06	2.065142e-04	1.549953e-04
## G0:0031253	15/420	301/23177	4.280832e-04	6.158875e-03	4.622427e-03
## G0:0016324	14/420	354/23177	5.508881e-03	4.387431e-02	3.292903e-02
## G0:0005901	13/420	97/23177	2.218810e-08	1.649315e-06	1.237862e-06
## G0:0030027	13/420	172/23177	1.626542e-05	3.627188e-04	2.722317e-04
## G0:0001726	12/420	149/23177	1.821335e-05	3.868168e-04	2.903180e-04
## G0:0030055	12/420	180/23177	1.153809e-04	1.979226e-03	1.485471e-03

## G0:0061695	12/420	241/23177	1.577657e-03	1.636360e-02	1.228139e-02
## G0:0042641	11/420	95/23177	1.211121e-06	5.699100e-05	4.277351e-05
## G0:0005912	11/420	143/23177	6.167692e-05	1.146163e-03	8.602307e-04
##					
## G0:0015629	Abl1/Arhgef2/Asap1/Auts2/Cdh2/Coro1c/Ctnna1/Dgkh/Dst/Ehbp1/Enah/Ezr/Fhod3/Flot2/Hdac4/Ilk/Ints6/Keap1/Limch1/Myh9/Myl6/Myl9/				
## G0:0005667	Arid4a/Atf4/Bdp1/Cbfb/Cdk4/Crem/Ddit3/E2f3/E2f7/Hdac4/Hdac5/Hivep2/Hnrnpab/Jun/Kat2a/Klf4/Mxi1/Nfatc1/Nfyb/Rxra				
## G0:0031252	Aak1/Abl1/Arf6/Arhgef2/Atf4/Cacna1d/Cdh2/Coro1c/Ctnna1/Dgkz/Dock8/Enah/Ezr/Flot2/Gnas/Ilk/Insr/I				
## G0:0045121	Atp1b1/Bmpr1a/Cavin1/Cblb/Cdh2/Cflar/Coro1c/Ctnna1/Dlg1/Ezr/Flot2/Gnas/Insr/Kcnd3/K				
## G0:0098857	Atp1b1/Bmpr1a/Cavin1/Cblb/Cdh2/Cflar/Coro1c/Ctnna1/Dlg1/Ezr/Flot2/Gnas/Insr/Kcnd3/K				
## G0:0043235	Acvr2b/Bmpr1a/Dlg1/Fgfr1/Ghr/Il11ra1/Il13ra1/Insr/Itga5/Itpr				
## G0:0090575	Atf4/Cbfb/Crem/Ddit3/E2f3/E2f7/Hdac4/Hdac5/Hivep2/Hnrnpab/Jun/Kat2a/Klf4/Mxi1/Nfatc1/Nfyb/Rxra				
## G0:0009925	Arrb2/Atp1b1/Cdh2/Dlg1/Dsp/Dsty/E				
## G0:0045178	Arrb2/Atp1b1/Cdh2/Dlg1/Dsp/Dsty/E				
## G0:0019898	Atg2a/Cdh2/Ctnna1/Eea1/Ezr/Farp1				
## G0:0045177	Atp1b1/Cacna1d/Cdh2/Clic4/Dsty/Ezr/Lrp1				
## G0:0016323	Arrb2/Atp1b1/Cdh2/Dlg1/Dsp/Dsty/E				
## G0:0016607	Brd2/Cry2/Dgkz/E2f7/Erbin/Gac				
## G0:0044853	Atp1b1/Bmpr1a/Cavin1/Cdh2/Coro1c/Ctnna1/Dlg1/Ezr/Flot2/Gnas/Insr/Kcnd3/K				
## G0:0150034	Aak1/Abl1/Agrn/Auts2				
## G0:0043209	Arf6/Atp1b1				
## G0:0031253	Arf6/Arhgef2/Asap1/Auts2/Cdh2/Coro1c/Ctnna1/Dlg1/Ezr/Flot2/Gnas/Insr/Kcnd3/K				
## G0:0016324	Atp1b1/Cacna1d/Cdh2/Clic4/Dsty/Ezr/Lrp1				
## G0:0005901	Atp1b1/Bmpr1a/Cavin1/Cdh2/Coro1c/Ctnna1/Dlg1/Ezr/Flot2/Gnas/Insr/Kcnd3/K				
## G0:0030027	Atp1b1/Bmpr1a/Cavin1/Cdh2/Coro1c/Ctnna1/Dlg1/Ezr/Flot2/Gnas/Insr/Kcnd3/K				
## G0:0001726	Cdh2				
## G0:0030055					
## G0:0061695					
## G0:0042641					
## G0:0005912					
##	Count	GeneRatio_decimal			
## G0:0015629	33	0.07857143			
## G0:0005667	32	0.07619048			
## G0:0031252	29	0.06904762			
## G0:0045121	25	0.05952381			
## G0:0098857	25	0.05952381			
## G0:0043235	22	0.05238095			
## G0:0090575	18	0.04285714			
## G0:0009925	18	0.04285714			
## G0:0045178	18	0.04285714			

```
## G0:0019898      18      0.04285714
## G0:0045177      18      0.04285714
## G0:0016323      17      0.04047619
## G0:0016607      17      0.04047619
## G0:0044853      16      0.03809524
## G0:0150034      16      0.03809524
## G0:0043209      15      0.03571429
## G0:0031253      15      0.03571429
## G0:0016324      14      0.03333333
## G0:0005901      13      0.03095238
## G0:0030027      13      0.03095238
## G0:0001726      12      0.02857143
## G0:0030055      12      0.02857143
## G0:0061695      12      0.02857143
## G0:0042641      11      0.02619048
## G0:0005912      11      0.02619048
```

```
ggplot(compG0_common_CC_diff_df_plot[1:25,], # you can replace the numbers to
#the row number of pathway of your interest
  aes(x = GeneRatio_decimal, y = reorder(Description, GeneRatio_decimal))) +
  geom_point(aes(size = Count, color = Count)) +
  scale_size_area()+
  theme_bw(base_size = 14) +
  scale_colour_gradient( low="blue", high = "red", limits= c(5,70)) +
  theme(axis.text.y = element_text(size=9.5))+
# scale_x_continuous(breaks=seq(0.02, 0.08, 0.01))+
  ylab(NULL) +
  ggtitle("G0 Cellular Component pathways enrichment")
```

GO Cellular Component pathways enrichment



```
dev.copy(
  svg,
  file = paste0("./results/chip_common_GO_CC_pathways.svg"),
  width = 10,
  height = 8
)
```

```
## svg
## 3
```

```
dev.off ()
```

```
## pdf
## 2
```

```
compGO_common_BP_diff_df_plot <- as.data.frame(compGO_common_BP_diff)
compGO_common_BP_diff_df_plot$GeneRatio_decimal <- compGO_common_BP_diff_df_plot$GeneRatio
compGO_common_BP_diff_df_plot$GeneRatio_decimal <-
  sapply(compGO_common_BP_diff_df_plot$GeneRatio_decimal,
    function(x) (eval(parse(text = as.character(x)))))
compGO_common_BP_diff_df_plot <- compGO_common_BP_diff_df_plot %>%
  arrange(desc(GeneRatio_decimal))

head(compGO_common_BP_diff_df_plot, 25)
```

```
##          ID
## GO:0060537 GO:0060537
## GO:0014706 GO:0014706
## GO:0048638 GO:0048638
## GO:0060562 GO:0060562
## GO:0001667 GO:0001667
## GO:0001503 GO:0001503
## GO:0042692 GO:0042692
## GO:0034329 GO:0034329
## GO:0060485 GO:0060485
## GO:0007178 GO:0007178
## GO:0048608 GO:0048608
```

G0:0061458 G0:0061458
 ## G0:0007517 G0:0007517
 ## G0:0001655 G0:0001655
 ## G0:0009896 G0:0009896
 ## G0:0001654 G0:0001654
 ## G0:0150063 G0:0150063
 ## G0:0048880 G0:0048880
 ## G0:0043087 G0:0043087
 ## G0:0072001 G0:0072001
 ## G0:0016055 G0:0016055
 ## G0:0198738 G0:0198738
 ## G0:0045786 G0:0045786
 ## G0:0048762 G0:0048762
 ## G0:0048738 G0:0048738

	Description
## G0:0060537	muscle tissue development
## G0:0014706	striated muscle tissue development
## G0:0048638	regulation of developmental growth
## G0:0060562	epithelial tube morphogenesis
## G0:0001667	ameboidal-type cell migration
## G0:0001503	ossification
## G0:0042692	muscle cell differentiation
## G0:0034329	cell junction assembly
## G0:0060485	mesenchyme development
## G0:0007178	transmembrane receptor protein serine/threonine kinase signaling pathway
## G0:0048608	reproductive structure development
## G0:0061458	reproductive system development
## G0:0007517	muscle organ development
## G0:0001655	urogenital system development
## G0:0009896	positive regulation of catabolic process
## G0:0001654	eye development
## G0:0150063	visual system development
## G0:0048880	sensory system development
## G0:0043087	regulation of GTPase activity
## G0:0072001	renal system development
## G0:0016055	Wnt signaling pathway
## G0:0198738	cell-cell signaling by wnt
## G0:0045786	negative regulation of cell cycle
## G0:0048762	mesenchymal cell differentiation

GO:0048738 cardiac muscle tissue development

##	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
----	-----------	---------	--------	----------	--------

```
## G0:0060537      42/414 477/23355 1.193341e-17 4.855704e-14 3.347635e-14
```

```
## G0:0014706      39/414 453/23355 3.927943e-16 7.991400e-13 5.509457e-13
```

```
## G0:0048638      34/414 403/23355 5.983140e-14 8.115133e-11 5.594761e-11
```

```
## G0:0060562      32/414 393/23355 8.845118e-13 8.997696e-10 6.203221e-10
```

```
## G0:0001667      32/414 439/23355 1.670249e-11 1.132707e-08 7.809145e-09
```

```
## G0:0001503      30/414 406/23355 5.256583e-11 2.138904e-08 1.474610e-08
```

```
## G0:0042692      30/414 410/23355 6.684829e-11 2.472779e-08 1.704791e-08
```

```
## G0:0034329      28/414 432/23355 4.438968e-09 9.031080e-07 6.226237e-07
```

```
## G0:0060485      27/414 288/23355 2.142513e-12 1.743577e-09 1.202062e-09
```

```
## G0:0007178      27/414 381/23355 1.242531e-09 3.889122e-07 2.681251e-07
```

```
## G0:0048608      27/414 478/23355 1.429853e-07 1.264799e-05 8.719816e-06
```

```
## G0:0061458      27/414 482/23355 1.688766e-07 1.431581e-05 9.869651e-06
```

```
## G0:0007517      26/414 357/23355 1.424674e-09 4.140714e-07 2.854705e-07
```

```
## G0:0001655      26/414 379/23355 4.989534e-09 9.463255e-07 6.524188e-07
```

```
## G0:0009896      26/414 457/23355 2.128049e-07 1.647415e-05 1.135766e-05
```

```
## G0:0001654      25/414 427/23355 2.138649e-07 1.647415e-05 1.135766e-05
```

```
## G0:0150063      25/414 430/23355 2.439101e-07 1.723792e-05 1.188422e-05
```

```
## G0:0048880      25/414 434/23355 2.900554e-07 2.000399e-05 1.379121e-05
```

```
## G0:0043087      24/414 318/23355 3.067536e-09 6.934335e-07 4.780692e-07
## G2:0070001      24/414 326/23355 3.015555e-09 4.533355e-07 4.251333e-07
```

##	GU:0072001	24/414	336/23355	9.017575e-09	1.528855e-06	1.054028e-06
##	GU:0012255	24/414	122/23355	1.222222e-08	1.122222e-06	1.115252e-06

##	GU	0016055	24/414	438/23355	1.220262e-06	6.448374e-05	4.445659e-05
##	00	0100700	04/414	440/23355	1.001700e-06	6.000774e-05	4.000410e-05

##	GU	0198738	24/414	440/23355	1.321730e-06	6.807745e-05	4.693418e-05
##	33	3345733	34/414	450/23355	6.111165e-06	3.411370e-05	3.503515e-05

```
## GU:0045786      24/414 452/23355 2.111465e-06 9.441972e-05 6.509515e-05
## GU:0048768      23/414 232/23355 3.050001e-11 1.372226e-03 1.225036e-03
```

```
## GU:0048762      23/414 232/23355 3.059091e-11 1.778206e-08 1.225936e-08
## GU:0048738      23/414 261/23355 3.265167e-10 1.107161e-07 7.623043e-08
```

```
## GU:0048738      23/414 261/23355 3.265167e-10 1.107164e-07 7.633043e-08
##
```

```
##
##  G0:0060527_Aldh1a2/Arid1a/Arrb2/Atf3/Bcl2/Bmpr1a/Btg2/Cent2/Cflar/Chd
```

```
## GU:0000337 AId11a2/AId1a/AId1b2/At1S/BC1z/Bmp11a/Btgz/Cc1t2/C11a1/Cnd
## GU:0014706 Aldb1a3/Arid1a/Arrb3/Atf3/Bcl3/Bmpr1a/B
```

```
## G0:0014700      A1G11A2/A1G11A/A1G12/A1G13/BG12/Bmp11A/B
## G0:0018638
```

```
## GD:0048038
## GD:0060562
```

```
## G0:00000502
## G0:0001667
```

```
## G0:0001007
## G0:0001503
```

```
## GG:0001000
## GQ:0042692
```

```
## G0:0012092
## G0:0034329
```

G0:0060485

```
## G0:0007178
```

G0:0048608

135

/Fhod3/Foxc1/Fzd7/Hdac4/Hdac5/Hdac7/Kat2a/Kdm6b/Lm
 /Chd7/Dsp/Fgfr1/Fhod3/Foxc1/Fzd7/Hdac4/Hdac5/Hdac7/
 /Bcl2/Bmpr1a/Cdk4/Celf1/Chd7/Dlg1/Dusp6/Ezr/Fgfr1/Fc
 /Arid1a/Bcl2/Brd2/Clic4/Dlg1/Enah/Hdac5/Ilk/Kat2a/Kd
 /Arsb/Cdh2/Coro1c/Fgfr1/Fstl1/Gadd45a/Gna12/Hdac5/
 Acvr2b/Akt1/Atf4/Bcl2/Bmpr1a/Cbfb/Fgfr1/Fndc3b/
 Akt1/Arid1a/Arrb2/Bcl2/Bmpr1a/Ccnt2/Cdh2/Cfl
 Abl1/Adgre5/Agrn/Arf6/Bcl2/Cdh2/Coro1c/
 Aldh1a2/Bcl2/Bmpr1a/Cdh2/Coro1c/Fgfr1/Foxc1/Hnrnp
 Abl1/Acvr2b/Arrb2/Bmpr1a/Fstl1/Got1
 Akt1/Arid1a/Arid4a/Arrb2/Bcl2/Ccnf/C

## G0:0061458			Akt1/Arid1a/Arid4a/Arrb2/Bcl2/Ccnf/C
## G0:0007517			Adarb1/Atf3/Bcl2/Bmpr1a/Btg2
## G0:0001655			Acvr2b/Aldh1a2/Bcl2/Cflar/Cu
## G0:0009896			Akt1/Btg2/Cblb/Celf1/Cpt
## G0:0001654			Acvr2b/Aldh1a2/Arid1a/
## G0:0150063			Acvr2b/Aldh1a2/Arid1a/
## G0:0048880			Acvr2b/Aldh1a2/Arid1a/
## G0:0043087			Agrrn/Arrb2/Asap1/Cblb/Cor
## G0:0072001			Acvr2b/Aldh1a2/Bcl2
## G0:0016055			Abl1/Calcc
## G0:0198738			Abl1/Calcc
## G0:0045786			Abl1/Bcl2/E
## G0:0048762			Aldh1a2/Bcl2/Bmpr1a/Cdh2/
## G0:0048738			Aldh1a2/Ari
##	Count	GeneRatio_decimal	
## G0:0060537	42	0.10144928	
## G0:0014706	39	0.09420290	
## G0:0048638	34	0.08212560	
## G0:0060562	32	0.07729469	
## G0:0001667	32	0.07729469	
## G0:0001503	30	0.07246377	
## G0:0042692	30	0.07246377	
## G0:0034329	28	0.06763285	
## G0:0060485	27	0.06521739	
## G0:0007178	27	0.06521739	
## G0:0048608	27	0.06521739	
## G0:0061458	27	0.06521739	
## G0:0007517	26	0.06280193	
## G0:0001655	26	0.06280193	
## G0:0009896	26	0.06280193	
## G0:0001654	25	0.06038647	
## G0:0150063	25	0.06038647	
## G0:0048880	25	0.06038647	
## G0:0043087	24	0.05797101	
## G0:0072001	24	0.05797101	
## G0:0016055	24	0.05797101	
## G0:0198738	24	0.05797101	
## G0:0045786	24	0.05797101	
## G0:0048762	23	0.05555556	


```
## G0:0048738      23      0.05555556
```

```
compG0_common_BP_diff_df_plot[1:25,]
```

```
##          ID
```

```
## G0:0060537 G0:0060537
```

```
## G0:0014706 G0:0014706
```

```
## G0:0048638 G0:0048638
```

```
## G0:0060562 G0:0060562
```

```
## G0:0001667 G0:0001667
```

```
## G0:0001503 G0:0001503
```

```
## G0:0042692 G0:0042692
```

```
## G0:0034329 G0:0034329
```

```
## G0:0060485 G0:0060485
```

```
## G0:0007178 G0:0007178
```

```
## G0:0048608 G0:0048608
```

```
## G0:0061458 G0:0061458
```

```
## G0:0007517 G0:0007517
```

```
## G0:0001655 G0:0001655
```

```
## G0:0009896 G0:0009896
```

```
## G0:0001654 G0:0001654
```

```
## G0:0150063 G0:0150063
```

```
## G0:0048880 G0:0048880
```

```
## G0:0043087 G0:0043087
```

```
## G0:0072001 G0:0072001
```

```
## G0:0016055 G0:0016055
```

```
## G0:0198738 G0:0198738
```

```
## G0:0045786 G0:0045786
```

```
## G0:0048762 G0:0048762
```

```
## G0:0048738 G0:0048738
```

```
##
```

```
Description
```

```
## G0:0060537
```

```
muscle tissue development
```

```
## G0:0014706
```

```
striated muscle tissue development
```

```
## G0:0048638
```

```
regulation of developmental growth
```

```
## G0:0060562
```

```
epithelial tube morphogenesis
```

```
## G0:0001667
```

```
ameboidal-type cell migration
```

```
## G0:0001503
```

```
ossification
```

```
## G0:0042692
```

```
muscle cell differentiation
```

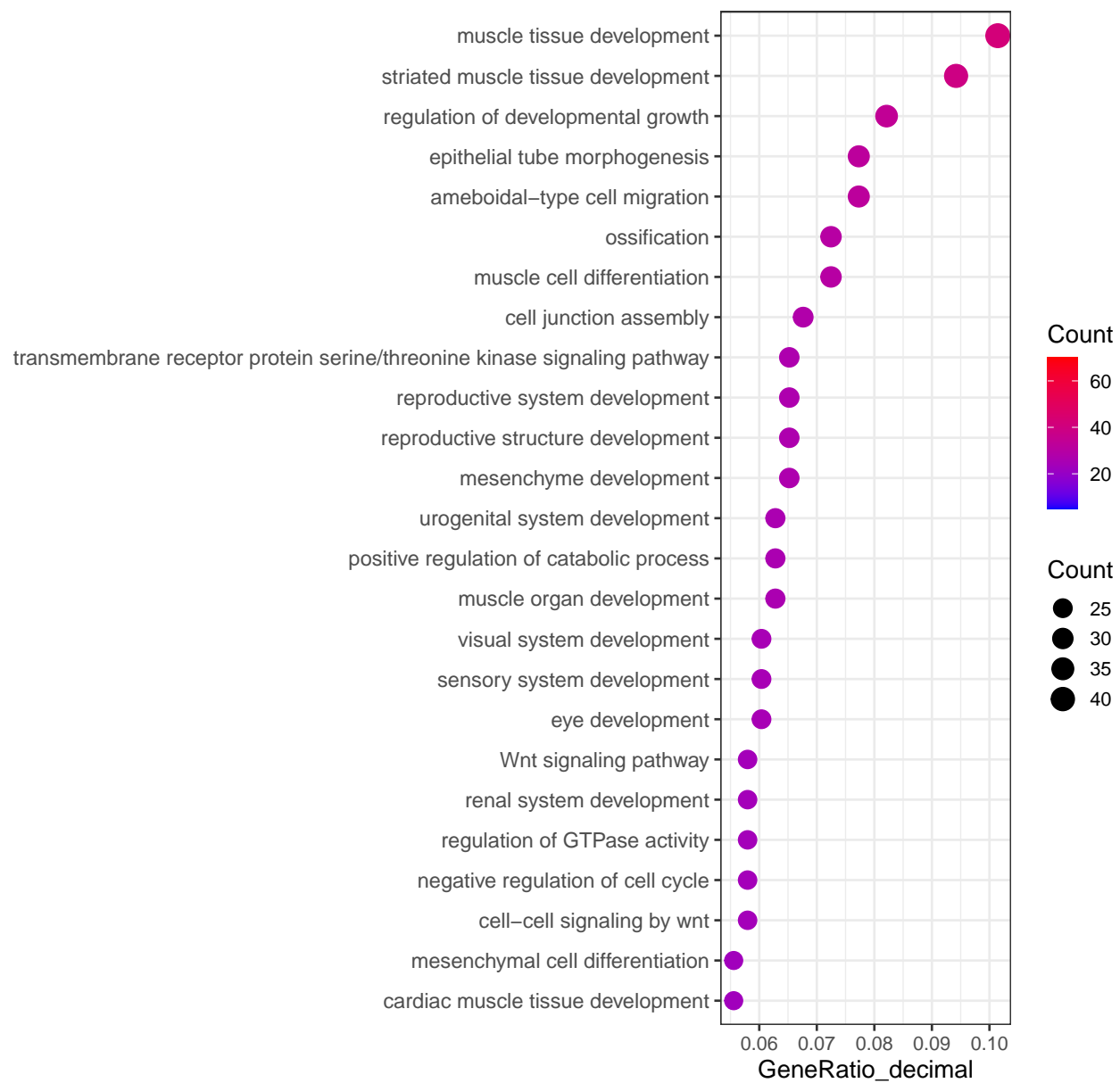
##	G0:0034329					cell junction assembly
##	G0:0060485					mesenchyme development
##	G0:0007178					transmembrane receptor protein serine/threonine kinase signaling pathway
##	G0:0048608					reproductive structure development
##	G0:0061458					reproductive system development
##	G0:0007517					muscle organ development
##	G0:0001655					urogenital system development
##	G0:0009896					positive regulation of catabolic process
##	G0:0001654					eye development
##	G0:0150063					visual system development
##	G0:0048880					sensory system development
##	G0:0043087					regulation of GTPase activity
##	G0:0072001					renal system development
##	G0:0016055					Wnt signaling pathway
##	G0:0198738					cell-cell signaling by wnt
##	G0:0045786					negative regulation of cell cycle
##	G0:0048762					mesenchymal cell differentiation
##	G0:0048738					cardiac muscle tissue development
##		GeneRatio	BgRatio	pvalue	p.adjust	qvalue
##	G0:0060537	42/414	477/23355	1.193341e-17	4.855704e-14	3.347635e-14
##	G0:0014706	39/414	453/23355	3.927943e-16	7.991400e-13	5.509457e-13
##	G0:0048638	34/414	403/23355	5.983140e-14	8.115133e-11	5.594761e-11
##	G0:0060562	32/414	393/23355	8.845118e-13	8.997696e-10	6.203221e-10
##	G0:0001667	32/414	439/23355	1.670249e-11	1.132707e-08	7.809145e-09
##	G0:0001503	30/414	406/23355	5.256583e-11	2.138904e-08	1.474610e-08
##	G0:0042692	30/414	410/23355	6.684829e-11	2.472779e-08	1.704791e-08
##	G0:0034329	28/414	432/23355	4.438968e-09	9.031080e-07	6.226237e-07
##	G0:0060485	27/414	288/23355	2.142513e-12	1.743577e-09	1.202062e-09
##	G0:0007178	27/414	381/23355	1.242531e-09	3.889122e-07	2.681251e-07
##	G0:0048608	27/414	478/23355	1.429853e-07	1.264799e-05	8.719816e-06
##	G0:0061458	27/414	482/23355	1.688766e-07	1.431581e-05	9.869651e-06
##	G0:0007517	26/414	357/23355	1.424674e-09	4.140714e-07	2.854705e-07
##	G0:0001655	26/414	379/23355	4.989534e-09	9.463255e-07	6.524188e-07
##	G0:0009896	26/414	457/23355	2.128049e-07	1.647415e-05	1.135766e-05
##	G0:0001654	25/414	427/23355	2.138649e-07	1.647415e-05	1.135766e-05
##	G0:0150063	25/414	430/23355	2.439101e-07	1.723792e-05	1.188422e-05
##	G0:0048880	25/414	434/23355	2.900554e-07	2.000399e-05	1.379121e-05
##	G0:0043087	24/414	318/23355	3.067536e-09	6.934335e-07	4.780692e-07
##	G0:0072001	24/414	336/23355	9.017575e-09	1.528855e-06	1.054028e-06

##	G0:0016055	24/414	438/23355	1.220262e-06	6.448374e-05	4.445659e-05
##	G0:0198738	24/414	440/23355	1.321730e-06	6.807745e-05	4.693418e-05
##	G0:0045786	24/414	452/23355	2.111465e-06	9.441972e-05	6.509515e-05
##	G0:0048762	23/414	232/23355	3.059091e-11	1.778206e-08	1.225936e-08
##	G0:0048738	23/414	261/23355	3.265167e-10	1.107164e-07	7.633043e-08
##						
##	G0:0060537	Aldh1a2/Arid1a/Arrb2/Atf3/Bcl2/Bmpr1a/Btg2/Ccnt2/Cflar/Chd7/Dlg1/Dsp/Fgfr1/Fhod3/Foxc1/Fzd7/Hdac4/Hdac5/Hdac7/Kat2a/Kdm6b/Lm				
##	G0:0014706	Aldh1a2/Arid1a/Arrb2/Atf3/Bcl2/Bmpr1a/Btg2/Ccnt2/Cflar/Chd7/Dsp/Fgfr1/Fhod3/Foxc1/Fzd7/Hdac4/Hdac5/Hdac7/				
##	G0:0048638	Abl1/Agrn/Akt1/Bcl2/Bmpr1a/Cdk4/Celf1/Chd7/Dlg1/Dusp6/Ezr/Fgfr1/Fc				
##	G0:0060562	Abl1/Aldh1a2/Arid1a/Bcl2/Brd2/Clic4/Dlg1/Enah/Hdac5/Ilk/Kat2a/Kd				
##	G0:0001667	Abl1/Akt1/Arf6/Arsb/Cdh2/Coro1c/Fgfr1/Fstl1/Gadd45a/Gna12/Hdac5/				
##	G0:0001503	Acvr2b/Akt1/Atf4/Bcl2/Bmpr1a/Cbfb/Fgfr1/Fndc3b/				
##	G0:0042692	Akt1/Arid1a/Arrb2/Bcl2/Bmpr1a/Ccnt2/Cdh2/Cfl				
##	G0:0034329	Abl1/Adgre5/Agrn/Arf6/Bcl2/Cdh2/Coro1c/				
##	G0:0060485	Aldh1a2/Bcl2/Bmpr1a/Cdh2/Coro1c/Fgfr1/Foxc1/Hnrnp				
##	G0:0007178	Abl1/Acvr2b/Arrb2/Bmpr1a/Fstl1/Got1				
##	G0:0048608	Akt1/Arid1a/Arid4a/Arrb2/Bcl2/Ccnf/C				
##	G0:0061458	Akt1/Arid1a/Arid4a/Arrb2/Bcl2/Ccnf/C				
##	G0:0007517	Adarb1/Atf3/Bcl2/Bmpr1a/Btg2				
##	G0:0001655	Acvr2b/Aldh1a2/Bcl2/Cflar/Cu				
##	G0:0009896	Akt1/Btg2/Cblb/Celf1/Cpt				
##	G0:0001654	Acvr2b/Aldh1a2/Arid1a/				
##	G0:0150063	Acvr2b/Aldh1a2/Arid1a/				
##	G0:0048880	Acvr2b/Aldh1a2/Arid1a/				
##	G0:0043087	Agrn/Arrb2/Asap1/Cblb/Cor				
##	G0:0072001	Acvr2b/Aldh1a2/Bcl2				
##	G0:0016055	Abl1/Calcc				
##	G0:0198738	Abl1/Calcc				
##	G0:0045786	Abl1/Bcl2/E				
##	G0:0048762	Aldh1a2/Bcl2/Bmpr1a/Cdh2/				
##	G0:0048738	Aldh1a2/Ari				
##		Count	GeneRatio_decimal			
##	G0:0060537	42	0.10144928			
##	G0:0014706	39	0.09420290			
##	G0:0048638	34	0.08212560			
##	G0:0060562	32	0.07729469			
##	G0:0001667	32	0.07729469			
##	G0:0001503	30	0.07246377			
##	G0:0042692	30	0.07246377			

```
## G0:0034329      28      0.06763285
## G0:0060485      27      0.06521739
## G0:0007178      27      0.06521739
## G0:0048608      27      0.06521739
## G0:0061458      27      0.06521739
## G0:0007517      26      0.06280193
## G0:0001655      26      0.06280193
## G0:0009896      26      0.06280193
## G0:0001654      25      0.06038647
## G0:0150063      25      0.06038647
## G0:0048880      25      0.06038647
## G0:0043087      24      0.05797101
## G0:0072001      24      0.05797101
## G0:0016055      24      0.05797101
## G0:0198738      24      0.05797101
## G0:0045786      24      0.05797101
## G0:0048762      23      0.05555556
## G0:0048738      23      0.05555556
```

```
ggplot(compG0_common_BP_diff_df_plot[1:25,], # you can replace the numbers
       # to the row number of pathway of your interest
       aes(x = GeneRatio_decimal, y = reorder(Description, GeneRatio_decimal))) +
  geom_point(aes(size = Count, color = Count)) +
  scale_size_area()+
  theme_bw(base_size = 14) +
  scale_colour_gradient( low="blue", high = "red", limits= c(5,70)) +
  theme(axis.text.y = element_text(size=12))+
  # scale_x_continuous(breaks=seq(0.02, 0.08, 0.01))+
  ylab(NULL) +
  ggtitle("G0 Biological Pathways enrichment")
```

GO Biological Pathways enrichr



```
dev.copy(  
  svg,  
  file = paste0("./results/chip_common_GO_BP_pathways.svg"),  
  width = 12,  
  height = 8  
)
```

```
## svg  
## 3
```

```
dev.off ()
```

```
## pdf  
## 2
```

Save RData and end of program

```
save.image("Marques_et_al_2021_ATACseq.RData")
```

```
sessionInfo()
```

```
## R version 4.1.0 (2021-05-18)  
## Platform: x86_64-pc-linux-gnu (64-bit)  
## Running under: Ubuntu 20.04.2 LTS  
##  
## Matrix products: default  
## BLAS: /usr/lib/x86_64-linux-gnu/blas/libblas.so.3.9.0  
## LAPACK: /usr/lib/x86_64-linux-gnu/lapack/liblapack.so.3.9.0  
##  
## locale:  
## [1] LC_CTYPE=en_US.UTF-8      LC_NUMERIC=C  
## [3] LC_TIME=de_CH.UTF-8      LC_COLLATE=en_US.UTF-8  
## [5] LC_MONETARY=de_CH.UTF-8  LC_MESSAGES=en_US.UTF-8  
## [7] LC_PAPER=de_CH.UTF-8     LC_NAME=C
```

```

## [9] LC_ADDRESS=C LC_TELEPHONE=C
## [11] LC_MEASUREMENT=de_CH.UTF-8 LC_IDENTIFICATION=C
##
## attached base packages:
## [1] parallel stats4 stats graphics grDevices utils datasets
## [8] methods base
##
## other attached packages:
## [1] biomaRt_2.48.0 org.Mm.eg.db_3.13.0
## [3] clusterProfiler_4.0.0 ReactomePA_1.36.0
## [5] viridis_0.6.1 viridisLite_0.4.0
## [7] RColorBrewer_1.1-2 ggpubr_0.4.0
## [9] ggplot2_3.3.5 rstatix_0.7.0
## [11] ChIPseeker_1.28.3 GenomicFeatures_1.44.0
## [13] AnnotationDbi_1.54.0 DiffBind_3.2.1
## [15] SummarizedExperiment_1.22.0 Biobase_2.52.0
## [17] MatrixGenerics_1.4.0 matrixStats_0.59.0
## [19] GenomicRanges_1.44.0 GenomeInfoDb_1.28.0
## [21] IRanges_2.26.0 S4Vectors_0.30.0
## [23] BiocGenerics_0.38.0 dplyr_1.0.6
##
## loaded via a namespace (and not attached):
## [1] rappdirs_0.3.3
## [2] rtracklayer_1.52.0
## [3] AnnotationForge_1.34.0
## [4] coda_0.19-4
## [5] tidyr_1.1.3
## [6] csaw_1.26.0
## [7] bit64_4.0.5
## [8] knitr_1.33
## [9] irlba_2.3.3
## [10] DelayedArray_0.18.0
## [11] data.table_1.14.0
## [12] hwriter_1.3.2
## [13] KEGGREST_1.32.0
## [14] RCurl_1.98-1.3
## [15] generics_0.1.0
## [16] cowplot_1.1.1
## [17] RSQLite_2.2.7

```

```
## [18] shadowtext_0.0.8
## [19] bit_4.0.4
## [20] enrichplot_1.12.2
## [21] base64url_1.4
## [22] xml2_1.3.2
## [23] lubridate_1.7.10
## [24] assertthat_0.2.1
## [25] batchtools_0.9.15
## [26] amap_0.8-18
## [27] apeglm_1.14.0
## [28] xfun_0.24
## [29] hms_1.1.0
## [30] evaluate_0.14
## [31] fansi_0.5.0
## [32] restfulr_0.0.13
## [33] progress_1.2.2
## [34] caTools_1.18.2
## [35] dbplyr_2.1.1
## [36] readxl_1.3.1
## [37] Rgraphviz_2.36.0
## [38] igraph_1.2.6
## [39] DBI_1.1.1
## [40] geneplotter_1.70.0
## [41] purrr_0.3.4
## [42] ellipsis_0.3.2
## [43] backports_1.2.1
## [44] V8_3.4.2
## [45] annotate_1.70.0
## [46] vctrs_0.3.8
## [47] abind_1.4-5
## [48] RMariaDB_1.1.1
## [49] cachem_1.0.5
## [50] withr_2.4.2
## [51] ggforce_0.3.3
## [52] DOT_0.1
## [53] BSgenome_1.60.0
## [54] bdsmatrix_1.3-4
## [55] checkmate_2.0.0
## [56] GenomicAlignments_1.28.0
```



```
## [57] treeio_1.16.1
## [58] prettyunits_1.1.1
## [59] DOSE_3.18.1
## [60] ape_5.5
## [61] lazyeval_0.2.2
## [62] crayon_1.4.1
## [63] genefilter_1.74.0
## [64] labeling_0.4.2
## [65] edgeR_3.34.0
## [66] pkgconfig_2.0.3
## [67] tweenr_1.0.2
## [68] nlme_3.1-152
## [69] rlang_0.4.11
## [70] lifecycle_1.0.0
## [71] downloader_0.4
## [72] filelock_1.0.2
## [73] BiocFileCache_2.0.0
## [74] GOstats_2.58.0
## [75] invgamma_1.1
## [76] cellranger_1.1.0
## [77] rsvg_2.1.2
## [78] polyclip_1.10-0
## [79] graph_1.70.0
## [80] Matrix_1.3-4
## [81] aplot_0.0.6
## [82] ashr_2.2-47
## [83] carData_3.0-4
## [84] boot_1.3-28
## [85] pheatmap_1.0.12
## [86] png_0.1-7
## [87] rjson_0.2.20
## [88] bitops_1.0-7
## [89] KernSmooth_2.23-20
## [90] Biostrings_2.60.0
## [91] blob_1.2.1
## [92] mixsqp_0.3-43
## [93] stringr_1.4.0
## [94] SQUAREM_2021.1
## [95] qvalue_2.24.0
```

```
## [96] ShortRead_1.50.0
## [97] brew_1.0-6
## [98] jpeg_0.1-8.1
## [99] ggsignif_0.6.2
## [100] reactome.db_1.76.0
## [101] scales_1.1.1
## [102] memoise_2.0.0
## [103] graphite_1.38.0
## [104] GSEABase_1.54.0
## [105] magrittr_2.0.1
## [106] plyr_1.8.6
## [107] gplots_3.1.1
## [108] zlibbioc_1.38.0
## [109] compiler_4.1.0
## [110] scatterpie_0.1.6
## [111] BiocIO_1.2.0
## [112] bbmle_1.0.23.1
## [113] plotrix_3.8-1
## [114] DESeq2_1.32.0
## [115] cli_2.5.0
## [116] Rsamtools_2.8.0
## [117] systemPipeR_1.26.0
## [118] XVector_0.32.0
## [119] Category_2.58.0
## [120] patchwork_1.1.1
## [121] MASS_7.3-54
## [122] tidyselect_1.1.1
## [123] stringi_1.6.2
## [124] forcats_0.5.1
## [125] highr_0.9
## [126] emdbook_1.3.12
## [127] yaml_2.2.1
## [128] GOSemSim_2.18.0
## [129] locfit_1.5-9.4
## [130] latticeExtra_0.6-29
## [131] ggrepel_0.9.1
## [132] grid_4.1.0
## [133] VariantAnnotation_1.38.0
## [134] fastmatch_1.1-0
```

```
## [135] tools_4.1.0
## [136] rio_0.5.27
## [137] rstudioapi_0.13
## [138] foreign_0.8-81
## [139] TxDb.Hsapiens.UCSC.hg19.knownGene_3.2.2
## [140] metapod_1.0.0
## [141] gridExtra_2.3
## [142] farver_2.1.0
## [143] gggraph_2.0.5
## [144] digest_0.6.27
## [145] rvcheck_0.1.8
## [146] BiocManager_1.30.15
## [147] Rcpp_1.0.6
## [148] car_3.0-11
## [149] broom_0.7.8
## [150] httr_1.4.2
## [151] colorspace_2.0-2
## [152] XML_3.99-0.6
## [153] truncnorm_1.0-8
## [154] splines_4.1.0
## [155] RBGL_1.68.0
## [156] tidytree_0.3.4
## [157] graphlayouts_0.7.1
## [158] xtable_1.8-4
## [159] jsonlite_1.7.2
## [160] ggtree_3.0.2
## [161] tidygraph_1.2.0
## [162] R6_2.5.0
## [163] pillar_1.6.1
## [164] htmltools_0.5.1.1
## [165] glue_1.4.2
## [166] fastmap_1.1.0
## [167] BiocParallel_1.26.0
## [168] fgsea_1.18.0
## [169] GreyListChIP_1.24.0
## [170] mvtnorm_1.1-1
## [171] utf8_1.2.1
## [172] lattice_0.20-44
## [173] tibble_3.1.2
```

```
## [174] numDeriv_2016.8-1.1
## [175] curl_4.3.2
## [176] gtools_3.9.2
## [177] zip_2.2.0
## [178] GO.db_3.13.0
## [179] openxlsx_4.2.4
## [180] survival_3.2-11
## [181] limma_3.48.0
## [182] rmarkdown_2.9
## [183] munsell_0.5.0
## [184] DO.db_2.9
## [185] GenomeInfoDbData_1.2.6
## [186] haven_2.4.1
## [187] reshape2_1.4.4
## [188] gtable_0.3.0
```