

# Class 14 RNA-Seq Analysis Mini Project

Ava Limtiaco (PID: A18007672)

## Table of contents

**Remove zero count genes** 6

**Run DESeq analysis** 7

Here we will perform a complete RNASeq analysis from counts to pathways and biological interpretation.

Trapnell C, Hendrickson DG, Sauvageau M, Goff L et al. "Differential analysis of gene regulation at transcript resolution with RNA-seq". Nat Biotechnol 2013 Jan;31(1):46-53. PMID: 23222703

The authors report on differential analysis of lung fibroblasts in response to loss of the developmental transcription factor HOXA1.

#Required Packages

```
library(DESeq2)
```

```
Loading required package: S4Vectors
```

```
Loading required package: stats4
```

```
Loading required package: BiocGenerics
```

```
Loading required package: generics
```

```
Attaching package: 'generics'
```

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

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

Attaching package: 'BiocGenerics'

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

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

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

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

Attaching package: 'S4Vectors'

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

```
findMatches
```

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

```
expand.grid, I, unname
```

Loading required package: IRanges

Loading required package: GenomicRanges

Loading required package: GenomeInfoDb

Loading required package: SummarizedExperiment

Loading required package: MatrixGenerics

Loading required package: matrixStats

Attaching package: 'MatrixGenerics'

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

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

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

```
library(AnnotationDbi)
library(org.Hs.eg.db)
```

```
library(pathview)
```

```
#####
Pathview is an open source software package distributed under GNU General
Public License version 3 (GPLv3). Details of GPLv3 is available at
http://www.gnu.org/licenses/gpl-3.0.html. Particullary, users are required to
formally cite the original Pathview paper (not just mention it) in publications
or products. For details, do citation("pathview") within R.
```

The pathview downloads and uses KEGG data. Non-academic uses may require a KEGG license agreement (details at <http://www.kegg.jp/kegg/legal.html>).

```
#####
```

```
library(gage)
```

```
library(gageData)
```

#Data Import

```
colData <- read.csv("GSE37704_metadata.csv", row.names=1)
countData <- read.csv("GSE37704_featurecounts.csv", row.names=1)
```

```
head(colData)
```

```
              condition
SRR493366 control_sirna
SRR493367 control_sirna
```

```
SRR493368 control_sirna
SRR493369      hoxa1_kd
SRR493370      hoxa1_kd
SRR493371      hoxa1_kd
```

```
head(countData)
```

	length	SRR493366	SRR493367	SRR493368	SRR493369	SRR493370
ENSG00000186092	918	0	0	0	0	0
ENSG00000279928	718	0	0	0	0	0
ENSG00000279457	1982	23	28	29	29	28
ENSG00000278566	939	0	0	0	0	0
ENSG00000273547	939	0	0	0	0	0
ENSG00000187634	3214	124	123	205	207	212
	SRR493371					
ENSG00000186092	0					
ENSG00000279928	0					
ENSG00000279457	46					
ENSG00000278566	0					
ENSG00000273547	0					
ENSG00000187634	258					

#Tidy up Check the correspondance of colData rows and countData columns.

```
rownames(colData)
```

```
[1] "SRR493366" "SRR493367" "SRR493368" "SRR493369" "SRR493370" "SRR493371"
```

```
colnames(countData)
```

```
[1] "length"      "SRR493366" "SRR493367" "SRR493368" "SRR493369" "SRR493370"
[7] "SRR493371"
```

```
counts <- countData[, -1]
```

```
head(counts)
```

	SRR493366	SRR493367	SRR493368	SRR493369	SRR493370	SRR493371
ENSG00000186092	0	0	0	0	0	0
ENSG00000279928	0	0	0	0	0	0
ENSG00000279457	23	28	29	29	28	46
ENSG00000278566	0	0	0	0	0	0
ENSG00000273547	0	0	0	0	0	0
ENSG00000187634	124	123	205	207	212	258

```
all(rownames(colData) == colnames(counts))
```

```
[1] TRUE
```

```
all(c(T,F,T))
```

```
[1] FALSE
```

## Remove zero count genes

We will have rows in 'counts' for genes that we can not say anything about because they have zero expression in the particular tissue we are looking at.

```
head(counts)
```

	SRR493366	SRR493367	SRR493368	SRR493369	SRR493370	SRR493371
ENSG00000186092	0	0	0	0	0	0
ENSG00000279928	0	0	0	0	0	0
ENSG00000279457	23	28	29	29	28	46
ENSG00000278566	0	0	0	0	0	0
ENSG00000273547	0	0	0	0	0	0
ENSG00000187634	124	123	205	207	212	258

If the 'rowSums()' is zero then a given gene (row) has no count data and we should exclude these genes from further consideration.

```
to.keep <- rowSums(counts) != 0
cleancounts <- counts[to.keep, ]
```

Q1. How many genes do we have left?

```
nrow(cleancounts)
```

```
[1] 15975
```

```
#Setup DESeq object for analysis
```

```
dds <- DESeqDataSetFromMatrix(countData = cleancounts, colData = colData, ~condition)
```

```
Warning in DESeqDataSet(se, design = design, ignoreRank): some variables in  
design formula are characters, converting to factors
```

## Run DESeq analysis

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

```
estimating size factors
```

```
estimating dispersions
```

```
gene-wise dispersion estimates
```

```
mean-dispersion relationship
```

```
final dispersion estimates
```

```
fitting model and testing
```

```
#Extract results
```

```
res <- results(dds)  
head(res)
```

log2 fold change (MLE): condition hoxa1 kd vs control sirna

Wald test p-value: condition hoxa1 kd vs control sirna

DataFrame with 6 rows and 6 columns

	baseMean	log2FoldChange	lfcSE	stat	pvalue
	<numeric>	<numeric>	<numeric>	<numeric>	<numeric>
ENSG00000279457	29.9136	0.1792571	0.3248216	0.551863	5.81042e-01
ENSG00000187634	183.2296	0.4264571	0.1402658	3.040350	2.36304e-03
ENSG00000188976	1651.1881	-0.6927205	0.0548465	-12.630158	1.43989e-36
ENSG00000187961	209.6379	0.7297556	0.1318599	5.534326	3.12428e-08
ENSG00000187583	47.2551	0.0405765	0.2718928	0.149237	8.81366e-01
ENSG00000187642	11.9798	0.5428105	0.5215599	1.040744	2.97994e-01
	padj				
	<numeric>				
ENSG00000279457	6.86555e-01				
ENSG00000187634	5.15718e-03				
ENSG00000188976	1.76549e-35				
ENSG00000187961	1.13413e-07				
ENSG00000187583	9.19031e-01				
ENSG00000187642	4.03379e-01				

#Add Gene annotation

```
library(AnnotationDbi)
library(org.Hs.eg.db)

res$symbol <- mapIds(org.Hs.eg.db,
                     keys = rownames(res),
                     column = "SYMBOL",
                     keytype = "ENSEMBL",
                     multiVals = "first")
```

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

```
res$entrez <- mapIds(org.Hs.eg.db,
                     keys = rownames(res),
                     column = "ENTREZID",
                     keytype = "ENSEMBL",
                     multiVals = "first")
```

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



```
res$name <- mapIds(org.Hs.eg.db,
                  keys = rownames(res),
                  column = "GENENAME",
                  keytype = "ENSEMBL",
                  multiVals = "first")
```

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

#Save my results to a CSV file

```
write.csv(res, file="results.csv")
```

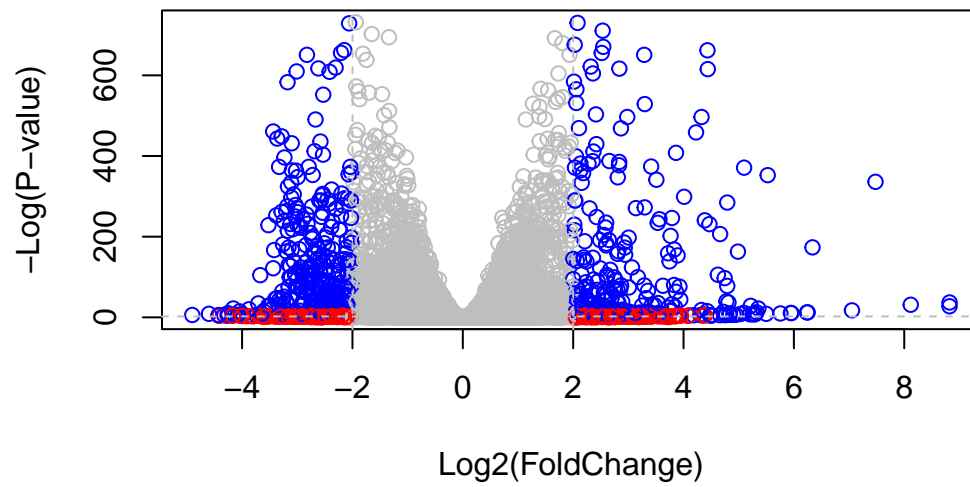
#Result visualization

```
mycols <- rep("gray", nrow(res))
mycols[ abs(res$log2FoldChange) > 2 ] <- "red"

inds <- (res$padj < 0.01) & (abs(res$log2FoldChange) > 2 )
mycols[ inds ] <- "blue"

plot( res$log2FoldChange, -log(res$padj),
      col=mycols, ylab="-Log(P-value)", xlab="Log2(FoldChange)" )

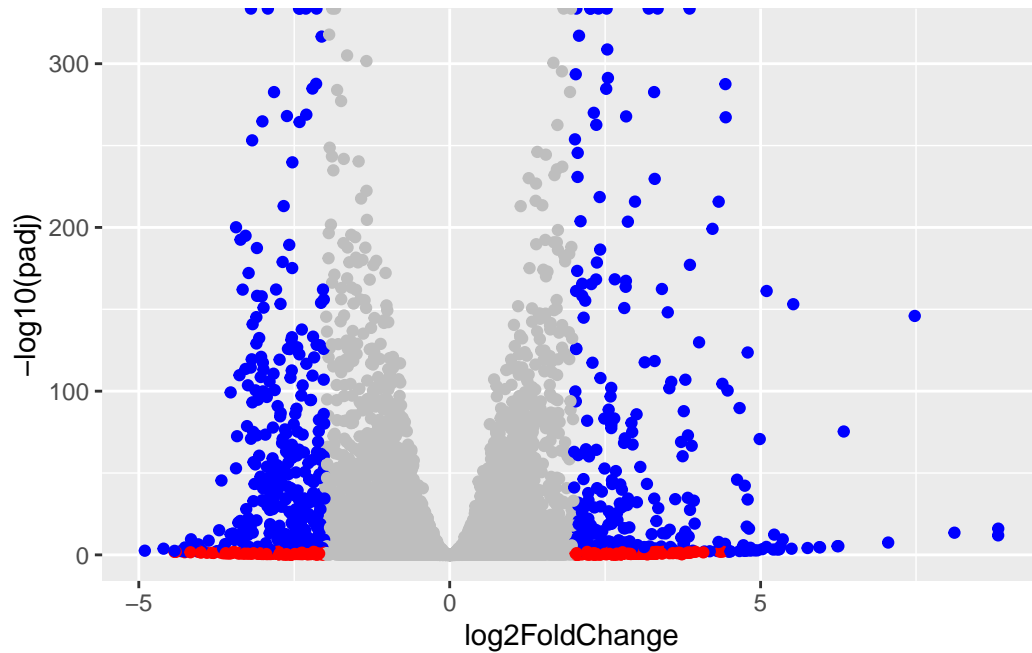
abline(v=c(-2,2), col="gray", lty=2)
abline(h=-log(0.1), col="gray", lty=2)
```



```
library(ggplot2)
library(ggrepel)

ggplot(res)+
  aes(x= log2FoldChange,y= -log10(padj), label= res$symbol)+
  geom_point(col=mycols)
```

Warning: Removed 1237 rows containing missing values or values outside the scale range (`geom\_point()`).



#Pathway analysis

```
data(kegg.sets.hs)
data(sigmet.idx.hs)
```

```
kegg.sets.hs=kegg.sets.hs[sigmet.idx.hs]
```

```
foldchanges=res$log2FoldChange
names(foldchanges) = res$entrez
head(foldchanges)
```

```
<NA>      148398      26155      339451      84069      84808
0.17925708 0.42645712 -0.69272046 0.72975561 0.04057653 0.54281049
```

```
keggres = gage(foldchanges, gsets=kegg.sets.hs)
```

```
attributes(keggres)
```

```
$names
[1] "greater" "less"    "stats"
```

```
head(keggres$less)
```

	p.geomean	stat.mean	p.val
hsa04110 Cell cycle	8.995727e-06	-4.378644	8.995727e-06
hsa03030 DNA replication	9.424076e-05	-3.951803	9.424076e-05
hsa03013 RNA transport	1.246882e-03	-3.059466	1.246882e-03
hsa03440 Homologous recombination	3.066756e-03	-2.852899	3.066756e-03
hsa04114 Oocyte meiosis	3.784520e-03	-2.698128	3.784520e-03
hsa00010 Glycolysis / Gluconeogenesis	8.961413e-03	-2.405398	8.961413e-03

	q.val	set.size	exp1
hsa04110 Cell cycle	0.001448312	121	8.995727e-06
hsa03030 DNA replication	0.007586381	36	9.424076e-05
hsa03013 RNA transport	0.066915974	144	1.246882e-03
hsa03440 Homologous recombination	0.121861535	28	3.066756e-03
hsa04114 Oocyte meiosis	0.121861535	102	3.784520e-03
hsa00010 Glycolysis / Gluconeogenesis	0.212222694	53	8.961413e-03

```
pathview(gene.data=foldchanges, pathway.id=keggres, species="hsa")
```

Info: Downloading xml files for hsac(0.00282277553699587, 0.00520206953646473, 0.007255098941, 0.115053569800308, 0.126460619211945, 0.132241873786472, 0.136920066131284, 0.13941320027347, 0.276993488830076, 0.28945232833203, 0.296088852305941, 0.300930811918732, 0.30269230994547, 0.519356159391074, 0.53081720433347, 0.531672041345871, 0.534884921031663, 0.537651485486221, 0.752246757197311, 0.752505053638346, 0.756524162296446, 0.758563851969868, 0.76149842342670, 0.894247160054425, 0.898264909409692, 0.909595445308512, 0.911173947840552, 0.92069143392553, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, 2.83336226490419, 2.58567325, 1.2479739087179, 1.2048877401891, 1.16738452466142, 1.11989152828197, 1.09939992687849, 1.08, 0.593947853026533, 0.597925393567583, 0.561096211945878, 0.536357335786319, 0.52288602482831, -0.0446533584931199, -0.0488976656749805, -0.0775366737169223, -0.080045560497172, -0.087745, -0.569471707020036, -0.597899003198084, -0.631868022146645, -0.688000149605839, -0.683682769, -1.12964145549338, -1.19738360661794, -1.18910601981577, -1.25652891970231, -1.2520941843637, -2.85289859438099, -3.05946647947393, -3.95180301718443, -4.3786443856141, NaN, NaN, NaN, NaN, 0.0704957890100001, 0.0705184151657714, 0.0722500516314278, 0.0803451561918264, 0.0825499055, 0.231260465450316, 0.232710177642139, 0.238667908557651, 0.240527105514867, 0.24084282599715, 0.38706972218425, 0.403734774607358, 0.410628638768613, 0.421624072068726, 0.432154972053367, 0.654228612083812, 0.666251293138561, 0.666926273728945, 0.667386746756933, 0.70182758714257, 0.84374203170927, 0.844191597789139, 0.846767241542369, 0.851493741402169, 0.870034512062492, 0.989454772961867, 0.990528974934311, 0.991038587218259, 0.99621548027014, 0.996933243692408, 0.553917062507613, 0.553917062507613, 0.553917062507613, 0.553917062507613, 0.553917062507613, 0.775513899710837, 0.775513899710837, 0.775513899710837, 0.775513899710837, 0.775513899710837, 0.853674318789922, 0.853674318789922, 0.853674318789922, 0.87839592853763, 0.881482811223289

0.999991004273293, 0.999991004273293, 0.999991004273293, 0.999991004273293, 0.999991004273293,  
0.999991004273293, 0.999991004273293, 0.999991004273293, 0.999991004273293, 0.999991004273293,  
0.999991004273293, 0.999991004273293, 0.999991004273293, 0.999991004273293, 0.999991004273293,  
127, 33, 88, 123, 83, 87, 54, 64, 15, 38, 36, 25, 21, 28, 56, 28, 85, 160, 14, 81, 20, 18, 4,  
0.0072550989425664, 0.0101073919374798, 0.0187472525444945, 0.0193997662548, 0.0197329944159,  
0.132241873786472, 0.136920066131284, 0.139413200273477, 0.143184308480604, 0.14515183518277,  
0.296088852305941, 0.300930811918732, 0.30269230994547, 0.308562955803806, 0.311095063984598,  
0.531672041345871, 0.534884921031663, 0.537651485486221, 0.549864831052667, 0.56513084211139,  
0.756524162296446, 0.758563851969868, 0.761498423426708, 0.770700896956808, 0.77600524032493,  
0.909595445308512, 0.911173947840552, 0.920691433925531, 0.925956183278906, 0.93559571695637,  
NA, NA, NA, NA, NA, NA, NA, NA), 1/1 pathways..

Warning in download.file(xml.url, xml.target, quiet = T): expanded path length 16392 would be  
./hsac(0.00282277553699587, 0.00520206953646473, 0.0072550989425664, 0.0101073919374798, 0.0115053569800308,  
0.115053569800308, 0.126460619211945, 0.132241873786472, 0.136920066131284, 0.139413200273477

Warning in download.file(xml.url, xml.target, quiet = T): URL https://rest.kegg.jp/get/hsac(0.00282277553699587,  
0.115053569800308, 0.126460619211945, 0.132241873786472, 0.136920066131284, 0.139413200273477

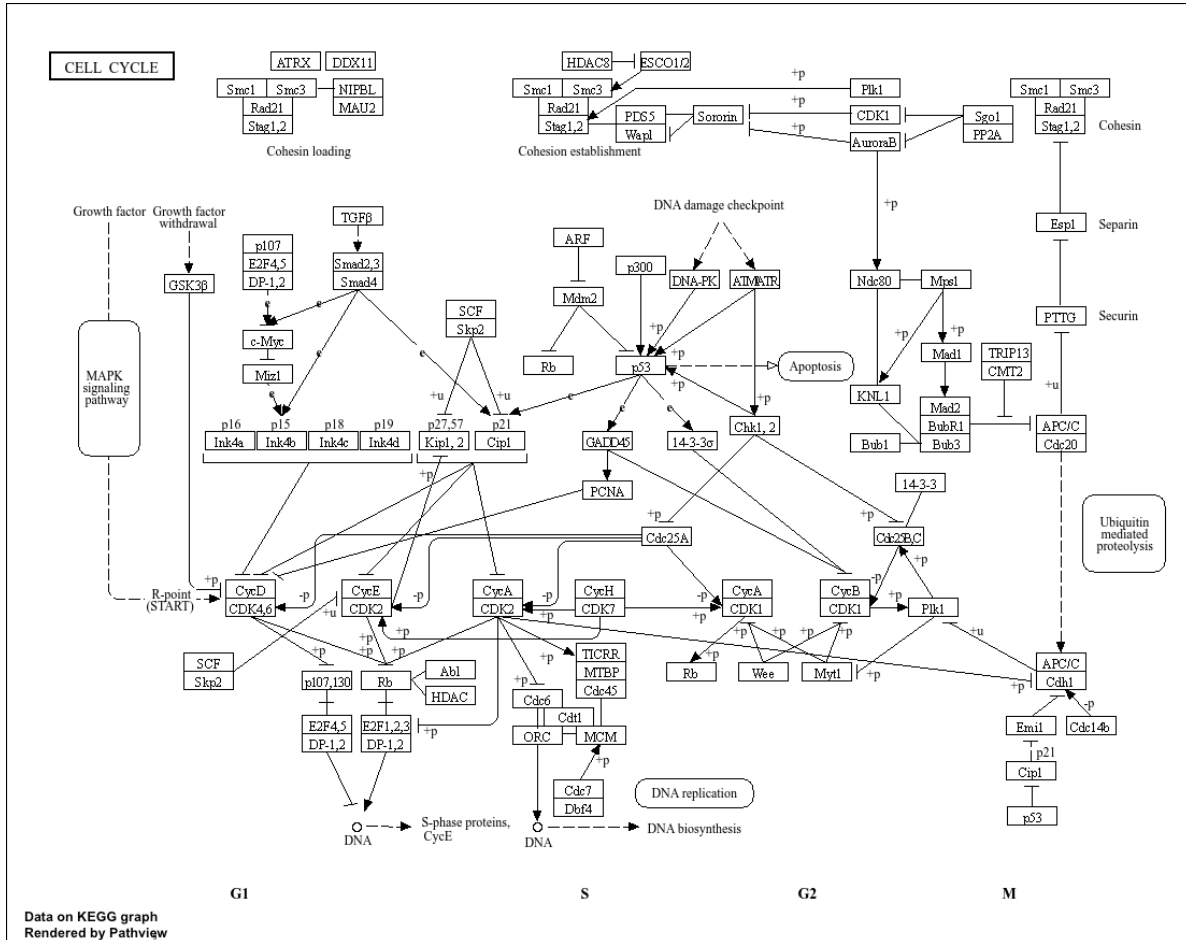
Warning in download.file(xml.url, xml.target, quiet = T): download had nonzero  
exit status

Info: Downloading png files for hsac(0.00282277553699587, 0.00520206953646473, 0.0072550989425664,  
0.115053569800308, 0.126460619211945, 0.132241873786472, 0.136920066131284, 0.139413200273477,  
0.276993488830076, 0.28945232833203, 0.296088852305941, 0.300930811918732, 0.30269230994547,  
0.519356159391074, 0.53081720433347, 0.531672041345871, 0.534884921031663, 0.537651485486221,  
0.752246757197311, 0.752505053638346, 0.756524162296446, 0.758563851969868, 0.761498423426708,  
0.894247160054425, 0.898264909409692, 0.909595445308512, 0.911173947840552, 0.920691433925531,  
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, 2.83336226490419, 2.58567325,  
1.2479739087179, 1.2048877401891, 1.16738452466142, 1.11989152828197, 1.09939992687849, 1.08,  
0.593947853026533, 0.597925393567583, 0.561096211945878, 0.536357335786319, 0.52288602482831,  
-0.0446533584931199, -0.0488976656749805, -0.0775366737169223, -0.080045560497172, -0.087745,  
-0.569471707020036, -0.597899003198084, -0.631868022146645, -0.688000149605839, -0.683682769,  
-1.12964145549338, -1.19738360661794, -1.18910601981577, -1.25652891970231, -1.2520941843637,  
-2.85289859438099, -3.05946647947393, -3.95180301718443, -4.3786443856141, NaN, NaN, NaN, NaN,  
0.0704957890100001, 0.0705184151657714, 0.0722500516314278, 0.0803451561918264, 0.0825499055,  
0.231260465450316, 0.232710177642139, 0.238667908557651, 0.240527105514867, 0.24084282599715,  
0.38706972218425, 0.403734774607358, 0.410628638768613, 0.421624072068726, 0.432154972053367,  
0.654228612083812, 0.666251293138561, 0.666926273728945, 0.667386746756933, 0.70182758714257,  
0.84374203170927, 0.844191597789139, 0.846767241542369, 0.851493741402169, 0.870034512062492,  
0.989454772961867, 0.990528974934311, 0.991038587218259, 0.99621548027014, 0.996933243692408

0.553917062507613, 0.553917062507613, 0.553917062507613, 0.553917062507613, 0.553917062507613,  
0.775513899710837, 0.775513899710837, 0.775513899710837, 0.775513899710837, 0.775513899710837,  
0.853674318789922, 0.853674318789922, 0.853674318789922, 0.87839592853763, 0.881482811223289,  
0.999991004273293, 0.999991004273293, 0.999991004273293, 0.999991004273293, 0.999991004273293,  
0.999991004273293, 0.999991004273293, 0.999991004273293, 0.999991004273293, 0.999991004273293,  
0.999991004273293, 0.999991004273293, 0.999991004273293, 0.999991004273293, 0.999991004273293,  
127, 33, 88, 123, 83, 87, 54, 64, 15, 38, 36, 25, 21, 28, 56, 28, 85, 160, 14, 81, 20, 18, 4,  
0.0072550989425664, 0.0101073919374798, 0.0187472525444945, 0.0193997662548, 0.0197329944159,  
0.132241873786472, 0.136920066131284, 0.139413200273477, 0.143184308480604, 0.14515183518277,  
0.296088852305941, 0.300930811918732, 0.30269230994547, 0.308562955803806, 0.311095063984598,  
0.531672041345871, 0.534884921031663, 0.537651485486221, 0.549864831052667, 0.56513084211139,  
0.756524162296446, 0.758563851969868, 0.761498423426708, 0.770700896956808, 0.77600524032493,  
0.909595445308512, 0.911173947840552, 0.920691433925531, 0.925956183278906, 0.93559571695637,  
NA, NA, NA, NA, NA, NA, NA, NA, NA), 1/1 pathways..

Warning: Failed to download KEGG xml/png files, hsac(0.00282277553699587, 0.0052020695364647,  
0.115053569800308, 0.126460619211945, 0.132241873786472, 0.136920066131284, 0.139413200273477,  
0.276993488830076, 0.28945232833203, 0.296088852305941, 0.300930811918732, 0.30269230994547,  
0.519356159391074, 0.53081720433347, 0.531672041345871, 0.534884921031663, 0.537651485486221,  
0.752246757197311, 0.752505053638346, 0.756524162296446, 0.758563851969868, 0.761498423426708,  
0.894247160054425, 0.898264909409692, 0.909595445308512, 0.911173947840552, 0.920691433925531,  
NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, 2.83336226490419, 2.58567325,  
1.2479739087179, 1.2048877401891, 1.16738452466142, 1.11989152828197, 1.09939992687849, 1.08,  
0.593947853026533, 0.597925393567583, 0.561096211945878, 0.536357335786319, 0.52288602482831,  
-0.0446533584931199, -0.0488976656749805, -0.0775366737169223, -0.080045560497172, -0.087745,  
-0.569471707020036, -0.597899003198084, -0.631868022146645, -0.688000149605839, -0.683682769,  
-1.12964145549338, -1.19738360661794, -1.18910601981577, -1.25652891970231, -1.2520941843637,  
-2.85289859438099, -3.05946647947393, -3.95180301718443, -4.3786443856141, NaN, NaN, NaN, NaN,  
0.0704957890100001, 0.0705184151657714, 0.0722500516314278, 0.0803451561918264, 0.0825499055,  
0.231260465450316, 0.232710177642139, 0.238667908557651, 0.240527105514867, 0.24084282599715,  
0.38706972218425, 0.403734774607358, 0.410628638768613, 0.421624072068726, 0.432154972053367,  
0.654228612083812, 0.666251293138561, 0.666926273728945, 0.667386746756933, 0.70182758714257,  
0.84374203170927, 0.844191597789139, 0.846767241542369, 0.851493741402169, 0.870034512062492,  
0.989454772961867, 0.990528974934311, 0.991038587218259, 0.99621548027014, 0.996933243692408,  
0.553917062507613, 0.553917062507613, 0.553917062507613, 0.553917062507613, 0.553917062507613,  
0.775513899710837, 0.775513899710837, 0.775513899710837, 0.775513899710837, 0.775513899710837,  
0.853674318789922, 0.853674318789922, 0.853674318789922, 0.87839592853763, 0.881482811223289,  
0.999991004273293, 0.999991004273293, 0.999991004273293, 0.999991004273293, 0.999991004273293,  
0.999991004273293, 0.999991004273293, 0.999991004273293, 0.999991004273293, 0.999991004273293,  
0.999991004273293, 0.999991004273293, 0.999991004273293, 0.999991004273293, 0.999991004273293,  
127, 33, 88, 123, 83, 87, 54, 64, 15, 38, 36, 25, 21, 28, 56, 28, 85, 160, 14, 81, 20, 18, 4,  
0.0072550989425664, 0.0101073919374798, 0.0187472525444945, 0.0193997662548, 0.0197329944159,

0.132241873786472, 0.136920066131284, 0.139413200273477, 0.143184308480604, 0.14515183518277  
0.296088852305941, 0.300930811918732, 0.30269230994547, 0.308562955803806, 0.311095063984598  
0.531672041345871, 0.534884921031663, 0.537651485486221, 0.549864831052667, 0.56513084211139  
0.756524162296446, 0.758563851969868, 0.761498423426708, 0.770700896956808, 0.77600524032493  
0.909595445308512, 0.911173947840552, 0.920691433925531, 0.925956183278906, 0.93559571695637  
NA, NA, NA, NA, NA, NA, NA, NA, NA) skipped!



#Gene ontology

```
data(go.sets.hs)
```

```
data(go.subs.hs)
```

```
# Focus on Biological Process subset of GO
```

```
gobpsets = go.sets.hs[go.subs.hs$BP]
```

```
gobpres = gage(foldchanges, gsets=gobpsets)
```

```
head(gobpres$less, 5)
```

	p.geomean	stat.mean	p.val
G0:0048285 organelle fission	1.536227e-15	-8.063910	1.536227e-15
G0:0000280 nuclear division	4.286961e-15	-7.939217	4.286961e-15
G0:0007067 mitosis	4.286961e-15	-7.939217	4.286961e-15
G0:0000087 M phase of mitotic cell cycle	1.169934e-14	-7.797496	1.169934e-14
G0:0007059 chromosome segregation	2.028624e-11	-6.878340	2.028624e-11

	q.val	set.size	expl
G0:0048285 organelle fission	5.841698e-12	376	1.536227e-15
G0:0000280 nuclear division	5.841698e-12	352	4.286961e-15
G0:0007067 mitosis	5.841698e-12	352	4.286961e-15
G0:0000087 M phase of mitotic cell cycle	1.195672e-11	362	1.169934e-14
G0:0007059 chromosome segregation	1.658603e-08	142	2.028624e-11

##Reactome analysis online

We need to make a little file of our significant genes that we can upload to the reactome webpage.

```
sig_genes <- res[res$padj <= 0.05 & !is.na(res$padj), "symbol"]
print(paste("Total number of significant genes:", length(sig_genes)))
```

```
[1] "Total number of significant genes: 8147"
```

```
sig_genes[6]
```

```
ENSG00000188157
  "AGRN"
```

```
write.table(sig_genes, file="significant_genes.txt", row.names=FALSE, col.names=FALSE, quote=
```

Then to perform pathway analysis online go to Reactome website (<https://reactome.org/PathwayBrowser/#TOC>)



