Aza_Toca_Pablo-PEC2

Pablo Aza

2023-12-21

Contents

1.Introduction

Microarray technologies have revolutionized genomic research by allowing the simultaneous measurement of thousands of genes' expression. These devices provide a comprehensive view of gene activity, offering valuable data to understand the biological processes underlying normal and pathological conditions.

However, the complexity and magnitude of the datasets generated by microarrays pose significant challenges in terms of analysis and extraction of relevant information. This is why the use of bioinformatics tools is becoming increasingly important to analyze such data. In this context, the Bioconductor project offers a set of tools in R designed for the analysis of high-dimensional biological data, such as those generated by microarrays.

In this work, we will leverage the potential of Bioconductor to explore a microarray from initial data exploration steps, through differential expression analysis, to identifying overrepresented biological processes.

The selected dataset is "GDS1736" from the "GSE3737" series of Rattus norvegicus in the study "Acute hypotension effect on kidneys." The study involves profiling the gene expression of kidneys from 12-week-old normotensive Wistar-Kyoto mice subjected to acute hypotension by withdrawing 3 ml of blood. The results provide insights into vasoregulatory mechanisms and the genetics and physiology of blood pressure regulation.

2. Objectives

The main objective is to analyze the differential gene expression between the Control (normotensive) and Hypotensive Mouse samples. Additionally, we will study, through enrichment analysis, which biological processes may have been affected.

3. Materials and Methods

3.1 Data Acquisition

The data is downloaded from the GEO database using R and the GEOquery package. The series number used as a query is GSE2401.

3.2 Microarray Exploration

General exploration is performed using BoxPlot and hierarchical clustering dendrograms. The arrayQualityMetrics() function and basic R functions are used for this purpose. We conduct a Principal Component Analysis using the prcomp() function and visualize the data using the ggplot() package.

3.3. Differential Analysis

Differential expression analysis of genes is performed using the Limma package.

3.4 Enrichment Analysis

We use the rat2302.db package from the Affymetrix Rat Genome 230 2.0 Array microarray platform. The clusterProfiler library is employed for enrichment analysis.

4. Results

4.1 Data Download

o obtain the data, we use GEOquery to access the public GEO database for biological experiments. With GEOquery, we can download this data directly from R. The selected GEO Dataset is GDS1251.

- GEO Dataset: GDS1251
- GEO Serie: GSE2401
- Platform: GPL341

```r
# Install all necessary packages for analysis

if (!requireNamespace("BiocManager", quietly = TRUE)) {
  install.packages("BiocManager")
}

if (!requireNamespace("ggplot2", quietly = TRUE)) {
  install.packages("ggplot2")
}


# List of packages
packages <- c("GEOquery","limma","rat2302.db","clusterProfiler")

# Install packages
missing_packages <- packages[!(packages %in% installed.packages()[,"Package"])]
if(length(missing_packages) > 0) {
  BiocManager::install(pkgs = missing_packages)
}


require(GEOquery)
require(limma)
require(rat2302.db)
require(clusterProfiler)

#Abrimos el dataset
gds<-getGEO("GDS1251")
```

Once we have initial information about the data, we generate the content matrix using the dataset number and creating a GDS file.

With the Meta() function, we extract information about the experiment.

```r
Meta(gds)
```

```
## $channel_count
## [1] "1"
##
## $dataset_id
## [1] "GDS1251" "GDS1251"
##
## $description
## [1] "Expression profiling of kidneys of 12 week old normotensive Wistar-Kyoto males subjected to acu
## [2] "control"
## [3] "acute hypotension"
##
## $email
## [1] "geo@ncbi.nlm.nih.gov"
##
## $feature_count
## [1] "15923"
##
```

```
## $institute
## [1] "NCBI NLM NIH"
##
## $name
## [1] "Gene Expression Omnibus (GEO)"
##
## $order
## [1] "none"
##
## $platform
## [1] "GPL341"
##
## $platform_organism
## [1] "Rattus norvegicus"
##
## $platform_technology_type
## [1] "in situ oligonucleotide"
##
## $pubmed_id
## [1] "15942020"
##
## $ref
## [1] "Nucleic Acids Res. 2005 Jan 1;33 Database Issue:D562-6"
##
## $reference_series
## [1] "GSE2401"
##
## $sample_count
## [1] "9"
##
## $sample_id
## [1] "GSM45184,GSM45186,GSM45187,GSM45189,GSM45193"
## [2] "GSM45188,GSM45190,GSM45191,GSM45192"
##
## $sample_organism
## [1] "Rattus norvegicus"
##
## $sample_type
## [1] "RNA"
##
## $title
## [1] "Acute hypotension effect on kidneys"
##
## $type
## [1] "Expression profiling by array" "stress"
## [3] "stress"
##
## $update_date
## [1] "Oct 10 2012"
##
## $value_type
## [1] "count"
##
## $web_link
```

```
## [1] "http://www.ncbi.nlm.nih.gov/geo"
```

We have a description of the study: "Expression profiling of kidneys of 12-week-old normotensive Wistar-Kyoto males subjected to acute hypotension by withdrawing 3 ml of blood. Results provide insight into vasoregulatory mechanisms and the genetics and physiology of blood pressure regulation."

We also observe the two groups that compose it:

-Control -Hypotense

The next step is to extract the expression set.

```
#extraer el expresionset y eliminar valores NA
matrix.data<-GDS2eSet(gds,do.log2 = TRUE)
matrix.data<-na.omit(matrix.data)
print(matrix.data)
```

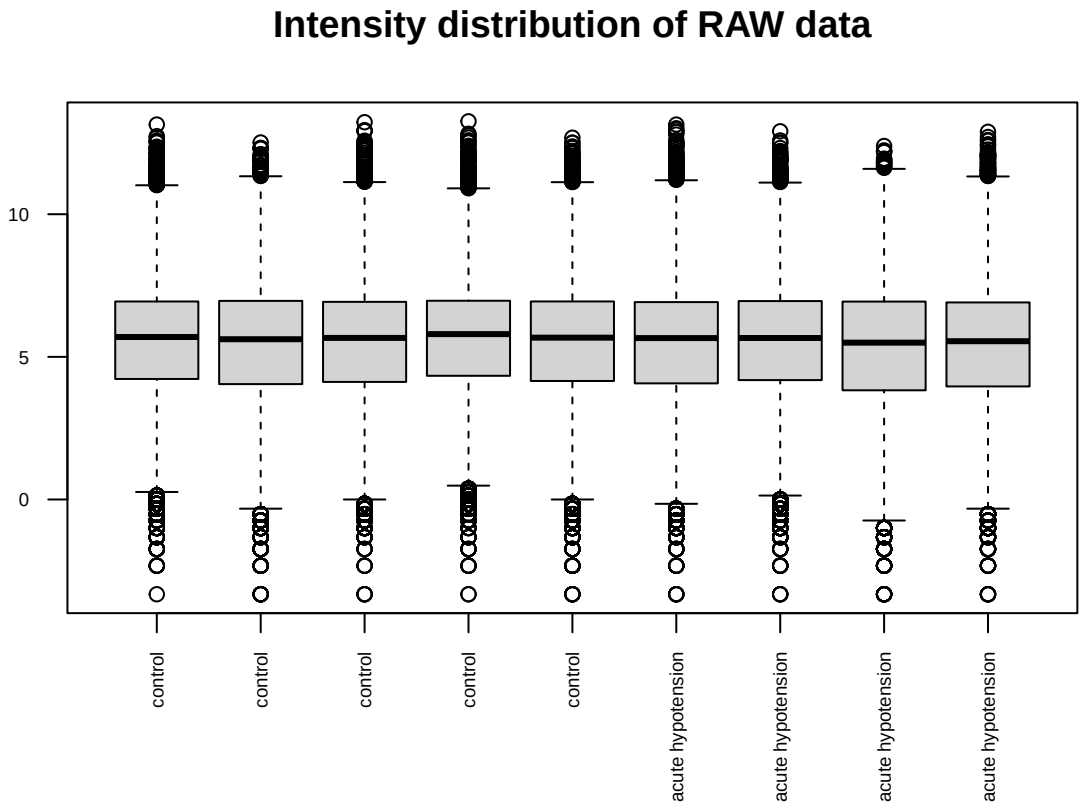
```
## ExpressionSet (storageMode: lockedEnvironment)
## assayData: 15923 features, 9 samples
##   element names: exprs
## protocolData: none
## phenoData
##   sampleNames: GSM45184 GSM45186 ... GSM45192 (9 total)
##   varLabels: sample stress description
##   varMetadata: labelDescription
## featureData
##   featureNames: 1367452_at 1367453_at ... AFFX_Rat_Hexokinase_M_at
##     (15923 total)
##   fvarLabels: ID Gene title ... GO:Component ID (21 total)
##   fvarMetadata: Column labelDescription
## experimentData: use 'experimentData(object)'
##   pubMedIds: 15942020
## Annotation:
```

4.2 Exploration, Quality Control

For exploration and quality control, we can use the "arrayQualityMetrics(raw_data, outdir ="Quality_Report_Kidney")" package. However, for convenience, we will perform data exploration with "ad-hoc" functions.

Use a boxplot to analyze the intensity distribution.

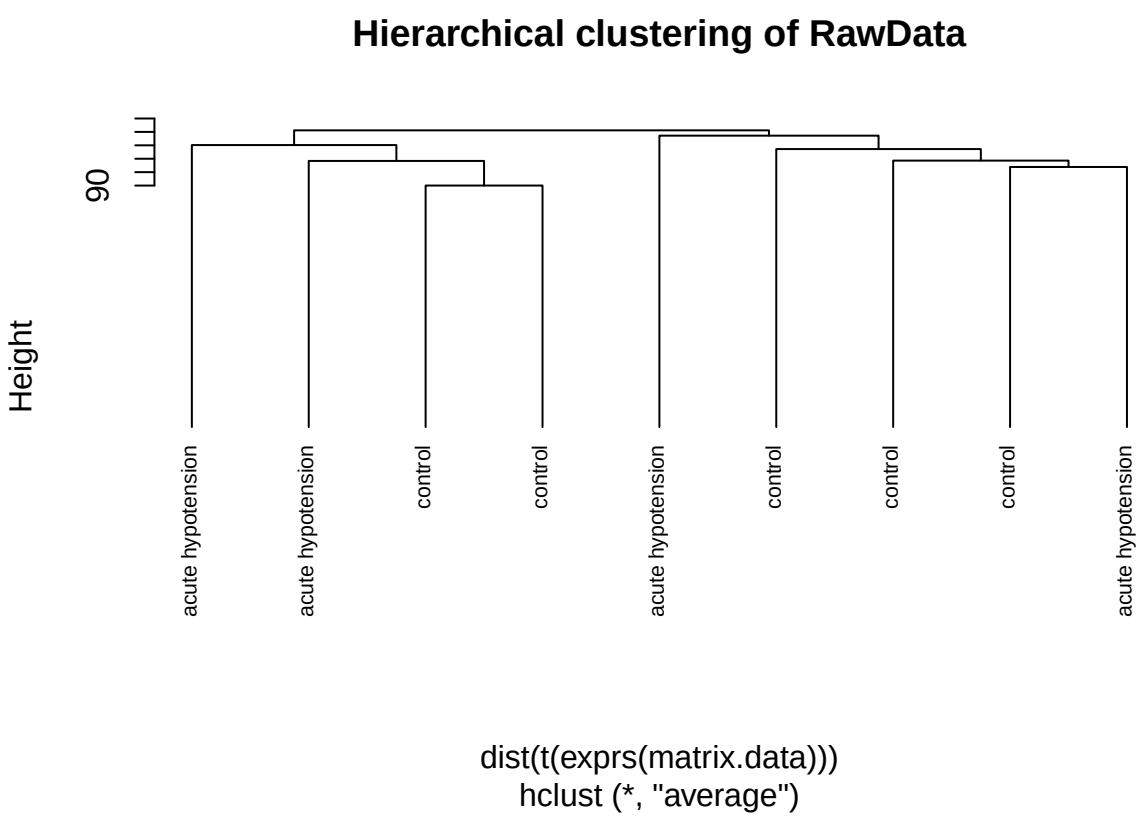
```
boxplot(exprs(matrix.data),names=pData(matrix.data)[,2], which="all", las = 2, main = "Intensity distril
    cex.axis = 0.6)
```



We observe that the distribution of signals is very similar between different samples. It is noteworthy that this is expected since the data was pre-normalized.

Next, we create a hierarchical clustering dendrogram to see how our microarrays would group.

```
# HIERARQUICAL CLUSTERING
clust.euclid.average <- hclust(dist(t(exprs(matrix.data))), method = "average")
plot(clust.euclid.average, labels = pData(matrix.data)[,2], main = "Hierarchical clustering of RawData"
    cex = 0.7, hang = -1)
```



The Cluster analysis reveals that they do not group as expected.

We perform a Principal Component Analysis (PCA) to see if we can simplify the data dimensions while maintaining the same variability. In our case, we consider genes as variables, so we transpose the dataframe. If we don't, we would see how genes are related to each other, which is not what we want.

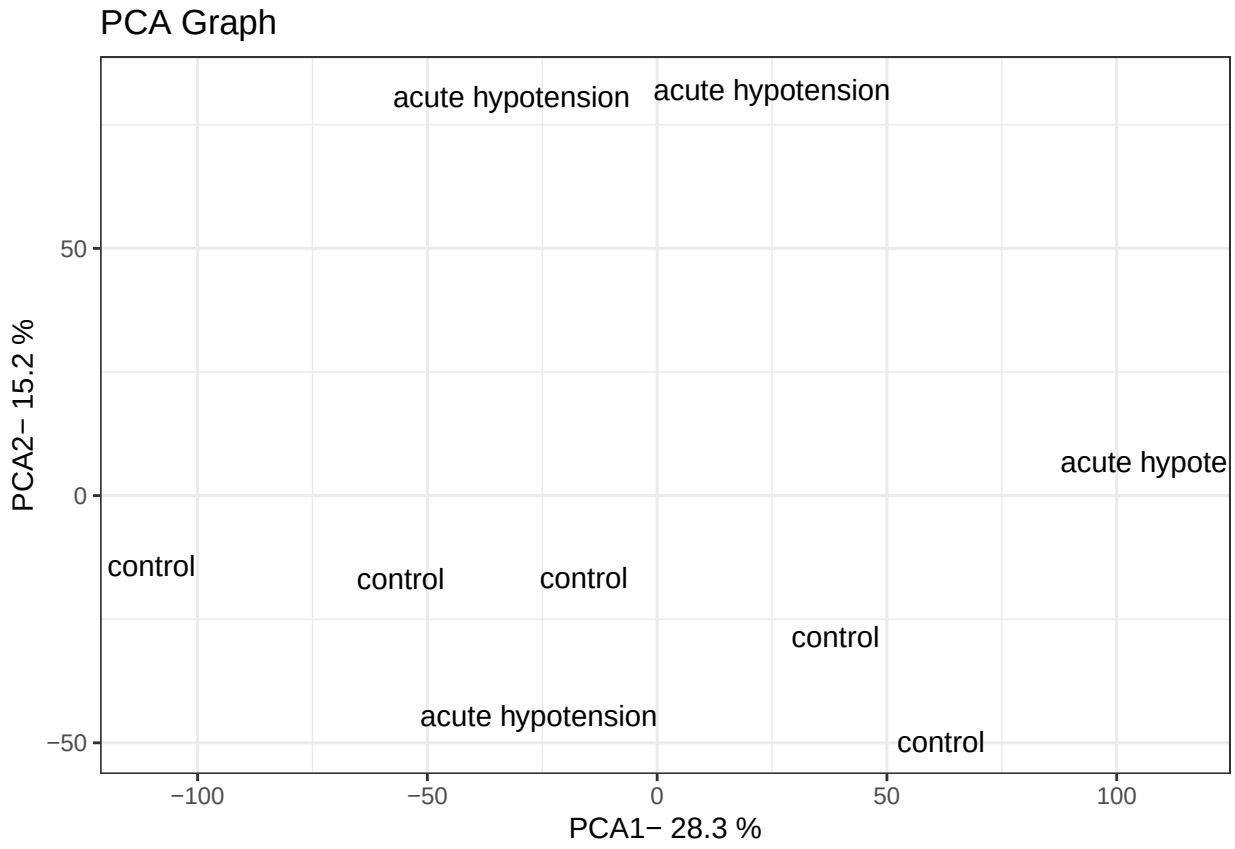
Since we have 9 samples in our analysis, we will see 9 PCs when we represent them. When we do, we will see that the first PC has the most variation in the original data, and so on.

```
# Find infinite or missing values that cause problems in PCA
any(is.na(exprs(matrix.data)))
```

```
## [1] TRUE
```

```
datos<-na.omit(exprs(matrix.data))

# Principal Component Analysis
pca<-prcomp(t(datos), scale=TRUE)

# Prepare data for ggplot
pca.data<-data.frame(Sample=pData(matrix.data)$stress,X=pca$x[,1],Y=pca$x[,2])


# Plot
library(ggplot2)
```

```
pca.var<-pca$sdev^2 #sacar desviacion estandar
pca.var.tot<-round(pca.var/sum(pca.var)*100,1) #los porcentages de DesEst explicados para cada PC

ggplot(data=pca.data, aes(x=X, y=Y, label=Sample))+ geom_text()+
  xlab(paste("PCA1-",pca.var.tot[1],"%",sep=" "))+
  ylab(paste("PCA2-",pca.var.tot[2],"%",sep=" "))+
  theme_bw()+
  ggtitle("PCA Graph")
```



The axes show the percentage of variability in the original data explained. As observed, we do not see the two clusters well differentiated as expected because controls and cases are intermingled.

PCA and dendrogram analysis could indicate a technical or experimental error, the day of sample collection, or their mixing. It could also be due to significant biological variability within each group, making it challenging to identify distinctive patterns in PCA. Additionally, there might be subtypes within the groups that were not considered in the study (e.g., gender).

Despite the above, we continue with the next section regarding the selection of differentially expressed genes.

4.3 Selection of Differentially Expressed Genes

To select differentially expressed genes, considering that we have a single factor with 2 levels (Control/Hypotension), we could use t-tests and select genes based on their adjusted p-value. However, we choose to use linear models for gene selection, given the activity's statement.

We formulate the specific question to be answered:

Which genes are differentially expressed between control and Hypotensive mice?

We display the model parameterization; a two-level factor.

$Yi, j = \alpha i + Ei, j$

Where i=1,2 because there can be up to 2 levels, and j=1, ... 5 since there can be up to 5 biological replicates.

We set up the design matrix and the linear model of the study. In our case, we have 5 controls and 4 Hypotension:

$$\begin{pmatrix} Y1 \\ Y2 \\ Y3 \\ Y4 \\ Y5 \\ Y6 \\ Y7 \\ Y8 \\ Y9 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} * \begin{pmatrix} \alpha 1 \\ \alpha 2 \end{pmatrix} + \begin{pmatrix} \epsilon 1 \\ \epsilon 2 \\ \epsilon 3 \\ \epsilon 4 \\ \epsilon 5 \\ \epsilon 6 \\ \epsilon 7 \\ \epsilon 8 \\ \epsilon 9 \end{pmatrix}$$

The contrast we are considering would be:

$\beta = \alpha 1 - \alpha 2$

With the following contrast matrix:

$$\begin{pmatrix} \beta 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \end{pmatrix} * \begin{pmatrix} \alpha 1 \\ \alpha 2 \end{pmatrix}$$

Once we have the formulation, we proceed to estimate the parameters and make the comparison. We use the limma package for this purpose.

The following code defines the design matrix and the contrasts.

```
#establecemos los niveles
lev<-factor(pData(matrix.data)$stress)

#diseñamos la matriz diseño
design<-model.matrix(~0+lev)
rownames(design)<-pData(matrix.data)$stress
colnames(design)<-c("hypo","control")
print(design)
```

```
##                   hypo control
## control              0       1
## control              0       1
## control              0       1
## control              0       1
## control              0       1
## acute hypotension    1       0
## acute hypotension    1       0
## acute hypotension    1       0
## acute hypotension    1       0
## attr(,"assign")
## [1] 1 1
## attr(,"contrasts")
```

```
## attr(,"contrasts")$lev
## [1] "contr.treatment"
```

```
#elaboramos la matriz de contraste
cont.matrix<-makeContrasts(Hypo.Control=(control-hypo),levels=design)
cont.matrix
```

```
##          Contrasts
## Levels    Hypo.Control
##   hypo              -1
##   control            1
```

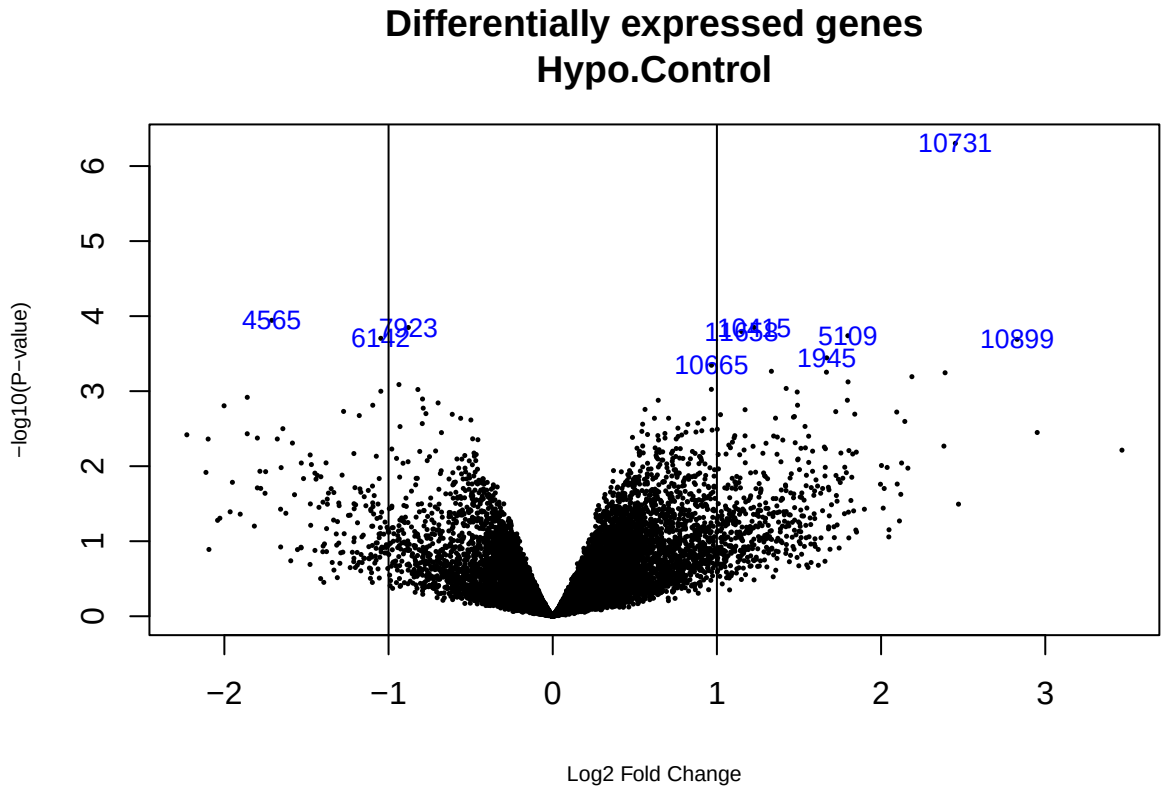
In the next step, we calculate the contrasts.

```
fit<-lmFit(matrix.data,design)#modelo linea
fit.main<-contrasts.fit(fit, cont.matrix)#contraste
fit.main<-eBayes(fit.main)#regularizar varianza
```

As a result, we have the Fold-changes and adjusted p-values, which allow us to design the "Volcano-plot" to visualize the data. It is worth noting that by using Bayes regularization, we are improving the estimation of gene variability, acting as "shrinkage," especially when the number of samples is small, as in our case: 5 for control and 4 for Hypotension.

Next, we visualize the data with a volcano plot and additionally, using the topTable function, we generate a list of genes that are differentially expressed for the contrast.

```
coefnum = 1
opt <- par(cex.lab = 0.7)
volcanoplot(fit.main, coef=coefnum, highlight=10, names=fit.main$ID,
main=paste("Differentially expressed genes",
colnames(cont.matrix)[coefnum], sep="\n"))
abline(v=c(-1,1))
```




```
par(opt)


#ordered data list
Hypo.Control <- topTable (fit.main, number=nrow(fit.main), coef="Hypo.Control", adjust="fdr")
```

4.4 Results Annotation.

As we are working with the genome of Rattus norvegicus, we searched for the reference "Annotation Packages" on Bioconductor. After searching, it seems that the appropriate package is the Affymetrix Rat230_2 Array annotation data, named "rat2302.db."

```
keytypes(rat2302.db)
```

```
##  [1] "ACCNUM"       "ALIAS"        "ENSEMBL"      "ENSEMBLPROT"  "ENSEMBLTRANS"
##  [6] "ENTREZID"     "ENZYME"       "EVIDENCE"     "EVIDENCEALL"  "GENENAME"
## [11] "GENETYPE"     "GO"           "GOALL"        "IPI"          "ONTOLOGY"
## [16] "ONTOLOGYALL"  "PATH"         "PFAM"         "PMID"         "PROBEID"
## [21] "PROSITE"      "REFSEQ"       "SYMBOL"       "UNIPROT"
```

Next, we annotate the data using identifiers such as "Symbol," "EntrezID," and "EnsemblID."

-SYMBOL: A readable identifier. -ENTREZID: A unique numerical identifier assigned to each gene by the National Center for Biotechnology Information (NCBI) through its Entrez Gene database. -ENSEMBL:

A unique numerical identifier assigned to each gene by the National Center for Biotechnology Information (NCBI) through its Entrez Gene database.

```
geneAnots <- AnnotationDbi::select(rat2302.db,rownames(Hypo.Control) , c("SYMBOL", "ENTREZID", "ENSEMBL
head(geneAnots)
```

PROBEID	SYMBOL	ENTREZID	ENSEMBL
1385507_at	Pheta1	288664	ENSRNOG00000028206
1372016_at	Gadd45b	299626	ENSRNOG00000019822
1382632_at	Robo2	84409	ENSRNOG00000029598
1375374_at	NA	NA	NA
1387539_at	Si	497756	ENSRNOG00000031067
1372560_at	NA	NA	NA

We perform an enrichment analysis, which is used to identify gene classes overrepresented in a set of interest (hypotensive mice) compared to a background set (controls).

We will conduct an "Over-Representation Analysis" using the "clusterProfiler" package. First, we select genes that are considered differentially expressed, i.e., those with a P.value less than 0.05 and a logFC greater than 1.

```
select.exp<-Hypo.Control[Hypo.Control$P.Value < 0.05 & Hypo.Control$logFC > 1, ]
dim(select.exp)
```

```
## [1] 212  27
```

Once we have the genes considered differentially expressed (a total of 212), we perform the enrichment analysis on the resulting dataset. It's worth noting that we will use "BP" (Biological Process) as the sub-ontology to focus on higher-level biological processes involving differentially expressed genes.

```
#Extract the ENTREZID of overexpressed genes, which is an argument in the function as a vector

entrez.select.exp <- AnnotationDbi::select(rat2302.db,rownames(select.exp) , c("ENTREZID"))

#Remove duplicates and NA values
entrez.select.exp<-na.omit(entrez.select.exp)



ego <- enrichGO(gene = as.integer(entrez.select.exp$ENTREZID), #a vector of entrez gene id
                keyType = "ENTREZID",
                OrgDb = rat2302.db,
                ont = "BP", #Subontology
                pAdjustMethod = "BH",
                qvalueCutoff = 0.25,
                readable = TRUE)
```

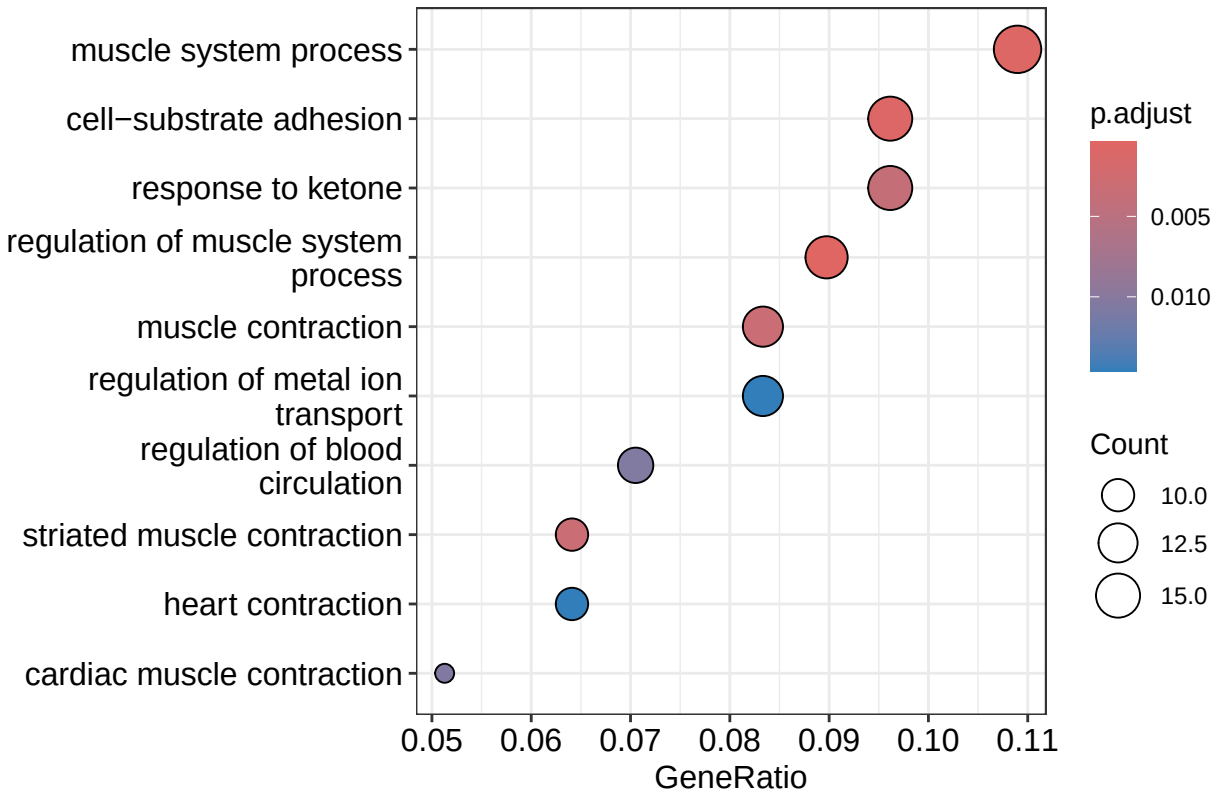
Export the results for access.

```r
# Get the current working directory
directorio <- getwd()

# Create "results" directory if it doesn't exist
if (!file.exists("results")) {
  dir.create("results")
}

# Define the path for the results directory
resultsDir <- file.path(directorio, "results")

# Set working directory
setwd(resultsDir)

# Data.frame of Enrichment Analysis
ego_results <- data.frame(ego)
write.csv(ego_results, "clusterProfiler_ORAresults_UpGO.csv")

setwd(directorio)
```

Let's visualize the enrichment analysis results to see which processes are more overexpressed.

- Dotplot of terms:

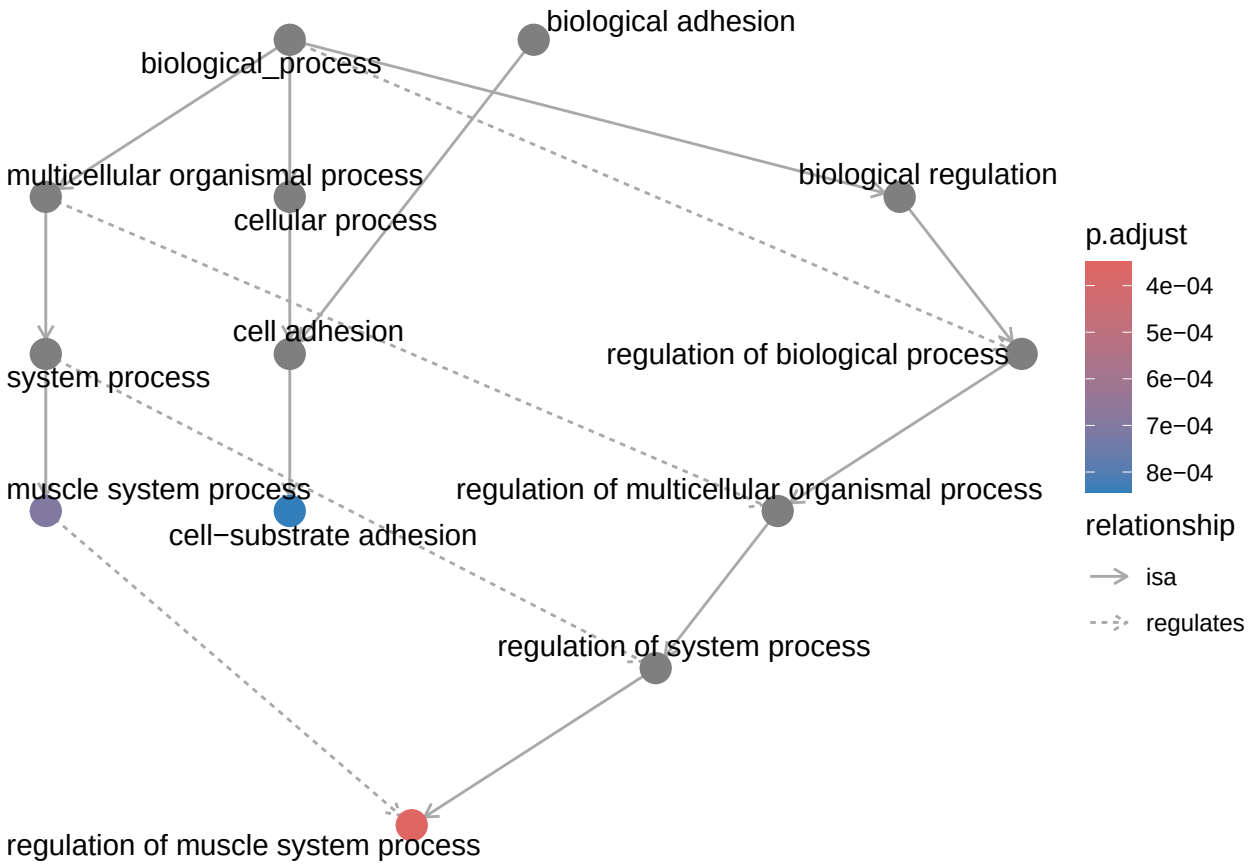
```r
dotplot(ego, showCategory=10)
```



In this Dotplot, we can see enriched biological processes, their significance, and the ratio of differentially expressed genes (X-axis) associated with these processes. In this case, processes related to the muscular system are prominent.

-Visualization of GO terms in hierarchy:

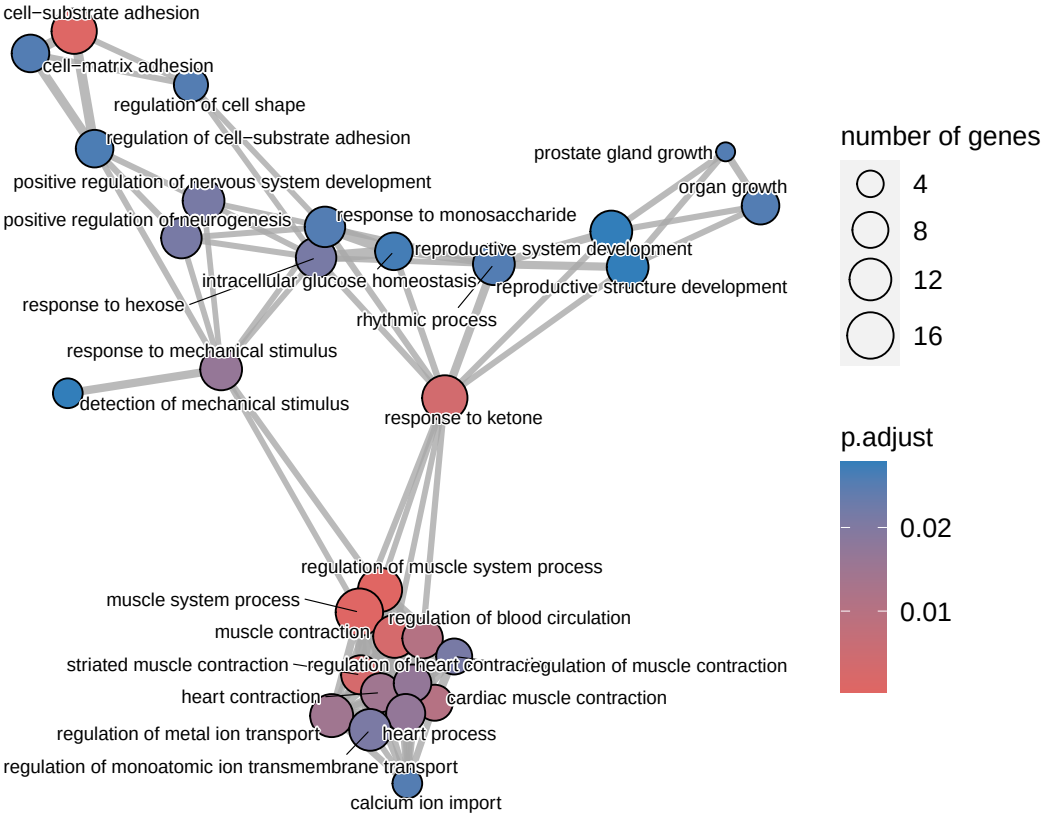
```
goplot(ego, showCategory=3)
```



Here, we explore the hierarchical structure of enriched terms. We focus on higher-level terms such as "Muscle System Process" and also observe "Cell-Substrate Adhesion."

- Enrichment Map

```
## Enrichmap clusters the 50 most significant (by adj.P.Va) GO terms to visualize relationships between
library(enrichplot)
ego_sim <- pairwise_termsim(ego)
emapplot(ego_sim, cex_label_category=0.5)
```



4.4.2 Gene Set Enrichment Analysis (GSEA)

Unlike the previous analysis, GSEA does not focus on the differential expression of genes. This method is used to interpret gene expression data and identify whether predefined sets of genes show statistically significant differences in their expression levels between our two conditions. In summary, we obtain information about biological pathways associated with observed changes in gene expression.

```
#Add ENTREZID to all our genes
entrezIDs <- AnnotationDbi::select(rat2302.db, rownames(Hypo.Control), c("ENTREZID"))


# Add ENTREZID to our genes by merging into the final geneList object
probeid.select.exp<- cbind( PROBEID= rownames(Hypo.Control), Hypo.Control)
geneList  <- merge(probeid.select.exp, entrezIDs, by="PROBEID")

# Remove genes with small logFC
geneList <- geneList[order(abs(geneList$logFC), decreasing=T),]
geneList <- geneList[ !duplicated(geneList$ENTREZ), ]  ### Keep highest

# Reorder for GSEA analysis
geneList <- geneList[order(geneList$logFC, decreasing=T),]
genesVector <- geneList$logFC
names(genesVector) <- geneList$ENTREZ
```

```
gseResulti <- gseKEGG(geneList = genesVector,
                      organism = "rno",
                      keyType = "kegg",
                      exponent = 1,
                      minGSSize = 10,maxGSSize = 500,
                      pvalueCutoff = 0.05,pAdjustMethod = "BH",
                      verbose = TRUE,
                      use_internal_data = FALSE,
                      seed = TRUE,
                      eps=0,
                      by = "fgsea"
                )

# keggResultsList[[i]] <- gseResulti
```

```
library(kableExtra)
gsea.result <- setReadable(gseResulti, OrgDb = rat2302.db, keyType ="ENTREZID" )

gsea.result.df <- as.data.frame(gsea.result)

gsea.result.df
```

	ID	Description	setSize	enrichmentScore	NES
rno05146	rno05146	Amoebiasis	77	0.6116974	1.923742
rno03010	rno03010	Ribosome	157	-0.3893174	-1.783624
rno04080	rno04080	Neuroactive ligand-receptor interaction	299	0.4420567	1.572937
rno04668	rno04668	TNF signaling pathway	93	0.5381232	1.751422
rno04935	rno04935	Growth hormone synthesis, secretion and action	105	0.5115203	1.686221
rno04211	rno04211	Longevity regulating pathway	75	0.5457281	1.717093
rno05322	rno05322	Systemic lupus erythematosus	75	-0.4263346	-1.713925
rno04740	rno04740	Olfactory transduction	59	0.5621447	1.713634
rno04933	rno04933	AGE-RAGE signaling pathway in diabetic complications	96	0.5234721	1.708342
rno04929	rno04929	GnRH secretion	62	0.5501140	1.696133
rno05161	rno05161	Hepatitis B	124	0.4842474	1.624074
rno04915	rno04915	Estrogen signaling pathway	107	0.4898793	1.619404
rno04068	rno04068	FoxO signaling pathway	116	0.4825152	1.611795
rno04510	rno04510	Focal adhesion	164	0.4611803	1.581940
rno05417	rno05417	Lipid and atherosclerosis	170	0.4533547	1.557812
rno04151	rno04151	PI3K-Akt signaling pathway	282	0.4111077	1.458293
rno04917	rno04917	Prolactin signaling pathway	68	0.5399372	1.679196
rno04380	rno04380	Osteoclast differentiation	109	0.5006398	1.658281
rno00983	rno00983	Drug metabolism - other enzymes	70	-0.4186753	-1.651090
rno01521	rno01521	EGFR tyrosine kinase inhibitor resistance	72	0.5235971	1.639775
rno04925	rno04925	Aldosterone synthesis and secretion	85	0.5116581	1.634783
rno04360	rno04360	Axon guidance	133	0.4645262	1.560747
rno04062	rno04062	Chemokine signaling pathway	142	0.4565224	1.544145
rno04020	rno04020	Calcium signaling pathway	229	0.4187631	1.468564
rno04024	rno04024	cAMP signaling pathway	189	0.4214732	1.459131
rno04022	rno04022	cGMP-PKG signaling pathway	149	0.4491546	1.522078
rno04662	rno04662	B cell receptor signaling pathway	59	0.5421706	1.652746
rno04066	rno04066	HIF-1 signaling pathway	101	0.4809799	1.579248
rno04550	rno04550	Signaling pathways regulating pluripotency of stem cells	99	0.4784239	1.567410
rno03020	rno03020	RNA polymerase	25	-0.5778724	-1.796810
rno00480	rno00480	Glutathione metabolism	58	-0.4258564	-1.612709
rno05410	rno05410	Hypertrophic cardiomyopathy	82	0.4933101	1.565170
rno01200	rno01200	Carbon metabolism	97	-0.3662625	-1.546090
rno04261	rno04261	Adrenergic signaling in cardiomyocytes	136	0.4537539	1.531093
rno01522	rno01522	Endocrine resistance	82	0.4902824	1.555564
rno04071	rno04071	Sphingolipid signaling pathway	112	0.4556617	1.517298
rno05165	rno05165	Human papillomavirus infection	246	0.4034729	1.417771
rno00620	rno00620	Pyruvate metabolism	41	-0.4771840	-1.646735
rno05164	rno05164	Influenza A	124	0.4524990	1.517596
rno05204	rno05204	Chemical carcinogenesis - DNA adducts	62	-0.4055925	-1.535411
rno05215	rno05215	Prostate cancer	84	0.4854945	1.549501
rno05162	rno05162	Measles	106	0.4610671	1.523094
rno04660	rno04660	T cell receptor signaling pathway	103	0.4591357	1.511364
rno05135	rno05135	Yersinia infection	113	0.4526346	1.506338
rno04713	rno04713	Circadian entrainment	82	0.4817534	1.528503
rno05206	rno05206	MicroRNAs in cancer	131	0.4420850	1.484822
rno01210	rno01210	2-Oxocarboxylic acid metabolism	27	-0.5266521	-1.668275
rno04928	rno04928	Parathyroid hormone synthesis, secretion and action	88	0.4851240	1.558663
rno04060	rno04060	Cytokine-cytokine receptor interaction	166	0.4142585	1.417779
rno05414	rno05414	Dilated cardiomyopathy	85	0.4778507	1.526766
rno04072	rno04072	Phospholipase D signaling pathway	129	0.4349676	1.457743
rno04911	rno04911	Insulin secretion	75	0.4899771	1.541677
rno00982	rno00982	Drug metabolism - cytochrome P450	50	-0.4222261	-1.536023
rno04666	rno04666	Fc gamma R-mediated phagocytosis	79	0.4813321	1.519149
rno01230	rno01230	Biosynthesis of amino acids	61	-0.3959662	-1.517334

```
print(kable(gsea.result.df[,c("Description","setSize","NES","p.adjust")])%>% scroll_box(height = "500px
```

```
##
## \begin{tabular}{l|l|r|r|r}
## \hline
##   & Description & setSize & NES & p.adjust\\
## \hline
## rno05146 & Amoebiasis & 77 & 1.923742 & 0.0005161\\
## \hline
## rno03010 & Ribosome & 157 & -1.783624 & 0.0005161\\
## \hline
## rno04080 & Neuroactive ligand-receptor interaction & 299 & 1.572937 & 0.0017098\\
## \hline
## rno04668 & TNF signaling pathway & 93 & 1.751422 & 0.0109243\\
## \hline
## rno04935 & Growth hormone synthesis, secretion and action & 105 & 1.686221 & 0.0109243\\
## \hline
## rno04211 & Longevity regulating pathway & 75 & 1.717093 & 0.0151331\\
## \hline
## rno05322 & Systemic lupus erythematosus & 75 & -1.713925 & 0.0151331\\
## \hline
## rno04740 & Olfactory transduction & 59 & 1.713634 & 0.0151331\\
## \hline
## rno04933 & AGE-RAGE signaling pathway in diabetic complications & 96 & 1.708342 & 0.0151331\\
## \hline
## rno04929 & GnRH secretion & 62 & 1.696133 & 0.0151331\\
## \hline
## rno05161 & Hepatitis B & 124 & 1.624074 & 0.0151331\\
## \hline
## rno04915 & Estrogen signaling pathway & 107 & 1.619404 & 0.0151331\\
## \hline
## rno04068 & FoxO signaling pathway & 116 & 1.611795 & 0.0151331\\
## \hline
## rno04510 & Focal adhesion & 164 & 1.581940 & 0.0151331\\
## \hline
## rno05417 & Lipid and atherosclerosis & 170 & 1.557812 & 0.0151331\\
## \hline
## rno04151 & PI3K-Akt signaling pathway & 282 & 1.458293 & 0.0151331\\
## \hline
## rno04917 & Prolactin signaling pathway & 68 & 1.679196 & 0.0168357\\
## \hline
## rno04380 & Osteoclast differentiation & 109 & 1.658281 & 0.0168357\\
## \hline
## rno00983 & Drug metabolism - other enzymes & 70 & -1.651090 & 0.0168357\\
## \hline
## rno01521 & EGFR tyrosine kinase inhibitor resistance & 72 & 1.639775 & 0.0198532\\
## \hline
## rno04925 & Aldosterone synthesis and secretion & 85 & 1.634783 & 0.0198532\\
## \hline
## rno04360 & Axon guidance & 133 & 1.560747 & 0.0198532\\
## \hline
## rno04062 & Chemokine signaling pathway & 142 & 1.544145 & 0.0198532\\
## \hline
```

```
## rno04020 & Calcium signaling pathway & 229 & 1.468564 & 0.0198532\\
## \hline
## rno04024 & cAMP signaling pathway & 189 & 1.459131 & 0.0198532\\
## \hline
## rno04022 & cGMP-PKG signaling pathway & 149 & 1.522078 & 0.0225619\\
## \hline
## rno04662 & B cell receptor signaling pathway & 59 & 1.652746 & 0.0234449\\
## \hline
## rno04066 & HIF-1 signaling pathway & 101 & 1.579248 & 0.0254315\\
## \hline
## rno04550 & Signaling pathways regulating pluripotency of stem cells & 99 & 1.567410 & 0.0278589\\
## \hline
## rno03020 & RNA polymerase & 25 & -1.796810 & 0.0307696\\
## \hline
## rno00480 & Glutathione metabolism & 58 & -1.612709 & 0.0307696\\
## \hline
## rno05410 & Hypertrophic cardiomyopathy & 82 & 1.565170 & 0.0307696\\
## \hline
## rno01200 & Carbon metabolism & 97 & -1.546090 & 0.0307696\\
## \hline
## rno04261 & Adrenergic signaling in cardiomyocytes & 136 & 1.531093 & 0.0307696\\
## \hline
## rno01522 & Endocrine resistance & 82 & 1.555564 & 0.0334317\\
## \hline
## rno04071 & Sphingolipid signaling pathway & 112 & 1.517298 & 0.0334317\\
## \hline
## rno05165 & Human papillomavirus infection & 246 & 1.417771 & 0.0334317\\
## \hline
## rno00620 & Pyruvate metabolism & 41 & -1.646735 & 0.0339340\\
## \hline
## rno05164 & Influenza A & 124 & 1.517596 & 0.0339340\\
## \hline
## rno05204 & Chemical carcinogenesis - DNA adducts & 62 & -1.535411 & 0.0365590\\
## \hline
## rno05215 & Prostate cancer & 84 & 1.549501 & 0.0366230\\
## \hline
## rno05162 & Measles & 106 & 1.523094 & 0.0366230\\
## \hline
## rno04660 & T cell receptor signaling pathway & 103 & 1.511364 & 0.0366230\\
## \hline
## rno05135 & Yersinia infection & 113 & 1.506338 & 0.0366230\\
## \hline
## rno04713 & Circadian entrainment & 82 & 1.528503 & 0.0379833\\
## \hline
## rno05206 & MicroRNAs in cancer & 131 & 1.484822 & 0.0398948\\
## \hline
## rno01210 & 2-Oxocarboxylic acid metabolism & 27 & -1.668275 & 0.0424789\\
## \hline
## rno04928 & Parathyroid hormone synthesis, secretion and action & 88 & 1.558663 & 0.0424789\\
## \hline
## rno04060 & Cytokine-cytokine receptor interaction & 166 & 1.417779 & 0.0486419\\
## \hline
## rno05414 & Dilated cardiomyopathy & 85 & 1.526766 & 0.0492221\\
## \hline
```

```
## rno04072 & Phospholipase D signaling pathway & 129 & 1.457743 & 0.0492221\\
## \hline
## rno04911 & Insulin secretion & 75 & 1.541677 & 0.0493265\\
## \hline
## rno00982 & Drug metabolism - cytochrome P450 & 50 & -1.536023 & 0.0493265\\
## \hline
## rno04666 & Fc gamma R-mediated phagocytosis & 79 & 1.519149 & 0.0493265\\
## \hline
## rno01230 & Biosynthesis of amino acids & 61 & -1.517334 & 0.0493265\\
## \hline
## \end{tabular}
```

The main difference between GSEA and ORA is that the latter compares the overlap between two gene sets (differential genes of control and hypotensive mice), using statistical tests to determine significance. GSEA, on the other hand, classifies all genes in the dataset from which we started and then checks if genes in a gene set (associated with a process) are enriched. While ORA focuses on the overlap between sets of genes that have been verified as differential, GSEA considers the entire distribution of changes in gene expression. In summary, GSEA considers the complete distribution of changes in gene expression in a dataset.

5. Discussion

We are dealing with a study involving two groups of mice: control and those with induced hypotension. Our analyses of differential gene expression, supported by enrichment analysis results, suggest that certain biological processes such as cell adhesion or the muscular system show differences. The association of differentially expressed genes with processes of the muscular system and cell-substrate adhesion could have interesting implications in the context of hypotension according to our results.

Muscle function, especially contractility and response to blood pressure, is essential for blood pressure regulation. Therefore, it is logical for hypotensive mice to have higher expression of these genes because they need an adaptation of the muscular system to changes in blood pressure. Furthermore, cell-substrate adhesion is crucial for the integrity and function of tissues, such as muscles. Therefore, these changes could reflect cellular adaptations to hypotension by hypotensive mice.

Additionally, the ketone response systems are also active in hypotensive individuals. This could be associated with the use of ketone bodies as an energy source. This finding suggests a metabolic reprogramming in response to hypotension. Changes in energy metabolism could be an adaptive strategy to maintain homeostasis under hypotensive conditions.

6. Conclusion

In summary, ORA results suggest that induced hypotension is associated with changes in gene expression affecting muscle function, cell adhesion, and metabolic response.