Class16: RNASeq Mini Project

Yuhan Zhang (PID: A13829264)

11/19/2021

```
library(DESeq2)
```

1. Data Import

Load data:

```
metaFile <- "GSE37704_metadata.csv"
countFile <- "GSE37704_featurecounts.csv"
```

```
colData <- read.csv(metaFile, row.names = 1)
head(colData)
```

```
##                condition
## SRR493366 control_sirna
## SRR493367 control_sirna
## SRR493368 control_sirna
## SRR493369      hoxa1_kd
## SRR493370      hoxa1_kd
## SRR493371      hoxa1_kd
```

```
countData <- read.csv(countFile, row.names = 1)
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
```

We need to remove the first column (i.e. `countData$length`) to match with metadata:

```
countData <- as.matrix(countData[, -1])
head(countData)
```

```
##                 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
```

We also need to remove entries that has no reading (0 across all columns)

```
row.rm = rowSums(countData) != 0
countData <- countData[row.rm,]
head(countData)
```

```
##                 SRR493366 SRR493367 SRR493368 SRR493369 SRR493370 SRR493371
## ENSG00000279457        23        28        29        29        28        46
## ENSG00000187634       124       123       205       207       212       258
## ENSG00000188976      1637      1831      2383      1226      1326      1504
## ENSG00000187961       120       153       180       236       255       357
## ENSG00000187583        24        48        65        44        48        64
## ENSG00000187642         4         9        16        14        16        16
```

```
nrow(countData)
```

```
## [1] 15975
```

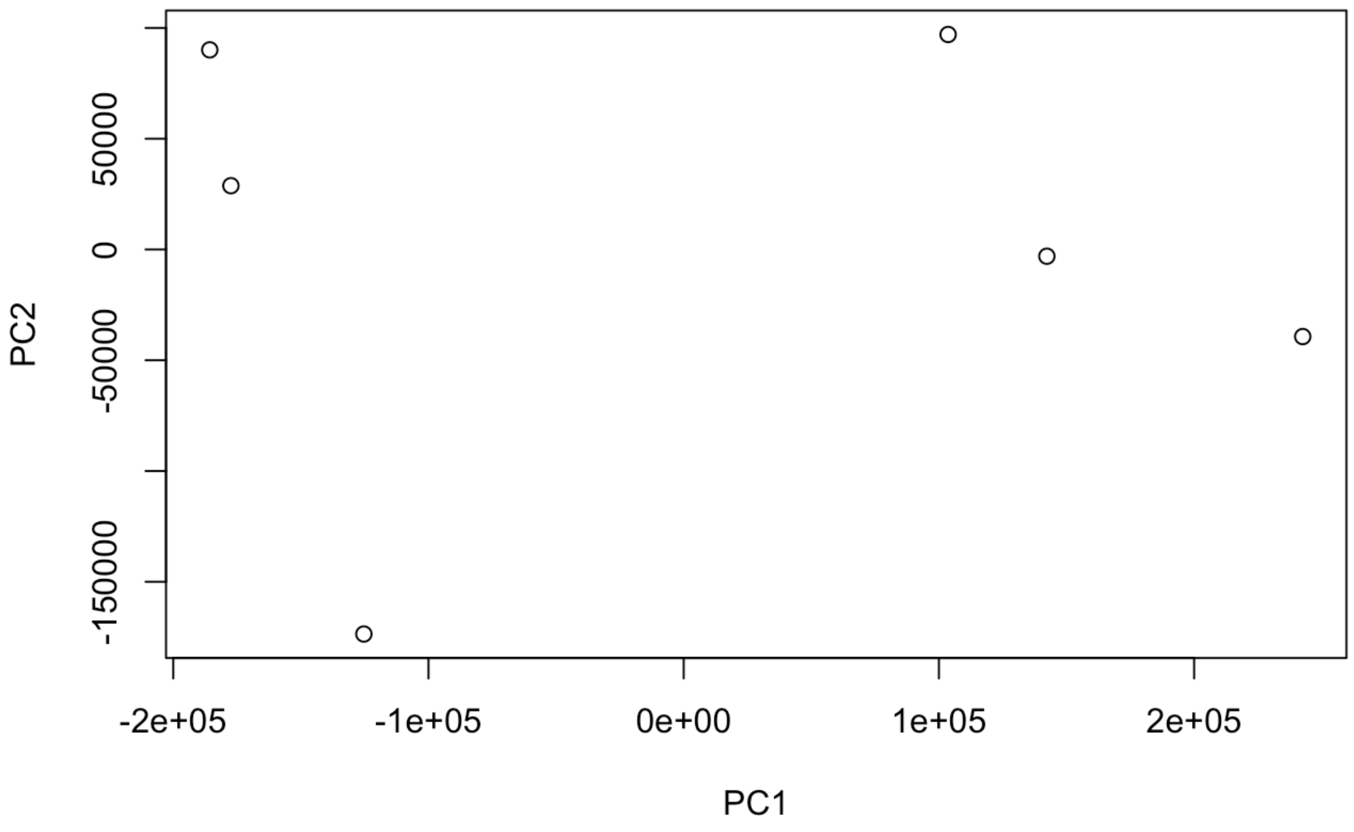
2. PCA for Quality Control

```
pca <- prcomp(t(countData))
summary(pca)
```

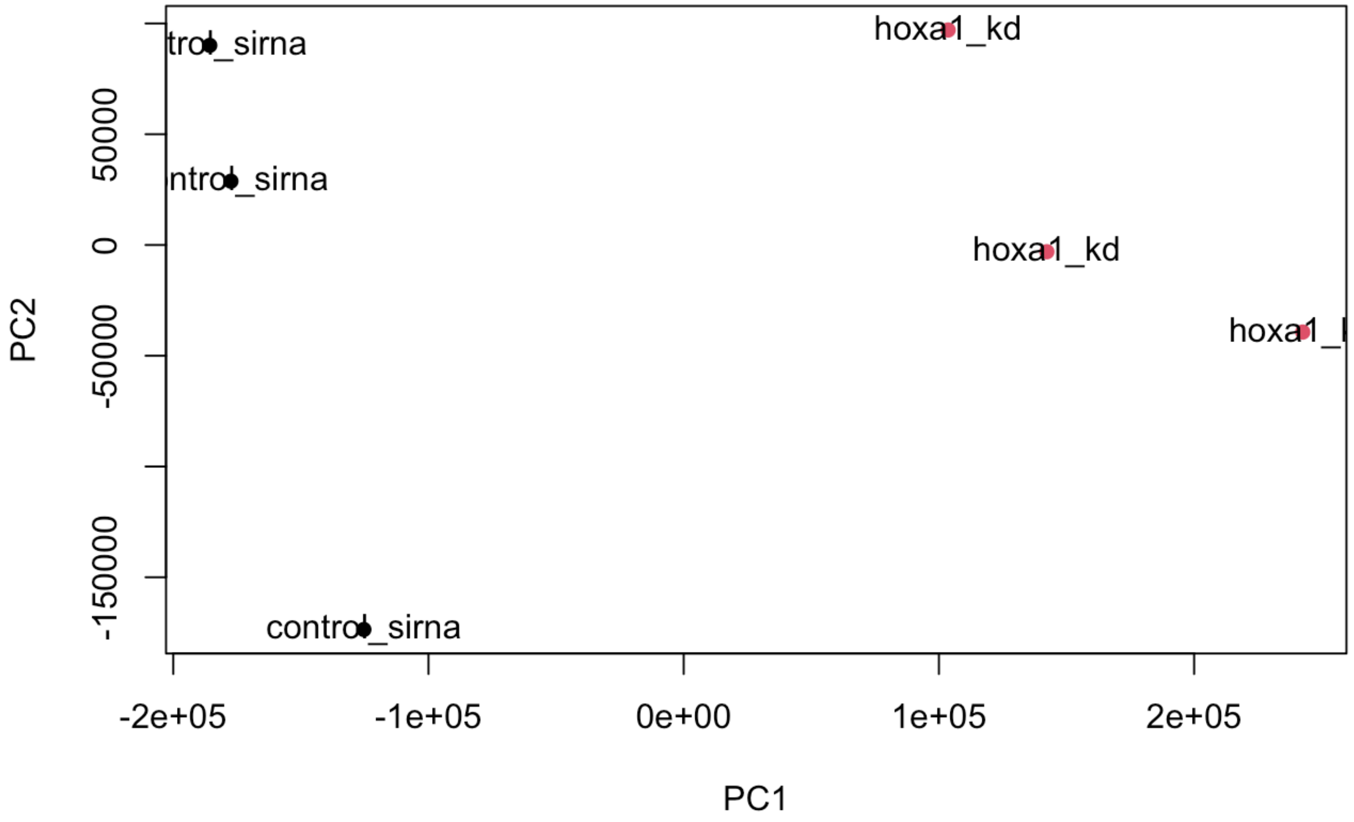
```
## Importance of components:
##                              PC1       PC2       PC3       PC4      PC5
## Standard deviation     1.852e+05 1.001e+05 1.998e+04 6.886e+03 5.15e+03
## Proportion of Variance 7.659e-01 2.235e-01 8.920e-03 1.060e-03 5.90e-04
## Cumulative Proportion  7.659e-01 9.894e-01 9.983e-01 9.994e-01 1.00e+00
##                              PC6
## Standard deviation     9.558e-10
## Proportion of Variance 0.000e+00
## Cumulative Proportion  1.000e+00
```

Plot first and second:

```
plot(pca$x)
```



```
plot(pca$x[, 1:2], pch = 16, col = as.factor(colData$condition))
text(pca$x[, 1:2], labels = colData$condition)
```

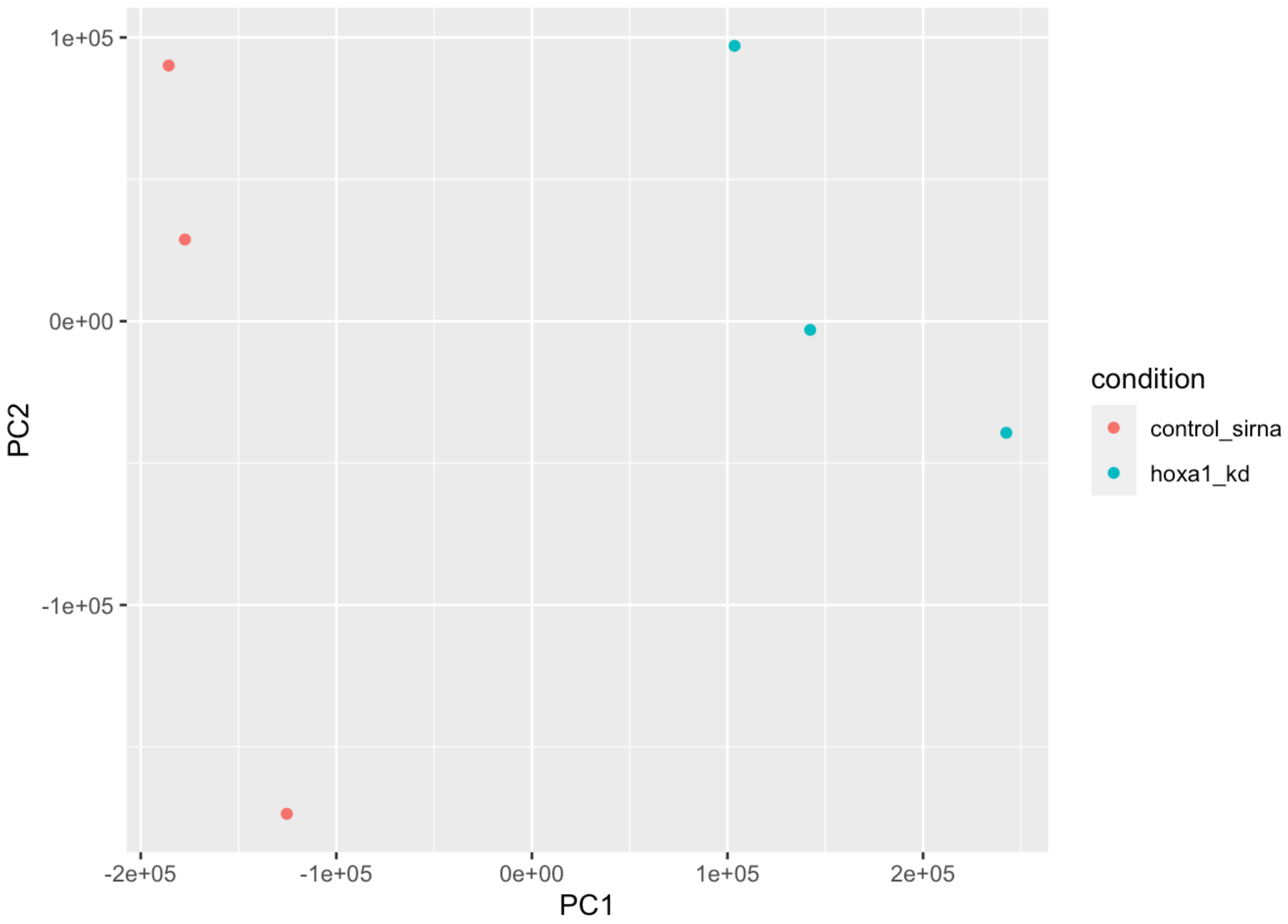


ggplot version:

```
library(ggplot2)

x <- as.data.frame(pca$x)
x$condition <- colData$condition

ggplot(x, aes(PC1, PC2, col=condition)) +
  geom_point()
```



3. Running DESeq2

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

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

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

```
## estimating size factors
```

```
## estimating dispersions
```

```
## gene-wise dispersion estimates
## mean-dispersion relationship
```

```
## final dispersion estimates
```

```
## fitting model and testing
```

```
dds
```

```
## class: DESeqDataSet
## dim: 15975 6
## metadata(1): version
## assays(4): counts mu H cooks
## rownames(15975): ENSG00000279457 ENSG00000187634 ... ENSG00000276345
##   ENSG00000271254
## rowData names(22): baseMean baseVar ... deviance maxCooks
## colnames(6): SRR493366 SRR493367 ... SRR493370 SRR493371
## colData names(2): condition sizeFactor
```

Get result from our DESeq data:

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

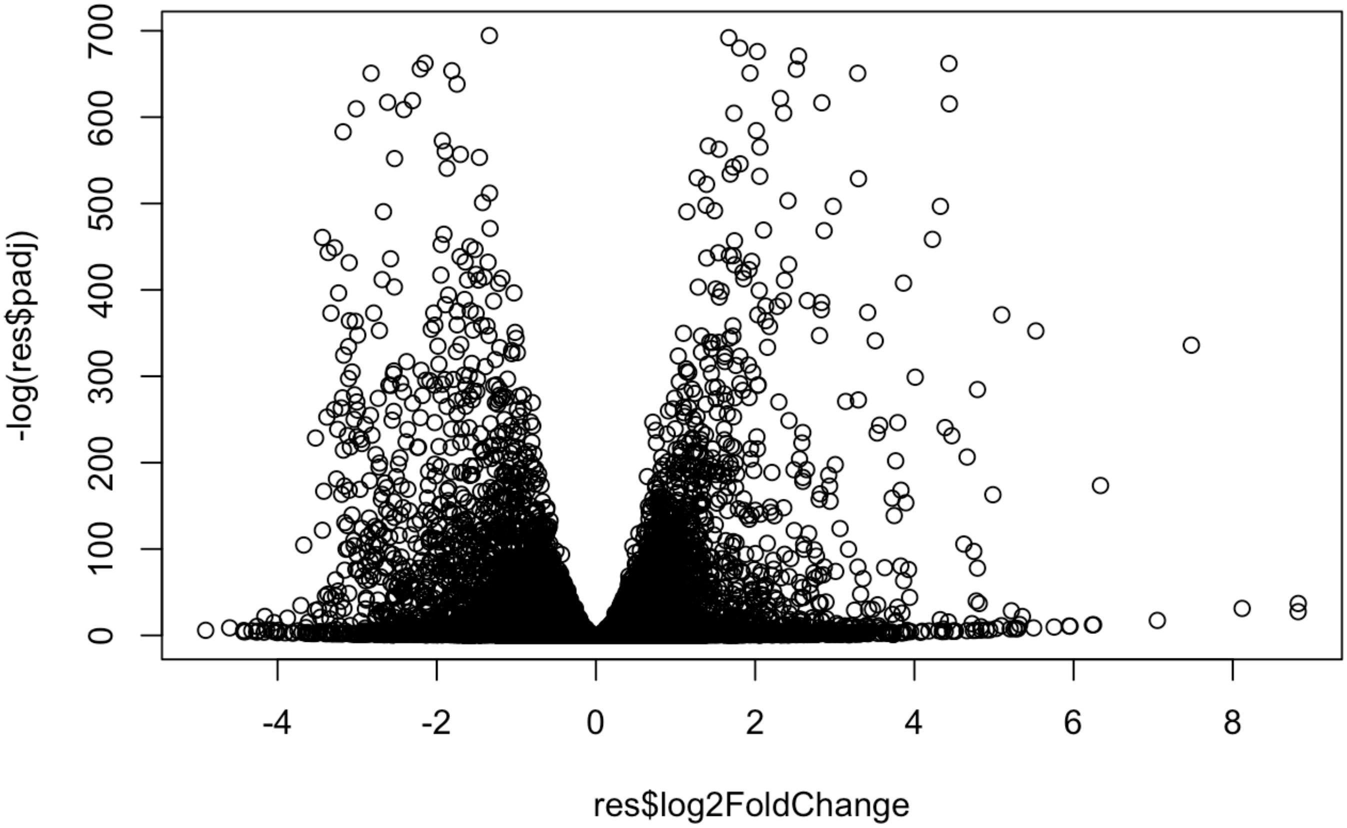
```
summary(res)
```

```
##
## out of 15975 with nonzero total read count
## adjusted p-value < 0.1
## LFC > 0 (up)       : 4349, 27%
## LFC < 0 (down)     : 4396, 28%
## outliers [1]       : 0, 0%
## low counts [2]     : 1237, 7.7%
## (mean count < 0)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

4. Volcano Plot

Let’s do the classic log2-FoldChange vs p-value volcano plot

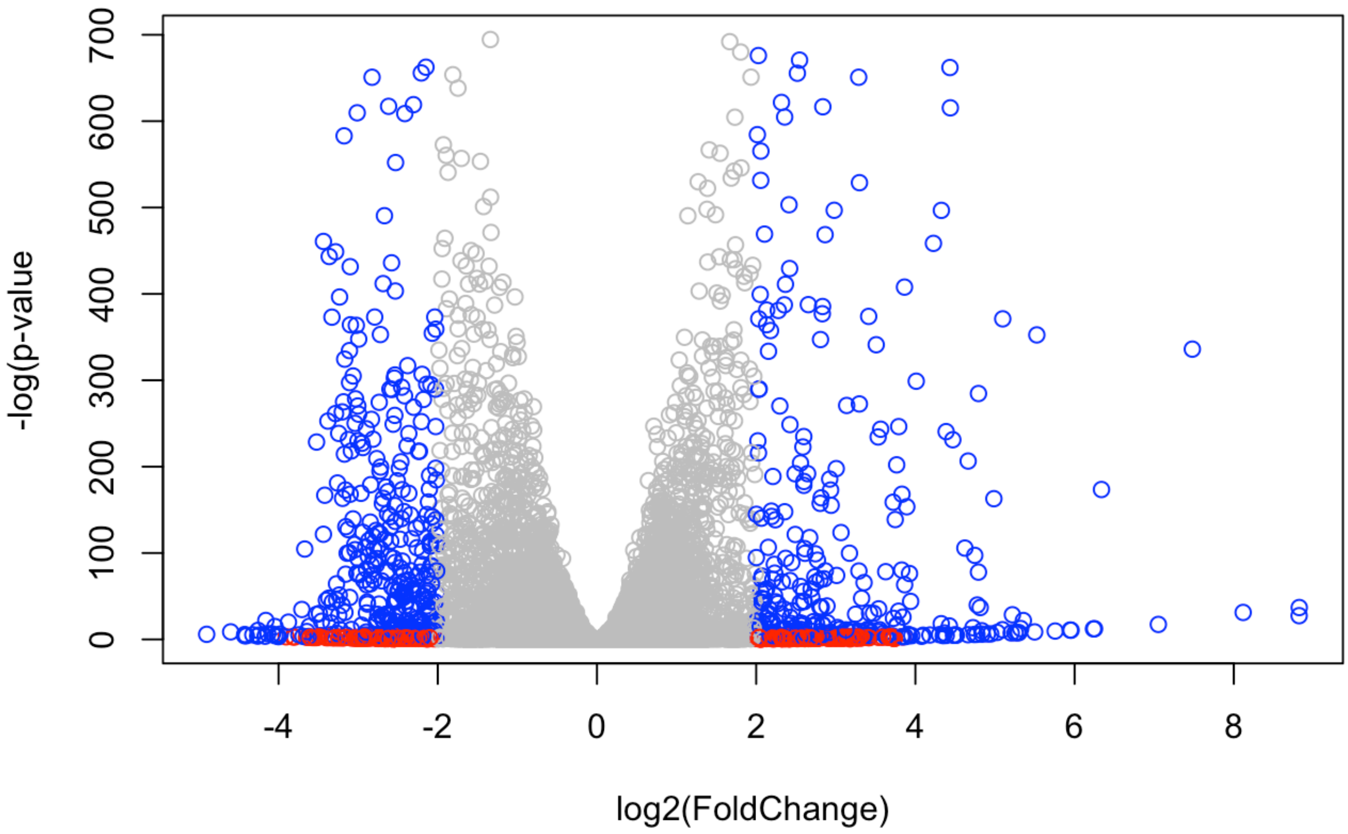
```
plot(res$log2FoldChange, -log(res$padj))
```



Add color

```
mycol <- rep("gray", nrow(res))
mycol[abs(res$log2FoldChange) > 2] <- "blue"
mycol[res$padj > 0.05 & abs(res$log2FoldChange) > 2] <- "red"

plot(res$log2FoldChange, -log(res$padj), col = mycol, xlab = "log2(FoldChange)",
     ylab = "-log(p-value")
```



5. Annotation

```
library("AnnotationDbi")
```

```
## Warning: package 'AnnotationDbi' was built under R version 4.1.2
```

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

```
columns(org.Hs.eg.db)
```

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

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

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

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

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

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

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

```
head(res, 10)
```

```
## log2 fold change (MLE): condition hoxa1 kd vs control sirna
## Wald test p-value: condition hoxa1 kd vs control sirna
## DataFrame with 10 rows and 9 columns
##                    baseMean log2FoldChange     lfcSE       stat      pvalue
##                   <numeric>      <numeric> <numeric>  <numeric>   <numeric>
## ENSG00000279457   29.913579      0.1792571 0.3248216   0.551863 5.81042e-01
## ENSG00000187634  183.229650      0.4264571 0.1402658   3.040350 2.36304e-03
## ENSG00000188976 1651.188076     -0.6927205 0.0548465 -12.630158 1.43990e-36
## ENSG00000187961  209.637938      0.7297556 0.1318599   5.534326 3.12428e-08
## ENSG00000187583   47.255123      0.0405765 0.2718928   0.149237 8.81366e-01
## ENSG00000187642   11.979750      0.5428105 0.5215598   1.040744 2.97994e-01
## ENSG00000188290  108.922128      2.0570638 0.1969053  10.446970 1.51282e-25
## ENSG00000187608  350.716868      0.2573837 0.1027266   2.505522 1.22271e-02
## ENSG00000188157 9128.439422      0.3899088 0.0467163   8.346304 7.04321e-17
## ENSG00000237330    0.158192      0.7859552 4.0804729   0.192614 8.47261e-01
##                        padj      symbol      entrez                   name
##                   <numeric> <character> <character>            <character>
## ENSG00000279457 6.86555e-01      WASH9P   102723897 WAS protein family h..
## ENSG00000187634 5.15718e-03      SAMD11      148398 sterile alpha motif ..
## ENSG00000188976 1.76549e-35       NOC2L       26155 NOC2 like nucleolar ..
## ENSG00000187961 1.13413e-07      KLHL17      339451 kelch like family me..
## ENSG00000187583 9.19031e-01     PLEKHN1       84069 pleckstrin homology ..
## ENSG00000187642 4.03379e-01       PERM1       84808 PPARGC1 and ESRR ind..
## ENSG00000188290 1.30538e-24        HES4       57801 hes family bHLH tran..
## ENSG00000187608 2.37452e-02       ISG15        9636 ISG15 ubiquitin like..
## ENSG00000188157 4.21963e-16        AGRN      375790                  agrin
## ENSG00000237330          NA      RNF223      401934 ring finger protein ..
```

6. Pathway Analysis

Use KEGG pathways:

```
library(pathview)
library(gage)
library(gageData)
```

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

# Focus on signaling and metabolic pathways only
kegg.sets.hs = kegg.sets.hs[sigmet.idx.hs]

# Examine the first 3 pathways
head(kegg.sets.hs, 3)
```

```
## $`hsa00232 Caffeine metabolism`
## [1] "10"   "1544" "1548" "1549" "1553" "7498" "9"
##
## $`hsa00983 Drug metabolism - other enzymes`
##  [1] "10"     "1066"   "10720"  "10941"  "151531" "1548"   "1549"   "1551"
##  [9] "1553"   "1576"   "1577"   "1806"   "1807"   "1890"   "221223" "2990"
## [17] "3251"   "3614"   "3615"   "3704"   "51733"  "54490"  "54575"  "54576"
## [25] "54577"  "54578"  "54579"  "54600"  "54657"  "54658"  "54659"  "54963"
## [33] "574537" "64816"  "7083"   "7084"   "7172"   "7363"   "7364"   "7365"
## [41] "7366"   "7367"   "7371"   "7372"   "7378"   "7498"   "79799"  "83549"
## [49] "8824"   "8833"   "9"      "978"
##
## $`hsa00230 Purine metabolism`
##   [1] "100"    "10201"  "10606"  "10621"  "10622"  "10623"  "107"    "10714"
##   [9] "108"    "10846"  "109"    "111"    "11128"  "11164"  "112"    "113"
##  [17] "114"    "115"    "122481" "122622" "124583" "132"    "158"    "159"
##  [25] "1633"   "171568" "1716"   "196883" "203"    "204"    "205"    "221823"
##  [33] "2272"   "22978"  "23649"  "246721" "25885"  "2618"   "26289"  "270"
##  [41] "271"    "27115"  "272"    "2766"   "2977"   "2982"   "2983"   "2984"
##  [49] "2986"   "2987"   "29922"  "3000"   "30833"  "30834"  "318"    "3251"
##  [57] "353"    "3614"   "3615"   "3704"   "377841" "471"    "4830"   "4831"
##  [65] "4832"   "4833"   "4860"   "4881"   "4882"   "4907"   "50484"  "50940"
##  [73] "51082"  "51251"  "51292"  "5136"   "5137"   "5138"   "5139"   "5140"
##  [81] "5141"   "5142"   "5143"   "5144"   "5145"   "5146"   "5147"   "5148"
##  [89] "5149"   "5150"   "5151"   "5152"   "5153"   "5158"   "5167"   "5169"
##  [97] "51728"  "5198"   "5236"   "5313"   "5315"   "53343"  "54107"  "5422"
## [105] "5424"   "5425"   "5426"   "5427"   "5430"   "5431"   "5432"   "5433"
## [113] "5434"   "5435"   "5436"   "5437"   "5438"   "5439"   "5440"   "5441"
## [121] "5471"   "548644" "55276"  "5557"   "5558"   "55703"  "55811"  "55821"
## [129] "5631"   "5634"   "56655"  "56953"  "56985"  "57804"  "58497"  "6240"
## [137] "6241"   "64425"  "646625" "654364" "661"    "7498"   "8382"   "84172"
## [145] "84265"  "84284"  "84618"  "8622"   "8654"   "87178"  "8833"   "9060"
## [153] "9061"   "93034"  "953"    "9533"   "954"    "955"    "956"    "957"
## [161] "9583"   "9615"
```

Make the input foldchange vector for KEGG and GO:

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

```
##   102723897      148398       26155      339451       84069       84808
##  0.17925708  0.42645712 -0.69272046  0.72975561  0.04057653  0.54281049
```

```
# Get the results
keggres = gage(foldchanges, gsets=kegg.sets.hs)
```

Look at the object return from `gage()`

```
attributes(keggres)
```

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

Downregulated pathway:

```
# Look at the first few down (less) pathways
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
```

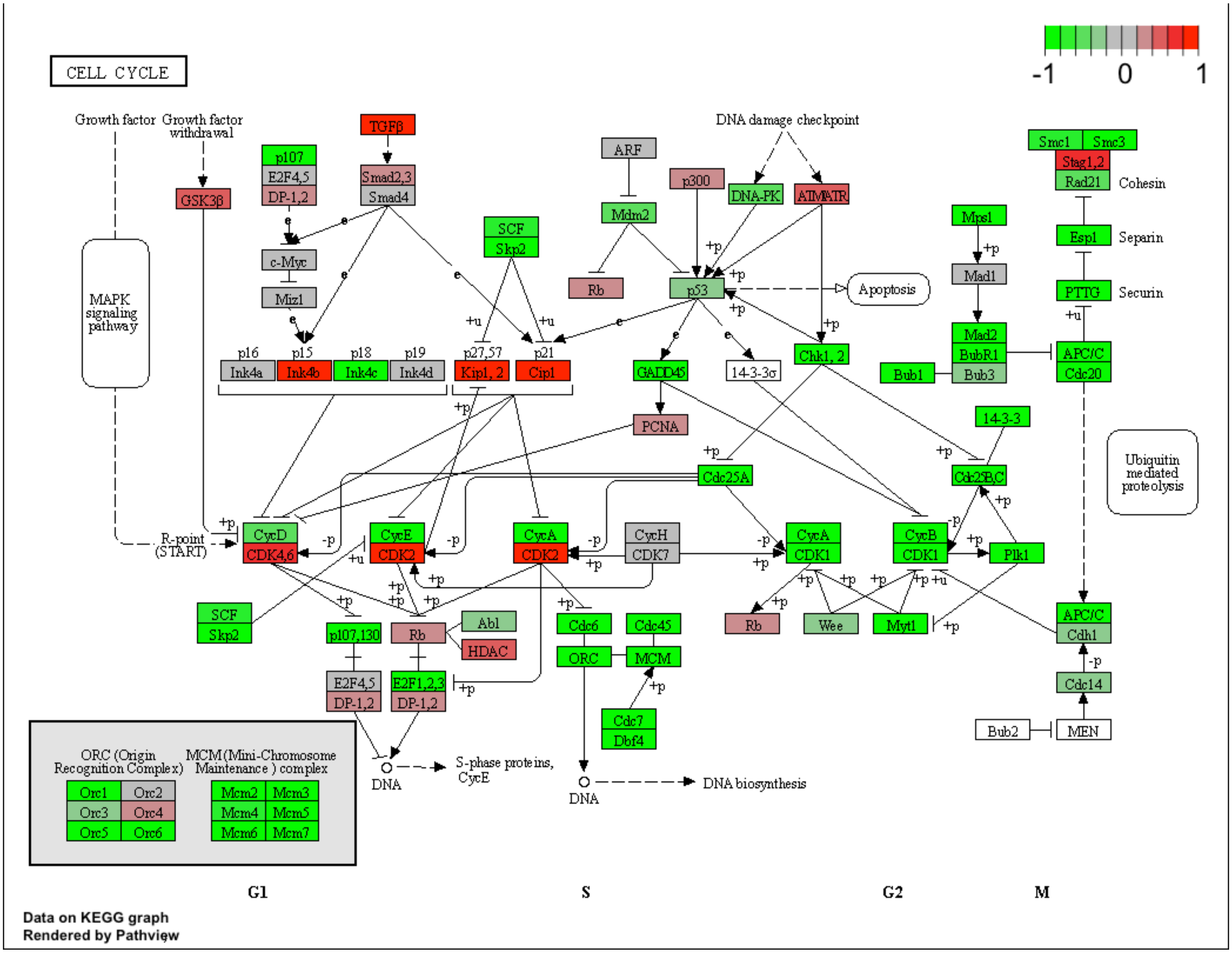
Let’s look at the first downregulated pathway:

```
pathview(gene.data=foldchanges, pathway.id="hsa04110")
```

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

```
## Info: Working in directory /Users/deka/Dropbox/My Mac (ciaiqinmachudeMacBook-Pro.l
ocal)/Documents/BGGN213_R/bggn213/class16
```

```
## Info: Writing image file hsa04110.pathview.png
```



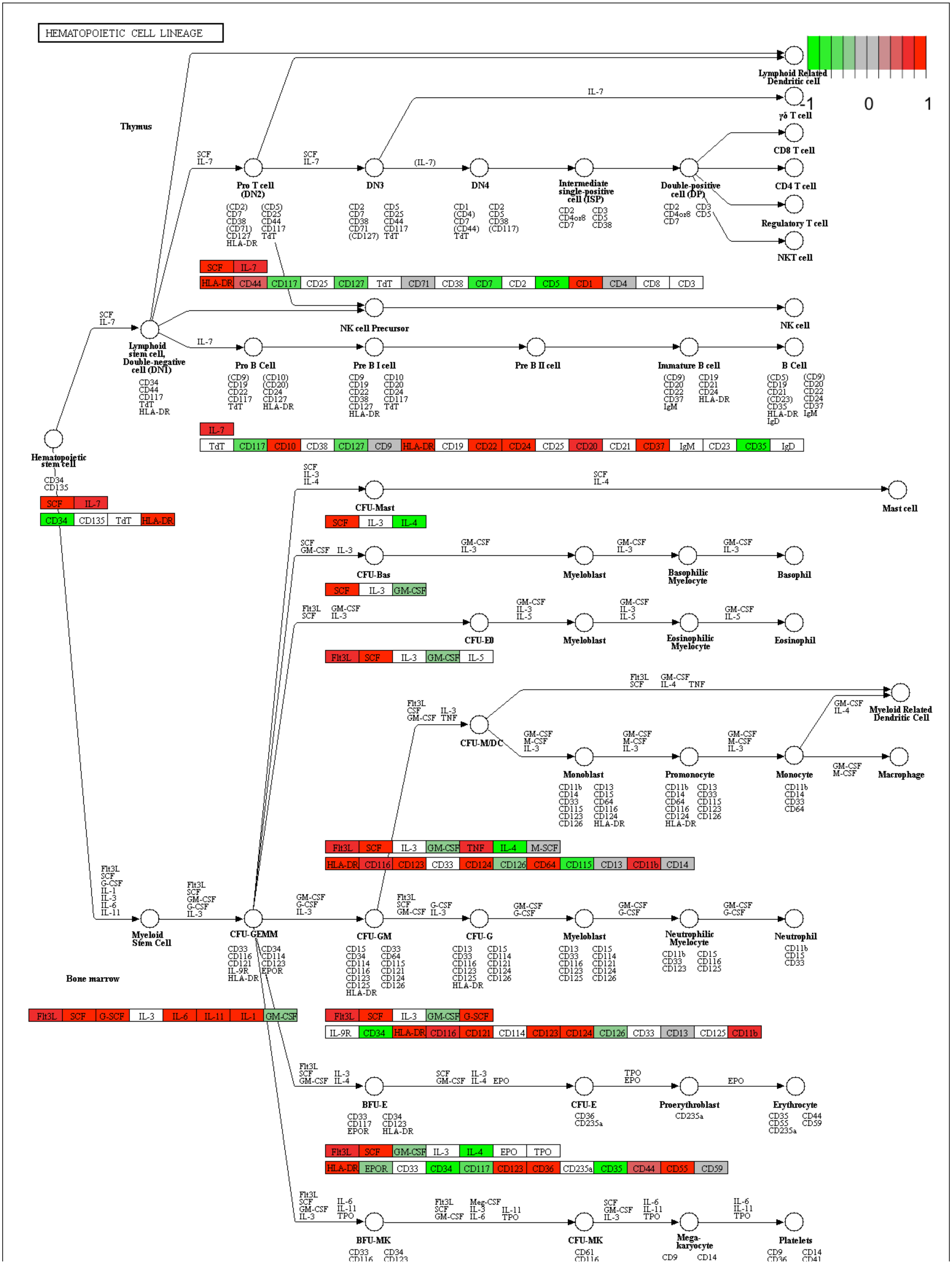
We can automatically pull up 5 upregulated pathway by doing so

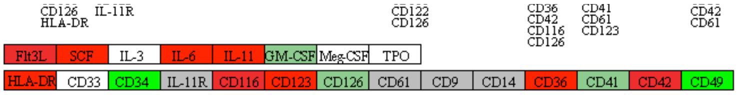
```
keggrespathways <- rownames(keggres$greater)[1:5]

keggresids = substr(keggrespathways, start=1, stop=8)
keggresids
```

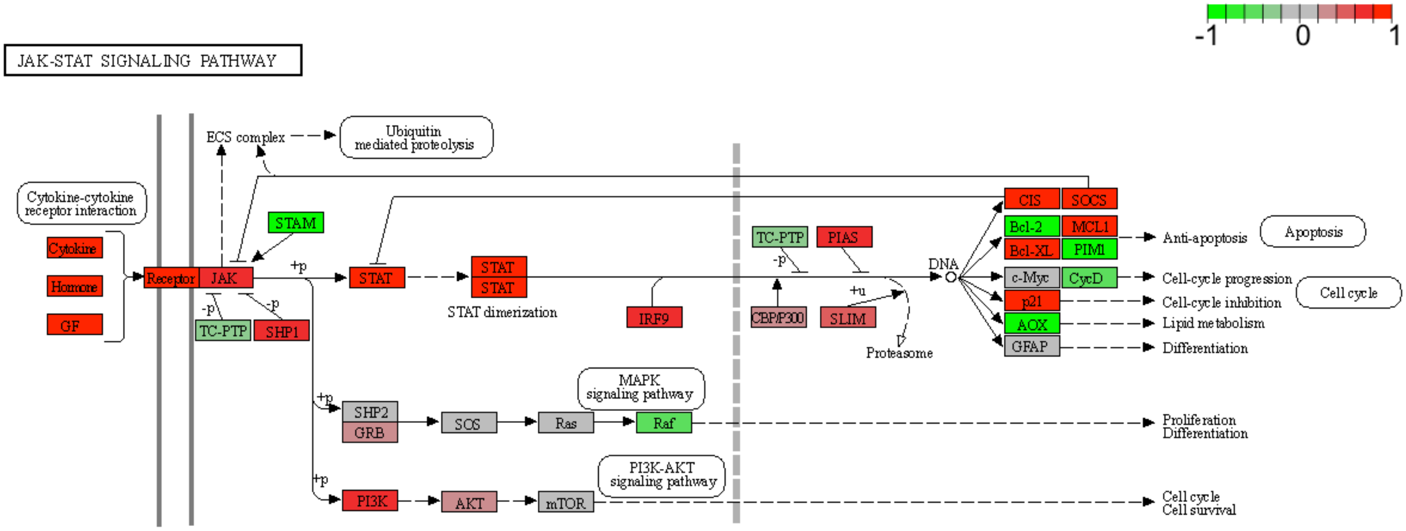
```
## [1] "hsa04640" "hsa04630" "hsa00140" "hsa04142" "hsa04330"
```

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

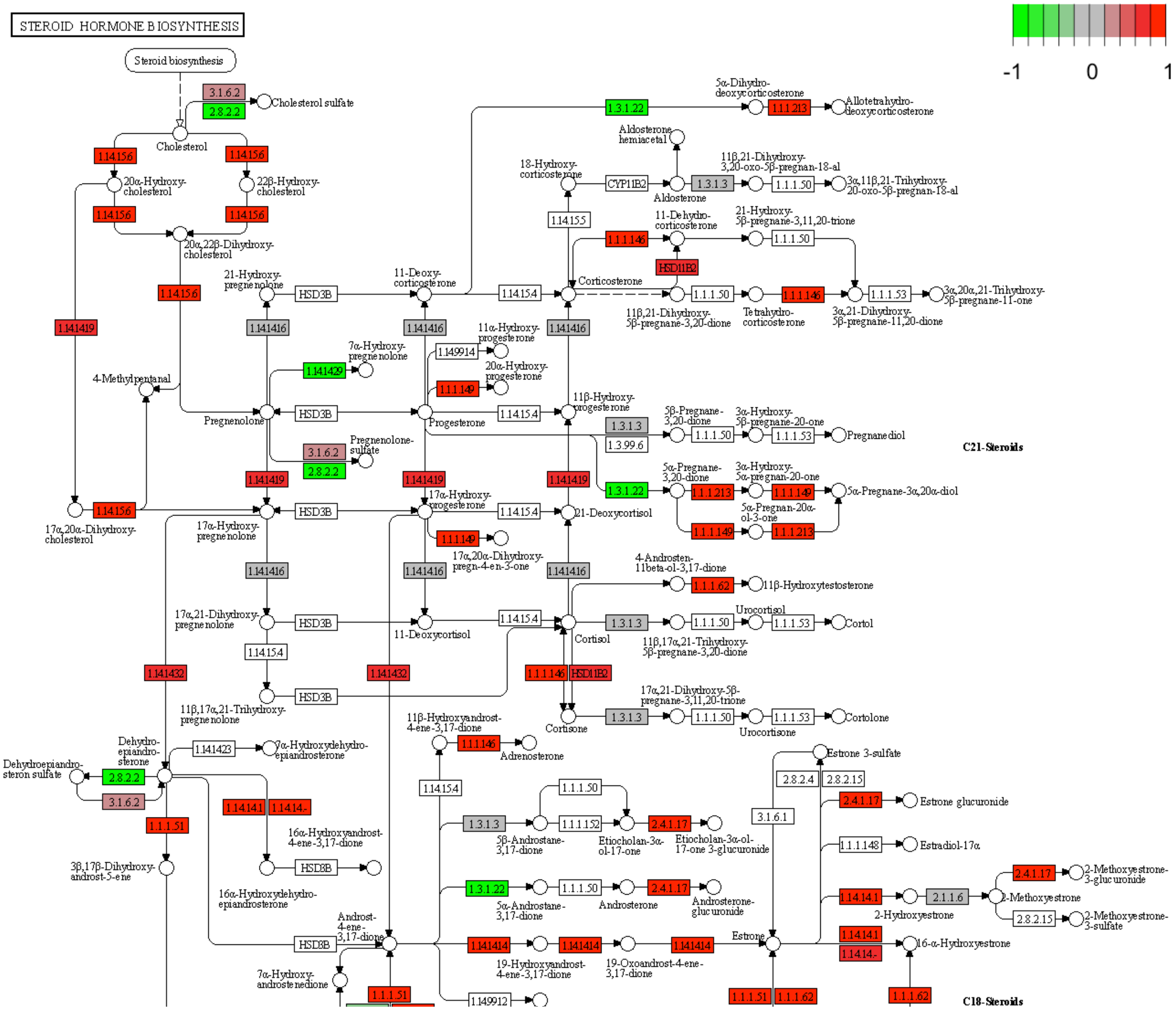

HEMATOPOIETIC CELL LINEAGE
-1
0
1
Thymus
Lymphoid Related Dendritic cell
IL-7
γδ T cell
SCF IL-7
SCF IL-7
(IL-7)
Pro T cell (DN2)
DN3
DN4
Intermediate single-positive cell (ISP)
Double-positive cell (DP)
CD8 T cell
CD4 T cell
Regulatory T cell
NKT cell
(CD2) (CD5) CD7 CD25 CD38 CD44 (CD71) CD117 CD127 TdT HLA-DR
CD2 CD5 CD7 CD25 CD38 CD44 CD71 CD117 (CD127) TdT
CD1 CD2 (CD4) CD5 CD7 CD38 (CD44) (CD117) TdT
CD2 CD3 CD4or8 CD5 CD7 CD38
CD2 CD3 CD4or8 CD5 CD7
SCF
IL-7
HLA-DR
CD44
CD117
CD25
CD127
TdT
CD71
CD38
CD7
CD2
CD5
CD1
CD4
CD8
CD3
NK cell Precursor
NK cell
SCF IL-7
IL-7
Lymphoid stem cell, Double-negative cell (DN1)
CD34 CD44 CD117 TdT HLA-DR
Pro B Cell
Pre B I cell
Pre B II cell
Immature B cell
B Cell
(CD9) (CD10) CD19 (CD20) CD22 CD24 CD117 CD127 TdT HLA-DR
CD9 CD10 CD19 CD20 CD22 CD24 CD38 CD117 CD127 TdT HLA-DR
(CD9) CD19 CD20 CD21 CD22 CD24 CD37 HLA-DR IgM
(CD5) (CD9) CD19 CD20 CD21 CD22 (CD23) CD24 CD35 CD37 HLA-DR IgM IgD
IL-7
TdT
CD117
CD10
CD38
CD127
CD9
HLA-DR
CD19
CD22
CD24
CD25
CD20
CD21
CD37
IgM
CD23
CD35
IgD
Hematopoietic stem cell
CD34 CD135
SCF
IL-7
CD34
CD135
TdT
HLA-DR
SCF IL-3 IL-4
SCF IL-4
CFU-Mast
Mast cell
SCF
IL-3
IL-4
SCF GM-CSF IL-3
GM-CSF IL-3
GM-CSF IL-3
GM-CSF IL-3
CFU-Bas
Myeloblast
Basophilic Myelocyte
Basophil
SCF
IL-3
GM-CSF
Flt3L SCF GM-CSF IL-3
GM-CSF IL-3 IL-5
GM-CSF IL-3 IL-5
GM-CSF IL-5
CFU-E0
Myeloblast
Eosinophilic Myelocyte
Eosinophil
Flt3L
SCF
IL-3
GM-CSF
IL-5
Flt3L SCF GM-CSF IL-4 TNF
Flt3L CSF GM-CSF IL-3 TNF
GM-CSF IL-4
CFU-M/DC
Myeloid Related Dendritic Cell
GM-CSF M-CSF IL-3
GM-CSF M-CSF IL-3
GM-CSF M-CSF IL-3
GM-CSF M-CSF
Monoblast
Promonocyte
Monocyte
Macrophage
CD11b CD13 CD14 CD15 CD33 CD64 CD115 CD116 CD123 CD124 CD126 HLA-DR
CD11b CD13 CD14 CD33 CD64 CD115 CD116 CD123 CD124 CD126 HLA-DR
CD11b CD14 CD33 CD64
Flt3L
SCF
IL-3
GM-CSF
TNF
IL-4
M-SCF
HLA-DR
CD116
CD123
CD33
CD124
CD126
CD64
CD115
CD13
CD11b
CD14
Flt3L SCF G-CSF IL-1 IL-3 IL-6 IL-11
Flt3L SCF GM-CSF G-CSF IL-3
GM-CSF G-CSF IL-3
Flt3L SCF GM-CSF G-CSF IL-3
GM-CSF G-CSF
GM-CSF G-CSF
GM-CSF G-CSF
Myeloid Stem Cell
CFU-GEMM
CFU-GM
CFU-G
Myeloblast
Neutrophilic Myelocyte
Neutrophil
CD33 CD34 CD116 CD114 CD121 CD123 IL-9R EPOR HLA-DR
CD15 CD33 CD34 CD64 CD114 CD115 CD116 CD121 CD123 CD124 CD125 CD126 HLA-DR
CD13 CD15 CD33 CD114 CD116 CD121 CD123 CD124 CD125 CD126 HLA-DR
CD13 CD15 CD33 CD114 CD116 CD121 CD123 CD124 CD125 CD126
CD11b CD15 CD33 CD116 CD123 CD125
CD11b CD15 CD33
Bone marrow
Flt3L
SCF
G-SCF
IL-3
IL-6
IL-11
IL-1
GM-CSF
Flt3L
SCF
IL-3
GM-CSF
G-SCF
IL-9R
CD34
HLA-DR
CD116
CD121
CD114
CD123
CD124
CD126
CD33
CD13
CD125
CD11b
Flt3L SCF GM-CSF IL-3 IL-4
SCF GM-CSF IL-3 IL-4 EPO
TPO EPO
EPO
BFU-E
CFU-E
Proerythroblast
Erythrocyte
CD33 CD34 CD117 CD123 EPOR HLA-DR
CD36 CD235a
CD235a
CD35 CD44 CD55 CD59 CD235a
Flt3L
SCF
GM-CSF
IL-3
IL-4
EPO
TPO
HLA-DR
EPOR
CD33
CD34
CD117
CD123
CD36
CD235a
CD35
CD44
CD55
CD59
Flt3L SCF GM-CSF IL-3 IL-6 IL-11 TPO
Flt3L SCF GM-CSF Meg-CSF IL-3 IL-6 IL-11 TPO
SCF GM-CSF IL-3 IL-6 IL-11 TPO
IL-6 IL-11 TPO
BFU-MK
CFU-MK
Mega-karyocyte
Platelets
CD33 CD34 CD116 CD123
CD61 CD116
CD9 CD14
CD9 CD14 CD36 CD41

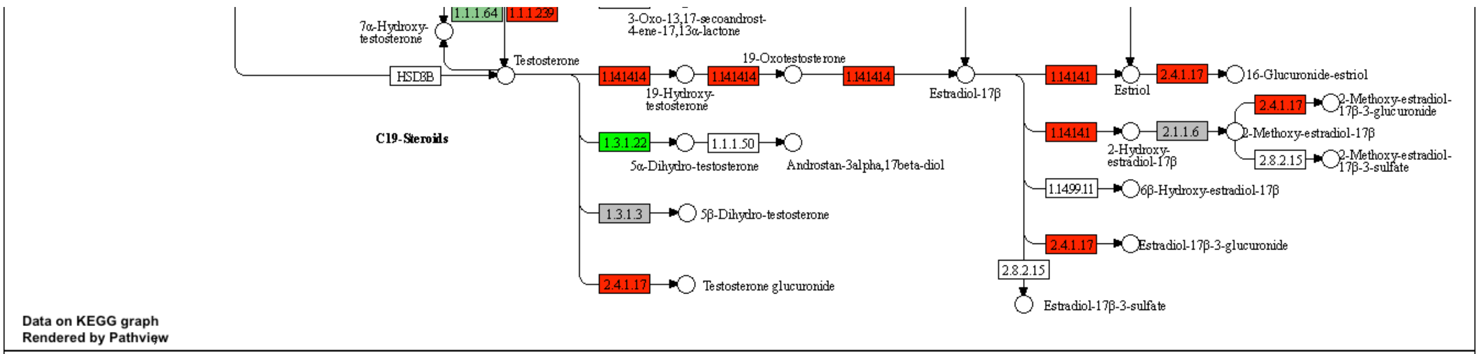


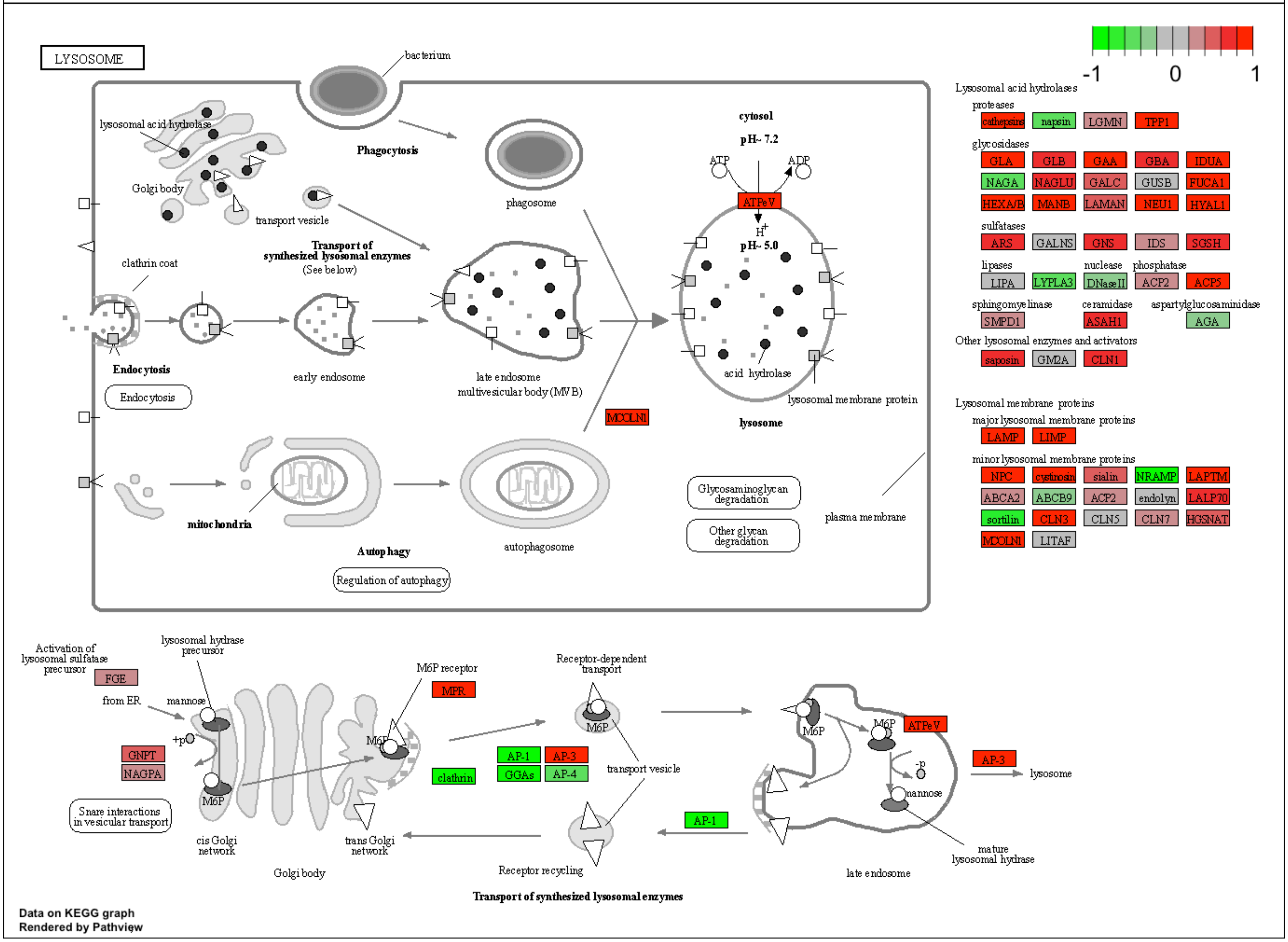
Data on KEGG graph
Rendered by Pathview

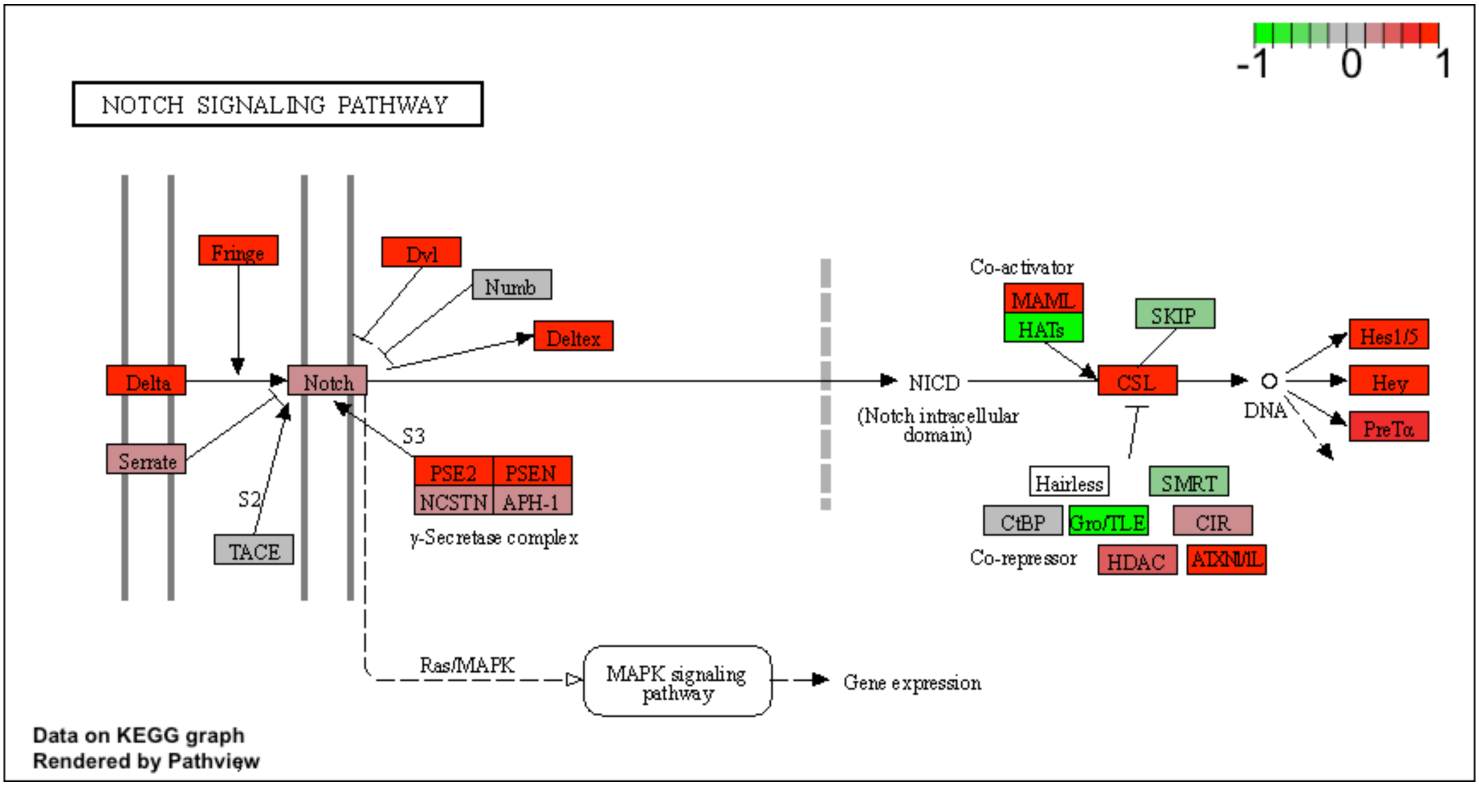


Data on KEGG graph
Rendered by Pathview




1.1.1.64
1.1.1.239
7α-Hydroxy-testosterone
3-Oxo-13,17-secoandrost-4-ene-17,13α-lactone
HSD8B
Testosterone
1.14.14.14
19-Hydroxy-testosterone
1.14.14.14
19-Oxotestosterone
1.14.14.14
Estradiol-17β
1.14.14.1
Estriol
2.4.1.17
16-Glucuronide-estriol
2.4.1.17
2-Methoxy-estradiol-17β-3-glucuronide
1.14.14.1
2-Hydroxy-estradiol-17β
2.1.1.6
2-Methoxy-estradiol-17β
2.8.2.15
2-Methoxy-estradiol-17β-3-sulfate
1.14.99.11
6β-Hydroxy-estradiol-17β
2.4.1.17
Estradiol-17β-3-glucuronide
2.8.2.15
Estradiol-17β-3-sulfate
C19-Steroids
1.3.1.22
5α-Dihydro-testosterone
1.1.1.50
Androstan-3alpha,17beta-diol
1.3.1.3
5β-Dihydro-testosterone
2.4.1.17
Testosterone glucuronide
Data on KEGG graph
Rendered by Pathview


LYSOSOME
-1
0
1
bacterium
lysosomal acid hydrolase
Golgi body
transport vesicle
Phagocytosis
phagosome
Transport of synthesized lysosomal enzymes
(See below)
clathrin coat
Endocytosis
Endocytosis
early endosome
late endosome
multivesicular body (MVB)
cytosol
pH- 7.2
ATP
ADP
ATPeV
H+
pH- 5.0
acid hydrolase
lysosomal membrane protein
lysosome
MCOLN1
mitochondria
Autophagy
Regulation of autophagy
autophagosome
Glycosaminoglycan degradation
Other glycan degradation
plasma membrane
Lysosomal acid hydrolases
proteases
cathepsins
napsin
LGMN
TPP1
glycosidases
GLA
GLB
GAA
GBA
IDUA
NAGA
NAGLU
GALC
GUSB
FUCA1
HEXA/B
MANB
LAMAN
NEU1
HYAL1
sulfatases
ARS
GALNS
GNS
IDS
SGSH
lipases
LIPA
LYPLA3
nuclease
DNaseII
phosphatase
ACP2
ACP5
sphingomyelinase
SMPD1
ceramidase
ASAH1
aspartylglucosaminidase
AGA
Other lysosomal enzymes and activators
saposin
GM2A
CLN1
Lysosomal membrane proteins
major lysosomal membrane proteins
LAMP
LIMP
minor lysosomal membrane proteins
NPC
cystinosin
sialin
NRAMP
LAPTM
ABCA2
ABCB9
ACP2
endolyn
LALP70
sortilin
CLN3
CLN5
CLN7
HGSNAT
MCOLN1
LITAF
Activation of lysosomal sulfatase precursor
FGE
lysosomal hydrase precursor
from ER
mannose
+p
GNPT
NAGPA
M6P
Snare interactions in vesicular transport
cis Golgi network
Golgi body
M6P receptor
MPR
M6P
trans Golgi network
clathrin
AP-1
AP-3
GGAs
AP-4
Receptor-dependent transport
M6P
transport vesicle
Receptor recycling
AP-1
M6P
M6P
ATPeV
-p
mannose
AP-3
lysosome
mature lysosomal hydrase
late endosome
Transport of synthesized lysosomal enzymes
Data on KEGG graph
Rendered by Pathview



We can also do similar thing using gene ontology. Focus on biological process:

```
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, same.dir=TRUE)

lapply(gobpres, head)
```

```
## $greater
##                                              p.geomean stat.mean        p.val
## GO:0007156 homophilic cell adhesion       8.519724e-05  3.824205 8.519724e-05
## GO:0002009 morphogenesis of an epithelium 1.396681e-04  3.653886 1.396681e-04
## GO:0048729 tissue morphogenesis           1.432451e-04  3.643242 1.432451e-04
## GO:0007610 behavior                       2.195494e-04  3.530241 2.195494e-04
## GO:0060562 epithelial tube morphogenesis  5.932837e-04  3.261376 5.932837e-04
## GO:0035295 tube development               5.953254e-04  3.253665 5.953254e-04
##                                               q.val set.size         exp1
## GO:0007156 homophilic cell adhesion       0.1951953      113 8.519724e-05
## GO:0002009 morphogenesis of an epithelium 0.1951953      339 1.396681e-04
## GO:0048729 tissue morphogenesis           0.1951953      424 1.432451e-04
## GO:0007610 behavior                       0.2243795      427 2.195494e-04
## GO:0060562 epithelial tube morphogenesis  0.3711390      257 5.932837e-04
## GO:0035295 tube development               0.3711390      391 5.953254e-04
## 
## $less
##                                             p.geomean stat.mean        p.val
## GO:0048285 organelle fission             1.536227e-15 -8.063910 1.536227e-15
## GO:0000280 nuclear division              4.286961e-15 -7.939217 4.286961e-15
## GO:0007067 mitosis                       4.286961e-15 -7.939217 4.286961e-15
## GO:0000087 M phase of mitotic cell cycle 1.169934e-14 -7.797496 1.169934e-14
## GO:0007059 chromosome segregation        2.028624e-11 -6.878340 2.028624e-11
## GO:0000236 mitotic prometaphase          1.729553e-10 -6.695966 1.729553e-10
##                                                 q.val set.size         exp1
## GO:0048285 organelle fission             5.841698e-12      376 1.536227e-15
## GO:0000280 nuclear division              5.841698e-12      352 4.286961e-15
## GO:0007067 mitosis                       5.841698e-12      352 4.286961e-15
## GO:0000087 M phase of mitotic cell cycle 1.195672e-11      362 1.169934e-14
## GO:0007059 chromosome segregation        1.658603e-08      142 2.028624e-11
## GO:0000236 mitotic prometaphase          1.178402e-07       84 1.729553e-10
## 
## $stats
##                                           stat.mean     exp1
## GO:0007156 homophilic cell adhesion        3.824205 3.824205
## GO:0002009 morphogenesis of an epithelium  3.653886 3.653886
## GO:0048729 tissue morphogenesis            3.643242 3.643242
## GO:0007610 behavior                        3.530241 3.530241
## GO:0060562 epithelial tube morphogenesis   3.261376 3.261376
## GO:0035295 tube development                3.253665 3.253665
```