Perform Enrichment Analysis

Loading required libraries

```
library(clusterProfiler)
library(org.Dm.eg.db)
library(pathview)
library(biomaRt)
library(ReactomePA)
library(ggplot2)
library(RColorBrewer)
library(ggsci)
```

Read significant genes into csv file

```
sig_genes <- function(file_path){
  genes <- read.csv(file_path, row.names = 1, stringsAsFactors = FALSE)
  final_genes <- na.omit(genes)
  return(final_genes)

}
significant_genes <- sig_genes("../Enrichment analysis in R/Data/Significant genes.csv")
head(significant_genes)
```

```
##              baseMean log2FoldChange     lfcSE      stat       pvalue
## FBgn0039155  730.5958       4.619014 0.1687068  27.37895 1.294976e-76
## FBgn0003360 4343.0354       3.179672 0.1435262  22.15395 6.158877e-32
## FBgn0039827  261.9162       4.162516 0.2325888  17.89646 1.212750e-30
## FBgn0025111 1501.4105      -2.899864 0.1269205 -22.84788 1.377576e-28
## FBgn0034736  225.8764       3.511439 0.2146721  16.35722 3.632828e-21
## FBgn0035085  638.2326       2.560412 0.1372952  18.64896 5.654840e-15
##                     padj
## FBgn0039155 1.600332e-72
## FBgn0003360 3.805570e-28
## FBgn0039827 4.995723e-27
## FBgn0025111 4.256022e-25
## FBgn0034736 8.978897e-18
## FBgn0035085 1.164708e-11
```

Convert FlyBase IDs (FBgn) to ENTREZ IDs

```
entrez_ids <- function(){
# copy the rownames of significant_genes and store it in flybase_ids
flybase_ids <- rownames(significant_genes)
# convert FLYBASE ids into ENTREZID
entrez_ids <- mapIds(org.Dm.eg.db,
```

```
                    keys = flybase_ids,
                    column = "ENTREZID",
                    keytype = "FLYBASE",
                    multiVals = "first")
# create column named ENTREZID in significant_genes that contain ENTREZID
# of sig genes
significant_genes$ENTREZID <- entrez_ids
# remove ENTREZID that contain NA
sign_genes <- significant_genes[!is.na(significant_genes$ENTREZID), ]
return(sign_genes)
}

sign_genes <- entrez_ids()
```

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

```
head(sign_genes)
```

```
##              baseMean log2FoldChange     lfcSE      stat       pvalue
## FBgn0039155  730.5958       4.619014 0.1687068  27.37895 1.294976e-76
## FBgn0003360 4343.0354       3.179672 0.1435262  22.15395 6.158877e-32
## FBgn0039827  261.9162       4.162516 0.2325888  17.89646 1.212750e-30
## FBgn0025111 1501.4105      -2.899864 0.1269205 -22.84788 1.377576e-28
## FBgn0034736  225.8764       3.511439 0.2146721  16.35722 3.632828e-21
## FBgn0035085  638.2326       2.560412 0.1372952  18.64896 5.654840e-15
##                      padj ENTREZID
## FBgn0039155 1.600332e-72    42865
## FBgn0003360 3.805570e-28    32007
## FBgn0039827 4.995723e-27    43689
## FBgn0025111 4.256022e-25    32008
## FBgn0034736 8.978897e-18    37572
## FBgn0035085 1.164708e-11    37991
```

Up-regulated significant genes

```
up_sig_genes <- function(){
  upregulated_genes <- sign_genes[ sign_genes$padj<0.01 & sign_genes$log2FoldChange>0 , ]
  return(upregulated_genes)
}

upregulated_sig_genes <- up_sig_genes()
upregulated_sig_genes
```

```
##              baseMean log2FoldChange      lfcSE      stat       pvalue
## FBgn0039155  730.59581       4.619014 0.16870676 27.378948 1.294976e-76
## FBgn0003360 4343.03540       3.179672 0.14352623 22.153947 6.158877e-32
## FBgn0039827  261.91624       4.162516 0.23258876 17.896463 1.212750e-30
## FBgn0034736  225.87636       3.511439 0.21467209 16.357221 3.632828e-21
## FBgn0035085  638.23261       2.560412 0.13729520 18.648955 5.654840e-15
```

```
## FBgn0029167 3706.11653       2.197000 0.09698887 22.652087 3.326229e-13
## FBgn0085359   68.61004       4.918127 0.49598137  9.915951 2.757558e-12
## FBgn0024288   58.85110       4.585836 0.46501165  9.861766 1.611111e-11
## FBgn0029896  489.89237       2.445024 0.15194279 16.091743 2.492025e-10
## FBgn0038832  301.92236       2.541149 0.20251626 12.547876 1.365942e-07
##                     padj ENTREZID
## FBgn0039155 1.600332e-72    42865
## FBgn0003360 3.805570e-28    32007
## FBgn0039827 4.995723e-27    43689
## FBgn0034736 8.978897e-18    37572
## FBgn0035085 1.164708e-11    37991
## FBgn0029167 5.138192e-10    39529
## FBgn0085359 3.786434e-09  2768869
## FBgn0024288 1.991011e-08    45039
## FBgn0029896 2.368957e-07    31612
## FBgn0038832 1.125354e-04    42468
```

ENTREZIDs of up-regulated significant genes

```
up_entrez_ids <- upregulated_sig_genes$ENTREZID
up_entrez_ids
```

```
##  [1] "42865"   "32007"   "43689"   "37572"   "37991"   "39529"   "2768869"
##  [8] "45039"   "31612"   "42468"
```

Down-regulated significant genes

```
down_sig_genes <- function(){
  downregulated_genes <- sign_genes[ sign_genes$padj<0.01 & sign_genes$log2FoldChange<0 , ]
  return(downregulated_genes)
}
downregulated_sig_genes <- down_sig_genes()
downregulated_sig_genes
```

```
##               baseMean log2FoldChange     lfcSE       stat       pvalue
## FBgn0025111 1501.41051      -2.899864 0.1269205 -22.847884 1.377576e-28
## FBgn0000071  342.23841      -2.679584 0.1826118 -14.673665 5.251900e-11
## FBgn0035189  215.64911      -2.974958 0.2535481 -11.733308 2.990802e-09
## FBgn0032405   79.61406      -2.672618 0.2397721 -11.146491 5.027699e-07
## FBgn0051092  153.06560      -2.327711 0.1767464 -13.169782 1.413227e-06
## FBgn0037290   67.09796      -3.010273 0.3300244  -9.121367 2.366946e-06
##                     padj ENTREZID
## FBgn0025111 4.256022e-25    32008
## FBgn0000071 5.408582e-08    40831
## FBgn0035189 2.640023e-06    38124
## FBgn0032405 3.883269e-04    34627
## FBgn0051092 1.027333e-03    43105
## FBgn0037290 1.625040e-03    40613
```

ENTREZIDs of Down-regulated significant genes

```
down_entrez_ids <- downregulated_sig_genes$ENTREZID
down_entrez_ids
```

```
## [1] "32008" "40831" "38124" "34627" "43105" "40613"
```

Gene Ontology

Group up-regulated significant genes that have similar BP GO terms

```
go_up <- function(){
    go <- groupGO( gene = up_entrez_ids,
                   OrgDb = org.Dm.eg.db,
                   ont = "BP",  # Biological Process
                   readable = TRUE)
    return(go)

}
go_terms_up <- go_up()
head(go_terms_up)
```

```
##                    ID           Description Count GeneRatio
## GO:0000003 GO:0000003          reproduction     2      2/10
## GO:0002376 GO:0002376 immune system process     1      1/10
## GO:0008152 GO:0008152     metabolic process     3      3/10
## GO:0009987 GO:0009987      cellular process     5      5/10
## GO:0016032 GO:0016032         viral process     0      0/10
## GO:0022414 GO:0022414  reproductive process     2      2/10
##                                    geneID
## GO:0000003                     sesB/Sox100B
## GO:0002376                              Hml
## GO:0008152              sesB/CG1544/Sox100B
## GO:0009987 Kal1/sesB/CG1544/Sox100B/CG3168
## GO:0016032
## GO:0022414                     sesB/Sox100B
```

Convert GO terms of up-regulated significant genes to DataFrame

```
dataframe_go_up <- function(){
df_go_terms_up <- as.data.frame(go_terms_up)
return(df_go_terms_up)
}
df_go_group_up <- dataframe_go_up()
```

```
df_go_group_up_top7 <- head(df_go_group_up,7)
```

Barplot of up-regulated BP GO terms

```
bar_plot_up <- function(){
p <- ggplot(df_go_group_up_top7, aes(x = reorder(Description, - Count),
                                     y = Count, fill = Description)) +
                                     geom_bar(stat = "identity") +
                                     ggtitle("BP of up-regulated GO terms") +
                                     coord_flip() +
                                     theme_bw() +
                                     scale_fill_jama()+
                                     theme(plot.title = element_text(size = 12,
                                     face = "bold", hjust = 0.5))+
                                     xlab("Description")
jpeg("../Enrichment analysis in R/outputs/BPgroup_up_barplot.jpeg")
print(p)
dev.off()
}
bar_plot_up()
```

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

Group down-regulated significant genes that have similar MF GO terms

```
go_down <- function(){
  go <- groupGO(gene = down_entrez_ids,
                OrgDb = org.Dm.eg.db,
                ont = "MF",  # Molecular function
                readable = TRUE)
  return(go)
}
go_terms_down <- go_down()
head(go_terms_down)
```

```
##                    ID                  Description Count GeneRatio
## GO:0003774 GO:0003774  cytoskeletal motor activity     0       0/6
## GO:0003824 GO:0003824           catalytic activity     2       2/6
## GO:0005198 GO:0005198 structural molecule activity     0       0/6
## GO:0005215 GO:0005215         transporter activity     1       1/6
## GO:0005488 GO:0005488                      binding     3       3/6
## GO:0009055 GO:0009055   electron transfer activity     0       0/6
##                     geneID
## GO:0003774
## GO:0003824     CG9119/firl
## GO:0005198
## GO:0005215            Ant2
## GO:0005488 Ama/CG9119/LpR2
## GO:0009055
```

Convert GO terms of down-regulated significant genes to DataFrame

```
dataframe_go_down <- function(){
df_go_terms_down <- as.data.frame(go_terms_down)
return(df_go_terms_down)
}
df_go_group_down <- dataframe_go_down()
```

```
df_go_group_down_top7 <- head(df_go_group_down,7)
```

Barplot of down-regulated MF GO terms

```
bar_plot_down <- function(){
p <- ggplot(df_go_group_down_top7, aes(x = reorder(Description, - Count),
                                       y = Count, fill = Description)) +
                                       geom_bar(stat = "identity") +
                                       ggtitle("MF of down-regulated GO terms")+
                                       coord_flip() +
                                       theme_bw() +
                                       scale_fill_jama()+
                                       theme(plot.title = element_text(size = 12,
                                       face = "bold", hjust = 0.5))+
                                       xlab("Description")
jpeg("../Enrichment analysis in R/outputs/MFgroup_down_barplot.jpeg")
print(p)
dev.off()
}
bar_plot_down()
```

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

Over-representation analysis

Go enrichment analysis

Enriched GO terms among up-regulated significant genes

```
enrich_go_up <- function(){
        ego_up <- enrichGO( gene          = up_entrez_ids,
                            OrgDb         = org.Dm.eg.db,
                            keyType       = "ENTREZID",
                            ont           = "CC",
                            pvalueCutoff  = 0.05,
                            pAdjustMethod = "BH",
                            readable      = TRUE)
        return(ego_up)
```

```
}
enrichment_go_up <- enrich_go_up()
head(enrichment_go_up)
```

```
##                    ID                         Description GeneRatio   BgRatio
## GO:0031012 GO:0031012                 extracellular matrix       2/7 271/12622
## GO:0031966 GO:0031966              mitochondrial membrane       2/7 300/12622
## GO:0030312 GO:0030312 external encapsulating structure       2/7 306/12622
## GO:0005740 GO:0005740              mitochondrial envelope       2/7 330/12622
## GO:0005918 GO:0005918                   septate junction       1/7  35/12622
## GO:0070160 GO:0070160                     tight junction       1/7  40/12622
##                 pvalue   p.adjust     qvalue      geneID Count
## GO:0031012 0.008981758 0.04674036 0.02025898    Kal1/Hml     2
## GO:0031966 0.010926205 0.04674036 0.02025898 sesB/CG1544     2
## GO:0030312 0.011350225 0.04674036 0.02025898    Kal1/Hml     2
## GO:0005740 0.013119498 0.04674036 0.02025898 sesB/CG1544     2
## GO:0005918 0.019254363 0.04674036 0.02025898      CG3770     1
## GO:0070160 0.021978871 0.04674036 0.02025898      CG3770     1
```

Dataframe of enriched GO terms for up-regulated significant genes

```
df_go_terms_up <- function(){
  df_go_term_up <- as.data.frame(enrichment_go_up)
  return(df_go_term_up)
}
df_go_up <- df_go_terms_up()
```

```
df_go_up_top7 <- head(df_go_up, 7)
```

BarPlot of up-regulated CC enriched GO terms

```
bar_plot_enriched_up <- function(){
jama_colors <- pal_jama("default")(7)
p <- ggplot(df_go_up_top7, aes(x = reorder(Description, -Count),
                               y = Count, fill = p.adjust)) +
                               geom_bar(stat = "identity") +
                               coord_flip() +
                               scale_fill_gradientn(name = "p.adjust",
                                                    colors = jama_colors) +
          labs(title = "Bar plot of CC of Enriched Up-regulated GO terms",
               x = "GO Term",
               y = "Gene Count") +
                              theme_bw() +
          theme(plot.title = element_text(size = 10,face = "bold", hjust = 0.5))
jpeg("../Enrichment analysis in R/outputs/CC_enriched_up_barplot.jpeg")
print(p)
dev.off()
}
bar_plot_enriched_up()
```

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

Dotplot of up-regulated CC enriched GO terms

```
dot_plot_enriched_up <- function(){
jama_colors <- pal_jama("default")(7)
p <- ggplot(df_go_up_top7, aes(x = reorder(Description, -Count),
                               y = Count, size = Count, color = p.adjust)) +
                               geom_point(alpha = 0.6) +
                               coord_flip() +
                               scale_size_continuous(range = c(3, 8),
                                                     name = "Gene Count") +
                               scale_color_gradientn(name = "p.adjust",
                                                     colors = jama_colors) +
               labs(title = "Dot plot of CC of enriched up-regulated genes",
                    x = "GO Term",
                    y = "Gene Count") +
                               theme_bw() +
                               theme(plot.title = element_text(size = 10,
                                                               face = "bold",
                                                               hjust = 0.5))
jpeg("../Enrichment analysis in R/outputs/CC_enriched_up_dotplot.jpeg")
print(p)
dev.off()
}
dot_plot_enriched_up()
```

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

Network plot for enriched GO terms among up-regulated genes

```
jpeg("../Enrichment analysis in R/outputs/Network_plot_up.jpeg")
cnet_plot_up <- cnetplot(enrichment_go_up, showCategory = 3, vertex.label.cex = 1.2)
print(cnet_plot_up)
dev.off()
```

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

GO graph for enriched GO terms among up-regulated significant genes

```
jpeg("../Enrichment analysis in R/outputs/GO_graph_up.jpeg")
go_graph_up <- plotGOgraph(enrichment_go_up)
```

```
##
## groupGOTerms:    GOBPTerm, GOMFTerm, GOCCTerm environments built.
```

```
##
## Building most specific GOs .....

##  ( 33 GO terms found. )

##
## Build GO DAG topology ..........

##  ( 33 GO terms and 50 relations. )

##
## Attaching package: 'SparseM'

## The following object is masked from 'package:base':
##
##     backsolve

##
## Annotating nodes ...............

##  ( 12622 genes annotated to the GO terms. )

## Loading required package: Rgraphviz

## Loading required package: graph

## Loading required package: grid

##
## Attaching package: 'Rgraphviz'

## The following objects are masked from 'package:IRanges':
##
##     from, to

## The following objects are masked from 'package:S4Vectors':
##
##     from, to
```

```
print(go_graph_up)
```

```
## $dag
## A graphNEL graph with directed edges
## Number of Nodes = 28
## Number of Edges = 42
##
## $complete.dag
## [1] "A graph with 28 nodes."
```

```
dev.off()
```

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

Enriched GO terms for down-regulated significant genes

```
enrich_go_down <- function(){
        ego_down <- enrichGO( gene          = down_entrez_ids,
                              OrgDb         = org.Dm.eg.db,
                              keyType       = "ENTREZID",
                              ont           = "MF",
                              pvalueCutoff  = 0.05,
                              pAdjustMethod = "BH",
                              readable      = TRUE)
        return(ego_down)

}

enrichment_go_down <- enrich_go_down()
head(enrichment_go_down)
```

```
##                    ID
## GO:0005346 GO:0005346
## GO:0015216 GO:0015216
## GO:0015215 GO:0015215
## GO:0015605 GO:0015605
## GO:1901505 GO:1901505
## GO:0015932 GO:0015932
##                                                                  Description
## GO:0005346          purine ribonucleotide transmembrane transporter activity
## GO:0015216              purine nucleotide transmembrane transporter activity
## GO:0015215                     nucleotide transmembrane transporter activity
## GO:0015605          organophosphate ester transmembrane transporter activity
## GO:1901505        carbohydrate derivative transmembrane transporter activity
## GO:0015932 nucleobase-containing compound transmembrane transporter activity
##            GeneRatio  BgRatio      pvalue   p.adjust     qvalue geneID Count
## GO:0005346       1/5 10/12435 0.004015092 0.02668318 0.01034805   Ant2     1
## GO:0015216       1/5 11/12435 0.004415890 0.02668318 0.01034805   Ant2     1
## GO:0015215       1/5 13/12435 0.005217101 0.02668318 0.01034805   Ant2     1
## GO:0015605       1/5 14/12435 0.005617513 0.02668318 0.01034805   Ant2     1
## GO:1901505       1/5 23/12435 0.009215419 0.03044605 0.01180733   Ant2     1
## GO:0015932       1/5 24/12435 0.009614543 0.03044605 0.01180733   Ant2     1
```

Dataframe of enriched GO terms among down-regulated significant genes

```
df_go_terms_down <- function(){
df_go_term_down <- as.data.frame(enrichment_go_down)
return(df_go_term_down)
```

```
}
df_go_down <- df_go_terms_down()
```

BarPlot of down-regulated MF enriched GO terms

```
bar_plot_enriched_down<- function(){
jama_colors <- pal_jama("default")(7)
p <- ggplot(df_go_down, aes(x = reorder(Description, - Count) ,
                            y = Count, fill = p.adjust)) +
                            geom_bar(stat = "identity") +
                            coord_flip() +
                            scale_fill_gradientn(name = "p.adjust",
                                                 colors = jama_colors) +
            labs(title = "Bar plot of MF of Enriched down-regulated Genes",
                            x = "GO Term",
                            y = "Gene Count") +
                            theme_bw() +
                            theme(plot.title = element_text(size = 10,
                                                            face = "bold",
                                                            hjust = 0.5))
jpeg("../Enrichment analysis in R/outputs/MF_enriched_down_barplot.jpeg",
width = 700, height = 800)
print(p)
dev.off()
}
bar_plot_enriched_down()
```

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

Dotplot of down-regulated MF enriched GO terms

```
dot_plot_enriched_down<- function(){
jama_colors <- pal_jama("default")(7)
p <- ggplot(df_go_down, aes(x = reorder(Description, -Count),
                            y = Count, size = Count,
                            color = p.adjust)) +
                            geom_point(alpha = 0.6) +
                            coord_flip() +
                            scale_size_continuous(range = c(3, 8),
                                                  name = "Gene Count") +
                            scale_color_gradientn(name = "p.adjust",
                                                  colors = jama_colors) +
             labs(title = "Dot plot of MF of enriched down-regulated genes",
                            x = "GO Term",
                            y = "Gene Count") +
                            theme_bw() +
                            theme(plot.title = element_text(size = 10,
                                                            face = "bold",
```

```
                                                              hjust = 0.5))
jpeg("../Enrichment analysis in R/outputs/MF_enriched_down_dotplot.jpeg",
width = 700, height = 800)
print(p)
dev.off()
}
dot_plot_enriched_down()
```

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

Network plot for enriched GO terms among down-regulated genes

```
jpeg("../Enrichment analysis in R/outputs/Network_plot_down.jpeg")
cnet_plot_down <- cnetplot(enrichment_go_down, showCategory = 7, vertex.label.cex = 1.2)
print(cnet_plot_down)
dev.off()
```

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

GO graph for enriched GO terms among down-regulated significant genes

```
jpeg("../Enrichment analysis in R/outputs/GO_graph_down.jpeg")
go_graph_down <- plotGOgraph(enrichment_go_down)
```

```
##
## groupGOTerms:    GOBPTerm, GOMFTerm, GOCCTerm environments built.

##
## Building most specific GOs .....

##  ( 41 GO terms found. )

##
## Build GO DAG topology ..........

##  ( 41 GO terms and 50 relations. )

##
## Annotating nodes ...............

##  ( 12435 genes annotated to the GO terms. )
```

```
print(go_graph_down)
```

```
## $dag
## A graphNEL graph with directed edges
## Number of Nodes = 12
## Number of Edges = 13
##
## $complete.dag
## [1] "A graph with 12 nodes."
```

```
dev.off()
```

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

Pathway enrichment analysis

Extract sig gene names from ensembl database

```
ensembl_gene_names <- function() {
  # Connect to the Ensembl database
  ensembl <- useMart("ensembl", dataset = "dmelanogaster_gene_ensembl")
  # list of Drosophila melanogaster FlyBase gene IDs
  flybase_gene_ids <- rownames(sign_genes)
  # Convert FlyBase gene IDs to external gene identifiers
  gene_info <- getBM(attributes = c("ensembl_gene_id", "external_gene_name"), filters = "flybase_gene_i
  return(gene_info)
}
gene_info_result <- ensembl_gene_names()
head(gene_info_result)
```

```
##   ensembl_gene_id external_gene_name
## 1     FBgn0000003     7SLRNA:CR32864
## 2     FBgn0000008                  a
## 3     FBgn0000014              abd-A
## 4     FBgn0000015              Abd-B
## 5     FBgn0000017                Abl
## 6     FBgn0000018                abo
```

Up-regulated Dmel_XXXX ids

```
dmel_up <- function(){
  filtered_up_sig_genes <- gene_info_result[gene_info_result$ensembl_gene_id %in%   rownames(upregulate
  dmel_up_identifiers <- paste("Dmel", filtered_up_sig_genes$external_gene_name,
                               sep = "_")
  return(dmel_up_identifiers)
}
dmel_up_genes <- dmel_up()
dmel_up_genes
```

```
##  [1] "Dmel_sesB"    "Dmel_Sox100B" "Dmel_Hml"     "Dmel_CG3168"  "Dmel_gas"
##  [6] "Dmel_CG3770"  "Dmel_CG15695" "Dmel_Kal1"    "Dmel_CG1544"  "Dmel_CG34330"
```

KEGG pathway enrichment analysis among up-regulated significant genes

```
kegg_enrichment_up <- function(){
kegg_up <- enrichKEGG(gene = dmel_up_genes,
                      organism = "dme", # Drosophila melanogaster
                      pvalueCutoff = 0.05)
  return(kegg_up)
}
kegg_enrich_up <- kegg_enrichment_up()
```

```
## Reading KEGG annotation online: "https://rest.kegg.jp/link/dme/pathway"...
```

```
## Reading KEGG annotation online: "https://rest.kegg.jp/list/pathway/dme"...
```

```
head(kegg_enrich_up)
```

```
##                                    category
## dme04512 Environmental Information Processing
## dme00785                          Metabolism
## dme00380                          Metabolism
## dme00310                          Metabolism
## dme01210                          Metabolism
##                                   subcategory       ID
## dme04512  Signaling molecules and interaction dme04512
## dme00785 Metabolism of cofactors and vitamins dme00785
## dme00380               Amino acid metabolism dme00380
## dme00310               Amino acid metabolism dme00310
## dme01210            Global and overview maps dme01210
##                                                                  Description
## dme04512        ECM-receptor interaction - Drosophila melanogaster (fruit fly)
## dme00785          Lipoic acid metabolism - Drosophila melanogaster (fruit fly)
## dme00380           Tryptophan metabolism - Drosophila melanogaster (fruit fly)
## dme00310              Lysine degradation - Drosophila melanogaster (fruit fly)
## dme01210 2-Oxocarboxylic acid metabolism - Drosophila melanogaster (fruit fly)
##          GeneRatio BgRatio      pvalue   p.adjust qvalue      geneID Count
## dme04512       1/2 11/3587 0.006124707 0.01946050     NA Dmel_CG3168     1
## dme00785       1/2 18/3587 0.010012453 0.01946050     NA Dmel_CG1544     1
## dme00380       1/2 21/3587 0.011676297 0.01946050     NA Dmel_CG1544     1
## dme00310       1/2 35/3587 0.019422401 0.01997453     NA Dmel_CG1544     1
## dme01210       1/2 36/3587 0.019974528 0.01997453     NA Dmel_CG1544     1
```

DataFrame of KEGG enriched among up-regulated significant genes

```
dataframe_kegg_up <- function(){
  df_kegg_up_genes <- as.data.frame(kegg_enrich_up)
```

```
  return(df_kegg_up_genes)
}
df_kegg_up <- dataframe_kegg_up()
head(df_kegg_up)
```

```
##                                     category
## dme04512 Environmental Information Processing
## dme00785                          Metabolism
## dme00380                          Metabolism
## dme00310                          Metabolism
## dme01210                          Metabolism
##                                   subcategory       ID
## dme04512  Signaling molecules and interaction dme04512
## dme00785 Metabolism of cofactors and vitamins dme00785
## dme00380                Amino acid metabolism dme00380
## dme00310                Amino acid metabolism dme00310
## dme01210             Global and overview maps dme01210
##                                                                  Description
## dme04512        ECM-receptor interaction - Drosophila melanogaster (fruit fly)
## dme00785          Lipoic acid metabolism - Drosophila melanogaster (fruit fly)
## dme00380           Tryptophan metabolism - Drosophila melanogaster (fruit fly)
## dme00310              Lysine degradation - Drosophila melanogaster (fruit fly)
## dme01210 2-Oxocarboxylic acid metabolism - Drosophila melanogaster (fruit fly)
##          GeneRatio BgRatio      pvalue   p.adjust qvalue      geneID Count
## dme04512       1/2 11/3587 0.006124707 0.01946050     NA Dmel_CG3168     1
## dme00785       1/2 18/3587 0.010012453 0.01946050     NA Dmel_CG1544     1
## dme00380       1/2 21/3587 0.011676297 0.01946050     NA Dmel_CG1544     1
## dme00310       1/2 35/3587 0.019422401 0.01997453     NA Dmel_CG1544     1
## dme01210       1/2 36/3587 0.019974528 0.01997453     NA Dmel_CG1544     1
```

Barplot for KEGG enriched up-regulated significant gene

```
bar_plot_kegg_enriched_up<- function(){
jama_colors <- pal_jama("default")(7)
p <- ggplot(df_kegg_up, aes(x = reorder(Description, - Count),
                            y = Count, fill = p.adjust)) +
                            geom_bar(stat = "identity", width = 0.8) +
                            coord_flip() +
                            scale_fill_gradientn(name = "p.adjust",
                                                 colors = jama_colors) +
      labs(title = "Bar plot of Enriched KEGG pathway among up-regulated Genes",
           x = "Enriched pathway",
           y = "Gene Count") +
                            theme_bw() +
                            theme(plot.title = element_text(size = 10,
                                                            face = "bold",
                                                            hjust = 0.5))
jpeg("../Enrichment analysis in R/outputs/kegg_enriched_up_barplot.jpeg",
width = 700, height = 800)
print(p)
dev.off()
```

```
}
bar_plot_kegg_enriched_up()
```

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

Dotplot for Enriched KEGG pathway among up-regulated significant gene

```
dot_plot_kegg_enriched_up<- function(){
jama_colors <- pal_jama("default")(7)
p <- ggplot(df_kegg_up, aes(x = reorder(Description, -Count),
                            y = Count, size = Count,
                            color = p.adjust))+
                            geom_point(alpha = 0.6) +
                            coord_flip() +
                            scale_size_continuous(range = c(3, 8),
                                                  name = "Gene Count") +
                            scale_color_gradientn(name = "p.adjust",
                                                  colors = jama_colors) +
      labs(title = "Dot plot of Enriched KEGG pathway among up-regulated genes",
                            x = "GO Term",
                            y = "Gene Count") +
                            theme_bw() +
                            theme(plot.title = element_text(size = 10,
                                                            face = "bold",
                                                            hjust = 0.5))
jpeg("../Enrichment analysis in R/outputs/kegg_enriched_up_dotplot.jpeg",
width = 800, height = 800)
print(p)
dev.off()
}
dot_plot_kegg_enriched_up()
```

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

Visualize a specific pathway

visualize the top enriched KEGG pathway among up-regulated sig genes

```
visualize_top_path_up <- function(){
  pathview(gene.data = up_entrez_ids,
           pathway.id = "dme04512",
           species = "dme",
           kegg.dir = "../Enrichment analysis in R/outputs/")
}
visualize_top_path_up()
```

```
## Info: Getting gene ID data from KEGG...

## Info: Done with data retrieval!

## Info: Working in directory C:/Users/HP/Desktop/Mentor M. Soudy/Bulk RNAseq analysis/Enrichment analys

## Info: Writing image file dme04512.pathview.png
```

Browse the top enriched pathway among up-regulated sig genes

```
browseKEGG(kegg_enrich_up, 'dme04512')
```

Downregulated Dmel_XXXX ids

```
dmel_down <- function(){
  filtered_down_sig_genes <- gene_info_result[gene_info_result$ensembl_gene_id %in% rownames(downregula

  dmel_down_identifiers <- paste("Dmel", filtered_down_sig_genes$external_gene_name,sep = "_")
  return(dmel_down_identifiers)
}
dmel_down_genes <- dmel_down()
dmel_down_genes
```

```
## [1] "Dmel_Ama"    "Dmel_Ant2"   "Dmel_firl"   "Dmel_CG9119" "Dmel_CG1124"
## [6] "Dmel_LpR2"
```

KEGG pathway enrichment analysis among down-regulated significant genes

```
kegg_enrichment_down <- function(){
  kegg_down <- enrichKEGG(gene = dmel_down_genes,
                          organism = "dme", # Drosophila melanogaster
                          pvalueCutoff = 0.05,
                          pAdjustMethod = "BH")
  if (is.null(kegg_down) || nrow(kegg_down) == 0) {
      print("No enriched pathways found")
    }
  return(kegg_down)
}
kegg_enrich_down <- kegg_enrichment_down()
```

```
## [1] "No enriched pathways found"
```

Reactome pathway enrichment analysis among down-regulated significant genes

```
reactome_down_genes <- function(){
  reactome_enrichment <- enrichPathway(gene = down_entrez_ids,
                                       organism = "fly", # Drosophila melanogaster
                                       pvalueCutoff = 0.05)

  return(reactome_enrichment)
}
reactome_enriched_path_down <- reactome_down_genes()
head(reactome_enriched_path_down)
```

```
##                          ID
## R-DME-8964038 R-DME-8964038
## R-DME-8964043 R-DME-8964043
## R-DME-1268020 R-DME-1268020
## R-DME-5365859 R-DME-5365859
## R-DME-425397   R-DME-425397
## R-DME-5362517 R-DME-5362517
##                                                                 Description
## R-DME-8964038                                                 LDL clearance
## R-DME-8964043                                 Plasma lipoprotein clearance
## R-DME-1268020                                 Mitochondrial protein import
## R-DME-5365859                                     RA biosynthesis pathway
## R-DME-425397  Transport of vitamins, nucleosides, and related molecules
## R-DME-5362517                                   Signaling by Retinoic Acid
##               GeneRatio BgRatio      pvalue   p.adjust      qvalue geneID Count
## R-DME-8964038       1/3 12/4698 0.007644902 0.03851761 0.009356504  43105     1
## R-DME-8964043       1/3 18/4698 0.011452699 0.03851761 0.009356504  43105     1
## R-DME-1268020       1/3 19/4698 0.012086383 0.03851761 0.009356504  32008     1
## R-DME-5365859       1/3 19/4698 0.012086383 0.03851761 0.009356504  34627     1
## R-DME-425397        1/3 26/4698 0.016514591 0.03851761 0.009356504  32008     1
## R-DME-5362517       1/3 28/4698 0.017777358 0.03851761 0.009356504  34627     1
```

DataFrame of Reactome enriched pathways among down-regulated significant genes

```
dataframe_reactome_down <- function(){
  df_reactome_down_genes <- as.data.frame(reactome_enriched_path_down)
  return(df_reactome_down_genes)
}
df_reactome_down <- dataframe_reactome_down()
head(df_reactome_down)
```

```
##                          ID
## R-DME-8964038 R-DME-8964038
## R-DME-8964043 R-DME-8964043
## R-DME-1268020 R-DME-1268020
## R-DME-5365859 R-DME-5365859
## R-DME-425397   R-DME-425397
## R-DME-5362517 R-DME-5362517
##                                                                 Description
```

```
## R-DME-8964038                                                LDL clearance
## R-DME-8964043                             Plasma lipoprotein clearance
## R-DME-1268020                             Mitochondrial protein import
## R-DME-5365859                                  RA biosynthesis pathway
## R-DME-425397  Transport of vitamins, nucleosides, and related molecules
## R-DME-5362517                               Signaling by Retinoic Acid
##               GeneRatio BgRatio      pvalue   p.adjust      qvalue geneID Count
## R-DME-8964038       1/3 12/4698 0.007644902 0.03851761 0.009356504  43105     1
## R-DME-8964043       1/3 18/4698 0.011452699 0.03851761 0.009356504  43105     1
## R-DME-1268020       1/3 19/4698 0.012086383 0.03851761 0.009356504  32008     1
## R-DME-5365859       1/3 19/4698 0.012086383 0.03851761 0.009356504  34627     1
## R-DME-425397        1/3 26/4698 0.016514591 0.03851761 0.009356504  32008     1
## R-DME-5362517       1/3 28/4698 0.017777358 0.03851761 0.009356504  34627     1
```

Visualize Reactome Pathway Enrichment Results

Barplot for Reactome enriched pathways among down-regulated significant gene

```
bar_plot_reactome_enriched_down<- function(){
jama_colors <- pal_jama("default")(7)
p <- ggplot(df_reactome_down, aes(x = reorder(Description, - Count) ,
                                  y = Count, fill = p.adjust)) +
                                  geom_bar(stat = "identity", width = 0.8) +
                                  coord_flip() +
                                  scale_fill_gradientn(name = "p.adjust",
                                                       colors = jama_colors) + labs(title = "Bar plot o
     x = "Enriched pathway",
     y = "Gene Count") +
                                  theme_bw() +
                                  theme(plot.title = element_text(size = 10,
                                                                  face = "bold",
jpeg("../Enrichment analysis in R/outputs/reactome_enriched_down_barplot.jpeg",
width = 700, height = 800)
print(p)
dev.off()
}
bar_plot_reactome_enriched_down()
```

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

Dotplot for Reactome enriched pathways among downregulated significant gene

```
dot_plot_reactome_enriched_down<- function(){
jama_colors <- pal_jama("default")(7)
p <- ggplot(df_reactome_down, aes(x = reorder(Description, -Count),
                                  y = Count, size = Count,
                                  color = p.adjust)) +
                                  geom_point(alpha = 0.6) +
```

```
                                  coord_flip() +
                                  scale_size_continuous(range = c(3, 8),
                                                        name = "Gene Count") +
                                  scale_color_gradientn(name = "p.adjust",
                                                        colors = jama_colors) +
labs(title = "Dot plot of enriched Reactome pathways among down-regulated genes",
     x = "GO Term",
     y = "Gene Count") +
                                  theme_bw() +
                                  theme(plot.title = element_text(size = 10,
                                                                  face = "bold",
jpeg("../Enrichment analysis in R/outputs/reactome_enriched_down_dotplot.jpeg",
width = 700, height = 800)
print(p)
dev.off()
}
dot_plot_reactome_enriched_down()
```

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

visualize the top enriched Reactome pathway for down-regulated sig genes

Take the generated URL and browse it

```
visualize_reactome_path_up <- function(){
  # Convert the Reactome ID to a URL for visualization
  reactome_url <- paste0("https://reactome.org/PathwayBrowser/#/", "R-DME-8964038")
  # Print the URL for manual review
  print(reactome_url)
}
visualize_reactome_path_up()
```

```
## [1] "https://reactome.org/PathwayBrowser/#/R-DME-8964038"
```