class16: RNASeq Mini Project

Ethan Harding (PID A15468670)

11/18/2021

1. Differential Expression Analysis

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

# Import metadata and take a peak
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
```

```
# Import countdata
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
```

Q. Complete the code below to remove the troublesome first column from countData

```
# Note we need to remove the odd first $length col
countData2 <- as.matrix(countData[,-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
```

This looks better but there are lots of zero entries in there so let's get rid of them as we have no data for these.

> Q. Complete the code below to filter countData to exclude genes (i.e. rows) where we have 0 read count across all samples (i.e. columns).

```
head(countData2)
```

```
##                 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
```

```
countsnozero <- countData2[rowSums(countData2) !=0, ]
head(countsnozero)
```

```
##                 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
```

2. DESeq Analysis

```
library(DESeq2)
```

```
## Loading required package: S4Vectors

## Loading required package: stats4

## Loading required package: BiocGenerics

## Loading required package: parallel

##
## Attaching package: 'BiocGenerics'

## The following objects are masked from 'package:parallel':
##
##     clusterApply, clusterApplyLB, clusterCall, clusterEvalQ,
##     clusterExport, clusterMap, parApply, parCapply, parLapply,
##     parLapplyLB, parRapply, parSapply, parSapplyLB

## The following objects are masked from 'package:stats':
##
##     IQR, mad, sd, var, xtabs

## The following objects are masked from 'package:base':
##
##     anyDuplicated, append, as.data.frame, basename, cbind, colnames,
##     dirname, do.call, duplicated, eval, evalq, Filter, Find, get, grep,
##     grepl, intersect, is.unsorted, lapply, Map, mapply, match, mget,
##     order, paste, pmax, pmax.int, pmin, pmin.int, Position, rank,
##     rbind, Reduce, rownames, sapply, setdiff, sort, table, tapply,
##     union, unique, unsplit, which.max, which.min

##
## Attaching package: 'S4Vectors'

## 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
```

Setup the objcet required by DESeq

```
dds = DESeqDataSetFromMatrix(countData=countData2,
                             colData=colData,
                             design=~condition)
```

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

```r
dds = DESeq(dds)
```

```
## estimating size factors
```

```
## estimating dispersions
```

```
## gene-wise dispersion estimates
```

```
## mean-dispersion relationship
```

```
## final dispersion estimates
```

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

Get our results

```r
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>
## ENSG00000186092    0.0000             NA        NA        NA         NA
## ENSG00000279928    0.0000             NA        NA        NA         NA
## ENSG00000279457   29.9136       0.179257  0.324822  0.551863 0.58104205
## ENSG00000278566    0.0000             NA        NA        NA         NA
## ENSG00000273547    0.0000             NA        NA        NA         NA
## ENSG00000187634  183.2296       0.426457  0.140266  3.040350 0.00236304
##                       padj
##                  <numeric>
## ENSG00000186092         NA
## ENSG00000279928         NA
## ENSG00000279457 0.68707978
## ENSG00000278566         NA
## ENSG00000273547         NA
## ENSG00000187634 0.00516278
```

```r
res = results(dds, contrast=c("condition", "hoxa1_kd", "control_sirna"))
```

> Q. Call the summary() function on your results to get a sense of how many genes are up or down-regulated at the default 0.1 p-value cutoff.

```r
summary(res)
```

```
## 
## out of 15975 with nonzero total read count
## adjusted p-value < 0.1
## LFC > 0 (up)       : 4349, 27%
## LFC < 0 (down)     : 4393, 27%
## outliers [1]       : 0, 0%
## low counts [2]     : 1221, 7.6%
## (mean count < 0)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

3. Annotation

> Q. Use the mapIDs() function multiple times to add SYMBOL, ENTREZID and GENENAME annotation to our results by completing the code below.

We want to add gene symbols, entrez ID's and gene names.

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

```
##
```

```r
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"
```

```r
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
```

```r
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
```

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

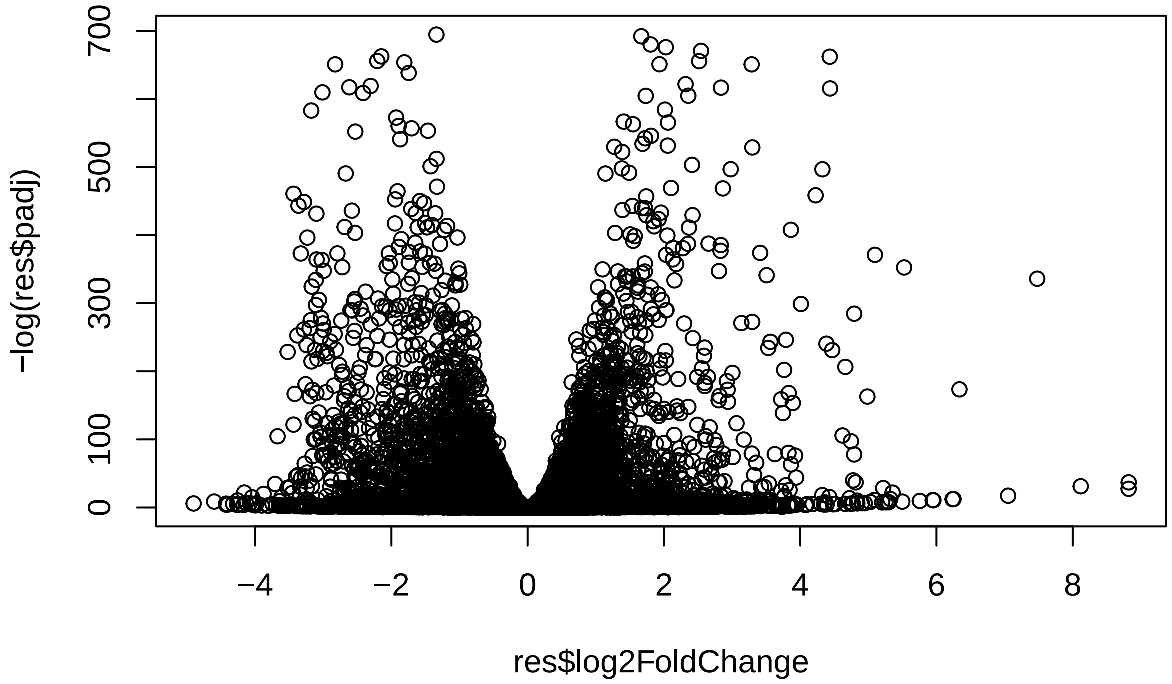
```
## 'select()' returned 1:many mapping between keys and columns
```

```
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>
## ENSG00000186092    0.0000             NA        NA         NA          NA
## ENSG00000279928    0.0000             NA        NA         NA          NA
## ENSG00000279457   29.9136      0.1792571 0.3248216   0.551863 5.81042e-01
## ENSG00000278566    0.0000             NA        NA         NA          NA
## ENSG00000273547    0.0000             NA        NA         NA          NA
## ENSG00000187634  183.2296      0.4264571 0.1402658   3.040350 2.36304e-03
## ENSG00000188976 1651.1881     -0.6927205 0.0548465 -12.630158 1.43990e-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 0.5215598   1.040744 2.97994e-01
##                        padj      symbol      entrez                   name
##                   <numeric> <character> <character>            <character>
## ENSG00000186092          NA       OR4F5       79501 olfactory receptor f..
## ENSG00000279928          NA          NA          NA                     NA
## ENSG00000279457 6.87080e-01      WASH9P   102723897 WAS protein family h..
## ENSG00000278566          NA          NA          NA                     NA
## ENSG00000273547          NA          NA          NA                     NA
## ENSG00000187634 5.16278e-03      SAMD11      148398 sterile alpha motif ..
## ENSG00000188976 1.76741e-35       NOC2L       26155 NOC2 like nucleolar ..
## ENSG00000187961 1.13536e-07      KLHL17      339451 kelch like family me..
## ENSG00000187583 9.18988e-01     PLEKHN1       84069 pleckstrin homology ..
## ENSG00000187642 4.03817e-01       PERM1       84808 PPARGC1 and ESRR ind..
```

4. Volcano Plot

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



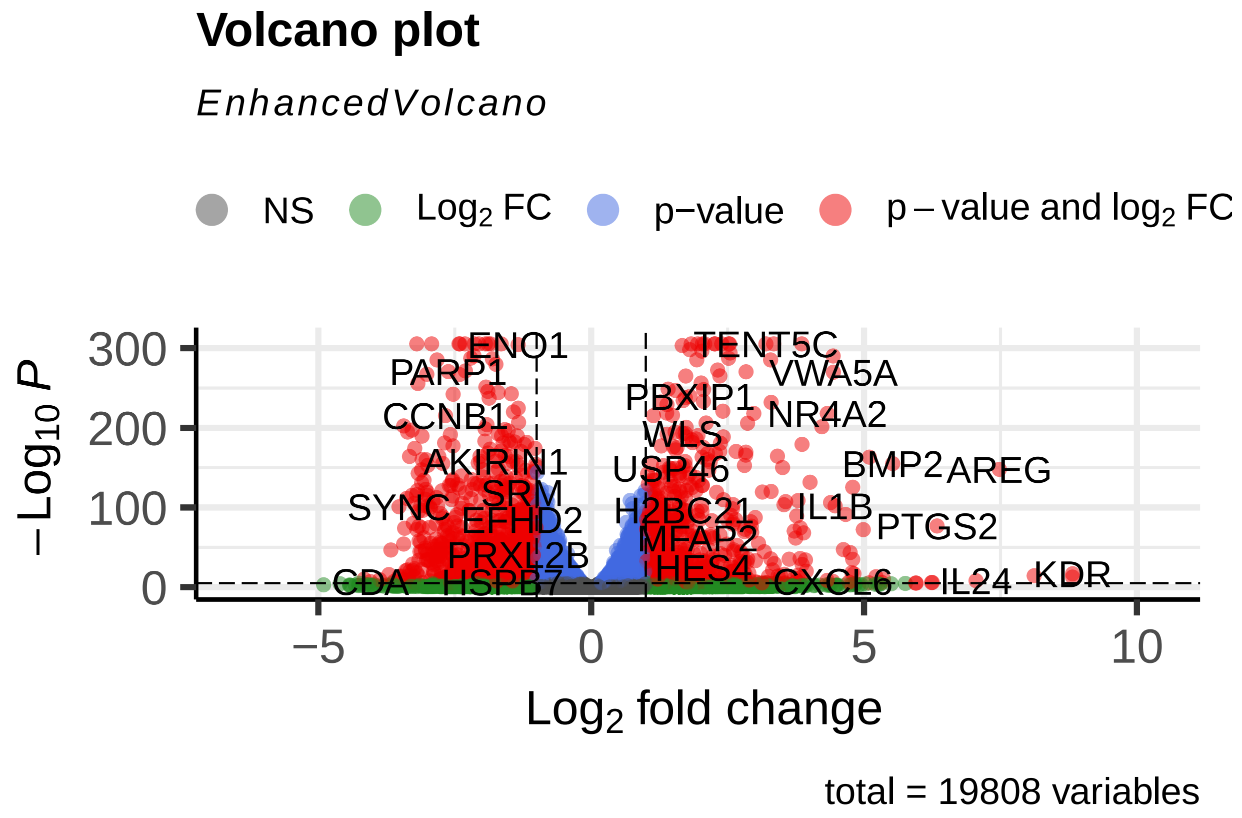
Q. Improve this plot by completing the below code, which adds color and axis labels

```
library(EnhancedVolcano)
```

```
## Loading required package: ggplot2

## Loading required package: ggrepel

## Registered S3 methods overwritten by 'ggalt':
##   method                  from
##   grid.draw.absoluteGrob  ggplot2
##   grobHeight.absoluteGrob ggplot2
##   grobWidth.absoluteGrob  ggplot2
##   grobX.absoluteGrob      ggplot2
##   grobY.absoluteGrob      ggplot2
```

```
x <- as.data.frame(res)

EnhancedVolcano(x,
    lab = x$symbol,
    x = 'log2FoldChange',
    y = 'pvalue')
```

```
## Warning: One or more p-values is 0. Converting to 10^-1 * current lowest non-
## zero p-value...
## Warning: Ignoring unknown parameters: xlim, ylim
```

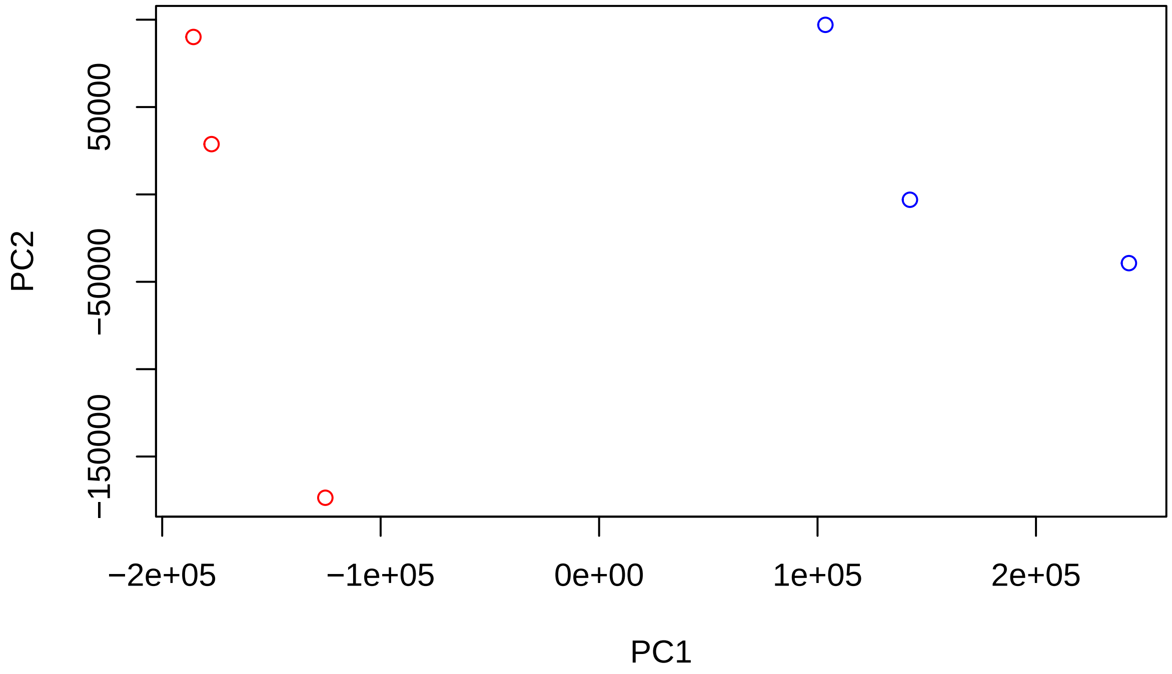


PCA Plot

```r
pca <- prcomp(t(countsnozero))
```

```r
mycols <- rep(c("red", "blue"), each=3)
```

```r
plot(pca$x[,1:2], col=mycols)
```



5. Pathway Analysis

```
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)
```

Focus on the signaling subset of KEGG

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

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

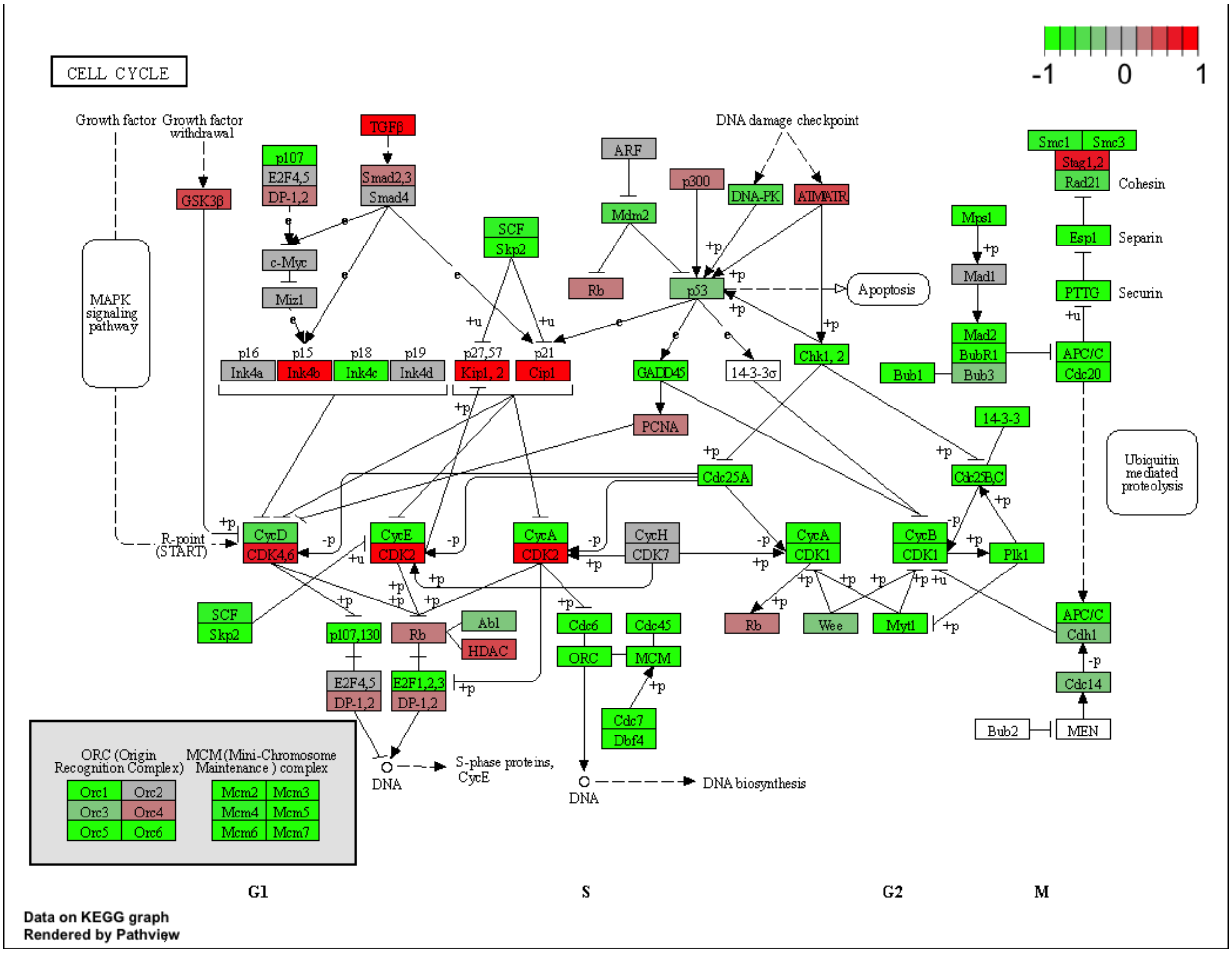
```
foldchanges = res$log2FoldChange
names(foldchanges) = res$entrez
head(foldchanges)
```

```
##      79501      <NA> 102723897      <NA>      <NA>    148398
##         NA        NA 0.1792571        NA        NA 0.4264571
```

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

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

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

## Info: Working in directory /Users/ethanharding/Desktop/BIMM 143/bimm143_github/class16

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



```r
# A different PDF based output of the same data
pathview(gene.data=foldchanges, pathway.id="hsa04110", kegg.native=FALSE)
```

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

```
## Info: Working in directory /Users/ethanharding/Desktop/BIMM 143/bimm143_github/class16
```

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

```r
## Focus on top 5 upregulated pathways here for demo purposes only
keggrespathways <- rownames(keggres$greater)[1:5]

# Extract the 8 character long IDs part of each string
keggresids = substr(keggrespathways, start=1, stop=8)
keggresids
```

```
## [1] "hsa04740" "hsa04640" "hsa00140" "hsa04630" "hsa04976"
```

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

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

```
## Info: Working in directory /Users/ethanharding/Desktop/BIMM 143/bimm143_github/class16
```

```
## Info: Writing image file hsa04740.pathview.png
```

```
## Info: some node width is different from others, and hence adjusted!
```

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

```
## Info: Working in directory /Users/ethanharding/Desktop/BIMM 143/bimm143_github/class16
```

```
## Info: Writing image file hsa04640.pathview.png
```

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

```
## Info: Working in directory /Users/ethanharding/Desktop/BIMM 143/bimm143_github/class16
```

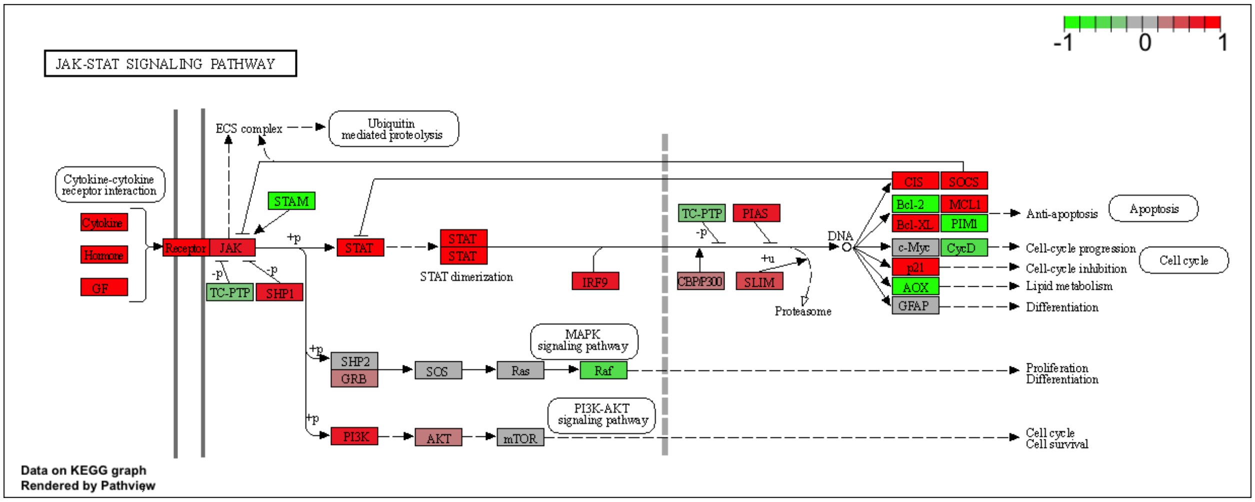
```
## Info: Writing image file hsa00140.pathview.png
```

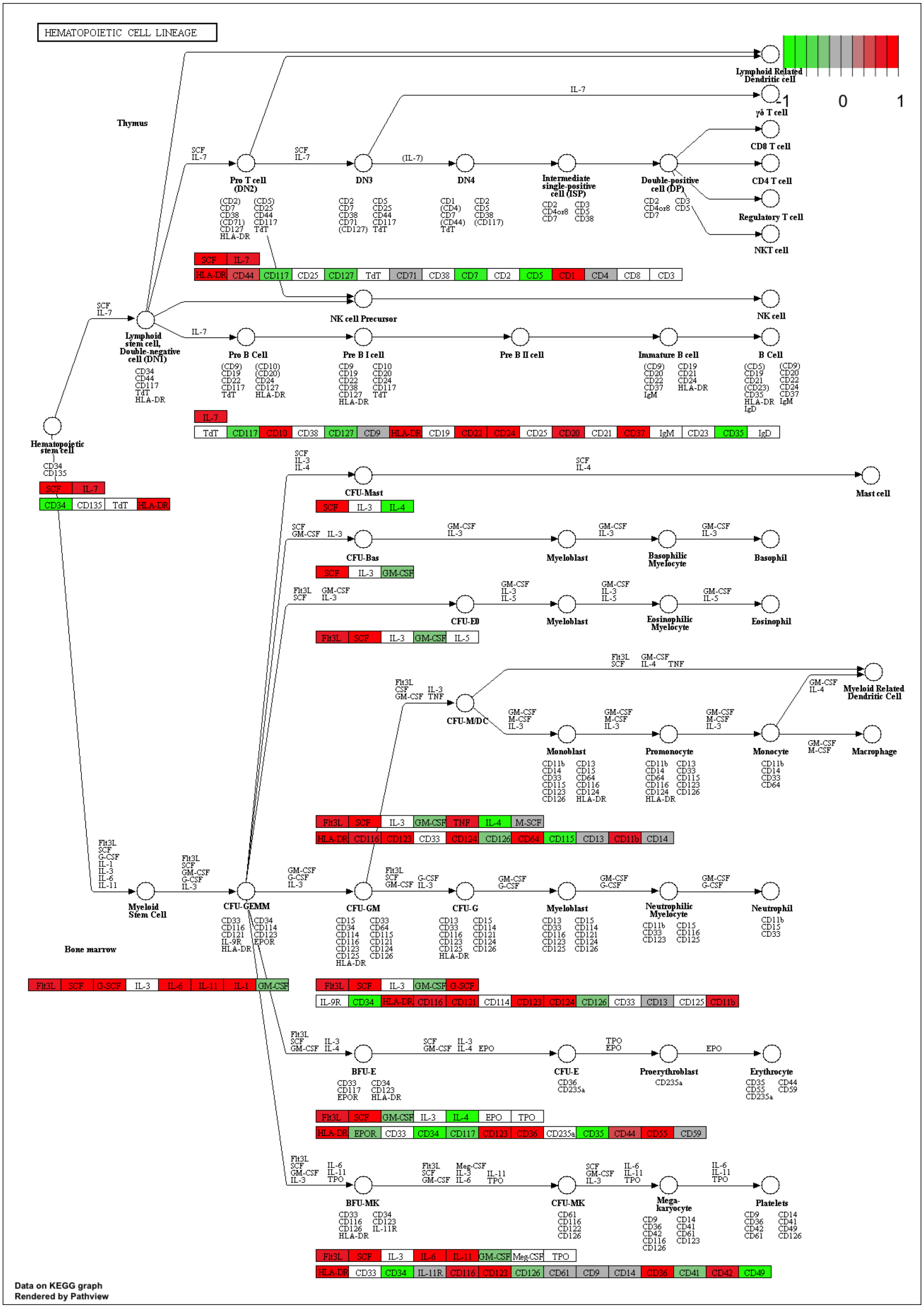
```
## ’select()’ returned 1:1 mapping between keys and columns
```

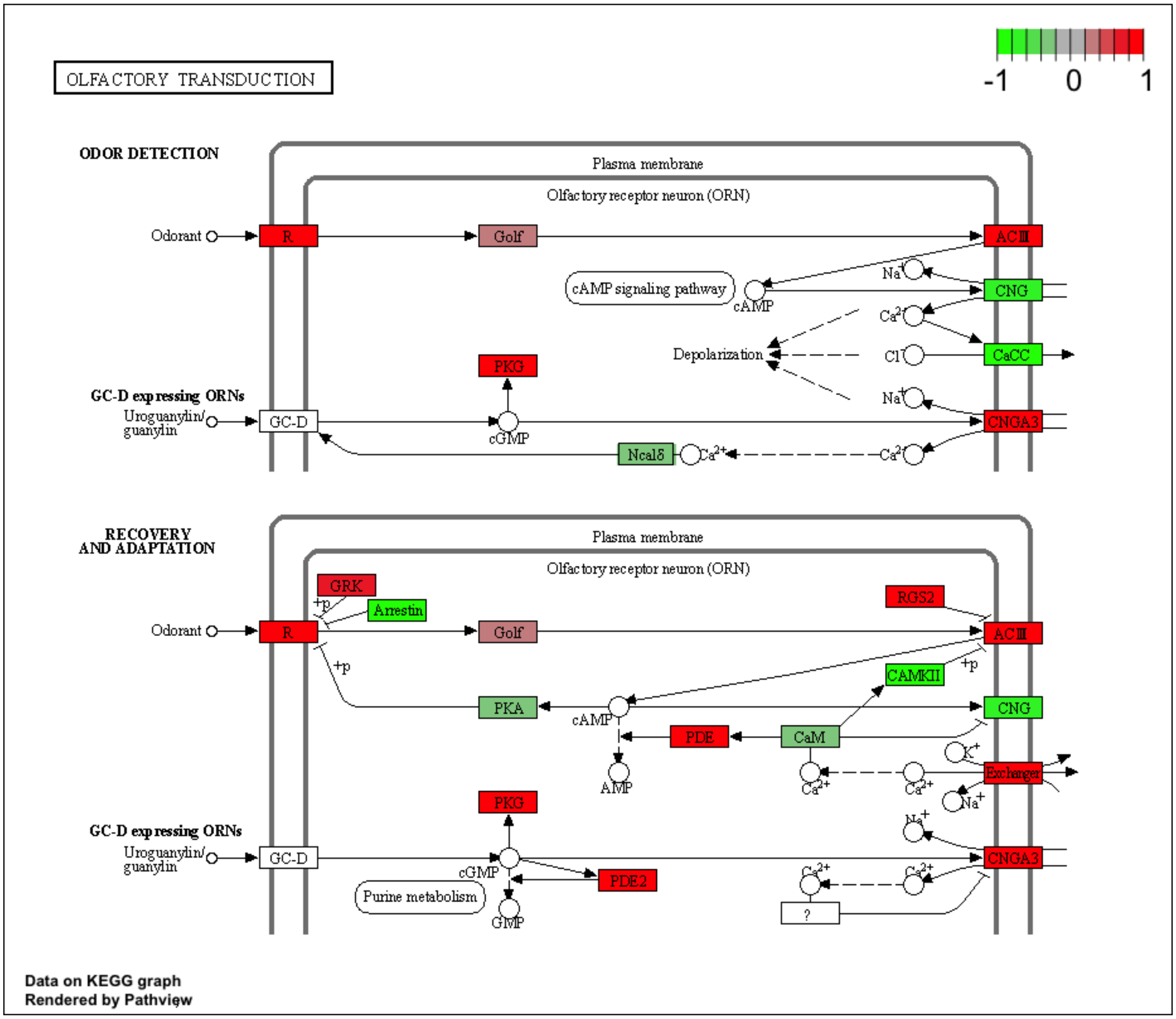
```
## Info: Working in directory /Users/ethanharding/Desktop/BIMM 143/bimm143_github/class16
```

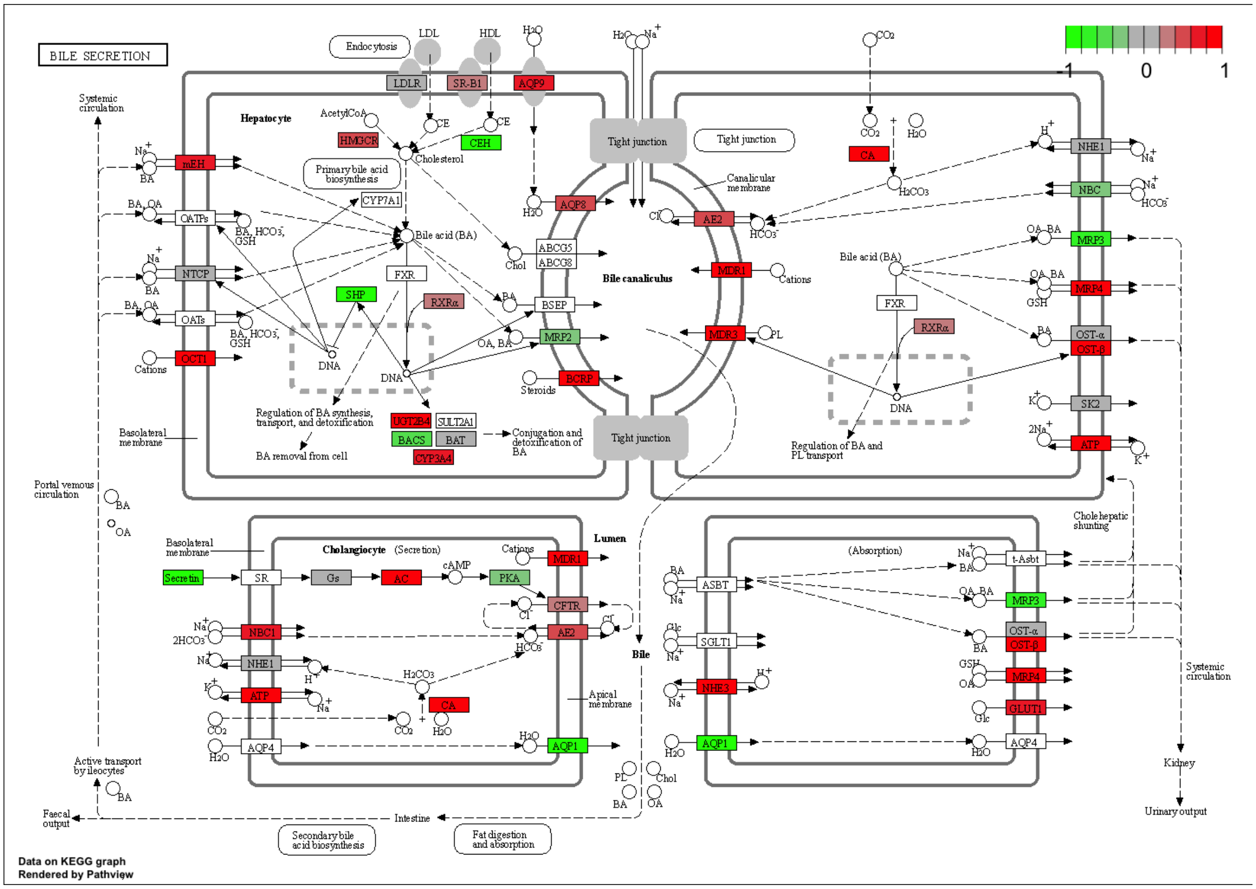
```
## Info: Writing image file hsa04630.pathview.png
```

```
## ’select()’ returned 1:1 mapping between keys and columns
## Info: Working in directory /Users/ethanharding/Desktop/BIMM 143/bimm143_github/class16

## Info: Writing image file hsa04976.pathview.png
```




HEMATOPOIETIC CELL LINEAGE
-1
0
1
Lymphoid Related Dendritic cell
IL-7
γδ T cell
Thymus
SCF IL-7
Pro T cell (DN2)
(CD2) CD7 CD38 (CD71) CD127 HLA-DR
(CD5) CD25 CD44 CD117 TdT
DN3
CD2 CD7 CD38 CD71 (CD127)
CD5 CD25 CD44 CD117 TdT
(IL-7)
DN4
CD1 (CD4) CD7 (CD44) TdT
CD2 CD5 CD38 (CD117)
Intermediate single-positive cell (ISP)
CD2 CD4or8 CD7
CD3 CD5 CD38
Double-positive cell (DP)
CD2 CD4or8 CD7
CD3 CD5
CD8 T cell
CD4 T cell
Regulatory T cell
NKT cell
SCF
IL-7
HLA-DR
CD44
CD117
CD25
CD127
TdT
CD71
CD38
CD7
CD2
CD5
CD1
CD4
CD8
CD3
NK cell Precursor
NK cell
Lymphoid stem cell, Double-negative cell (DN1)
CD34 CD44 CD117 TdT HLA-DR
IL-7
Pro B Cell
(CD9) CD19 CD22 CD117 TdT
(CD10) (CD20) CD24 CD127 HLA-DR
Pre B I cell
CD9 CD19 CD22 CD38 CD127 HLA-DR
CD10 CD20 CD24 CD117 TdT
Pre B II cell
Immature B cell
(CD9) CD20 CD22 CD37 IgM
CD19 CD21 CD24 HLA-DR
B Cell
(CD5) CD19 CD21 (CD23) CD35 HLA-DR IgD
(CD9) CD20 CD22 CD24 CD37 IgM
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
CFU-Mast
SCF IL-4
Mast cell
SCF
IL-3
IL-4
SCF GM-CSF IL-3
CFU-Bas
GM-CSF IL-3
Myeloblast
GM-CSF IL-3
Basophilic Myelocyte
GM-CSF IL-3
Basophil
SCF
IL-3
GM-CSF
Flt3L SCF
GM-CSF IL-3
CFU-E0
GM-CSF IL-3 IL-5
Myeloblast
GM-CSF IL-3 IL-5
Eosinophilic Myelocyte
GM-CSF IL-5
Eosinophil
Flt3L
SCF
IL-3
GM-CSF
IL-5
Flt3L SCF
GM-CSF IL-4 TNF
Myeloid Related Dendritic Cell
Flt3L CSF GM-CSF
IL-3 TNF
CFU-M/DC
GM-CSF M-CSF IL-3
Monoblast
CD11b CD14 CD33 CD115 CD123 CD126
CD13 CD15 CD64 CD116 CD124 HLA-DR
GM-CSF M-CSF IL-3
Promonocyte
CD11b CD14 CD64 CD116 CD124 HLA-DR
CD13 CD33 CD115 CD123 CD126
GM-CSF M-CSF IL-3
Monocyte
CD11b CD14 CD33 CD64
GM-CSF IL-4
GM-CSF M-CSF
Macrophage
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
Myeloid Stem Cell
Flt3L SCF GM-CSF G-CSF IL-3
CFU-GEMM
CD33 CD116 CD121 IL-9R HLA-DR
CD34 CD114 CD123 EPOR
GM-CSF G-CSF IL-3
CFU-GM
CD15 CD34 CD114 CD116 CD123 CD125 HLA-DR
CD33 CD64 CD115 CD121 CD124 CD126
Flt3L SCF GM-CSF
G-CSF IL-3
CFU-G
CD13 CD33 CD116 CD123 CD125 HLA-DR
CD15 CD114 CD121 CD124 CD126
GM-CSF G-CSF
Myeloblast
CD13 CD33 CD116 CD123 CD125
CD15 CD114 CD121 CD124 CD126
GM-CSF G-CSF
Neutrophilic Myelocyte
CD11b CD33 CD123
CD15 CD116 CD125
GM-CSF G-CSF
Neutrophil
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
BFU-E
CD33 CD117 EPOR
CD34 CD123 HLA-DR
SCF GM-CSF
IL-3 IL-4
EPO
CFU-E
CD36 CD235a
TPO EPO
Proerythroblast
CD235a
EPO
Erythrocyte
CD35 CD55 CD235a
CD44 CD59
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
BFU-MK
CD33 CD116 CD126 HLA-DR
CD34 CD123 IL-11R
Flt3L SCF GM-CSF
Meg-CSF IL-3 IL-6
IL-11 TPO
CFU-MK
CD61 CD116 CD122 CD126
SCF GM-CSF IL-3
IL-6 IL-11 TPO
Mega-karyocyte
CD9 CD36 CD42 CD116 CD126
CD14 CD41 CD61 CD123
IL-6 IL-11 TPO
Platelets
CD9 CD36 CD42 CD61
CD14 CD41 CD49 CD126
Flt3L
SCF
IL-3
IL-6
IL-11
GM-CSF
Meg-CSF
TPO
HLA-DR
CD33
CD34
IL-11R
CD116
CD123
CD126
CD61
CD9
CD14
CD36
CD41
CD42
CD49
Data on KEGG graph
Rendered by Pathview


OLFACTORY TRANSDUCTION
-1
0
1
ODOR DETECTION
Plasma membrane
Olfactory receptor neuron (ORN)
Odorant
R
Golf
ACIII
Na+
cAMP signaling pathway
cAMP
CNG
Ca2+
Depolarization
Cl-
CaCC
PKG
Na+
GC-D expressing ORNs
Uroguanylin/guanylin
GC-D
cGMP
CNGA3
Ncalδ
Ca2+
Ca2+
RECOVERY AND ADAPTATION
Plasma membrane
Olfactory receptor neuron (ORN)
GRK
+p
Arrestin
RGS2
Odorant
R
Golf
ACIII
+p
CAMKII
+p
PKA
cAMP
CNG
PDE
CaM
AMP
K+
Ca2+
Ca2+
Exchanger
Na+
PKG
Na+
GC-D expressing ORNs
Uroguanylin/guanylin
GC-D
cGMP
CNGA3
PDE2
Ca2+
Ca2+
Purine metabolism
GMP
?
Data on KEGG graph
Rendered by Pathview



GO

```
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        1.624062e-05  4.226117 1.624062e-05
## GO:0048729 tissue morphogenesis            5.407952e-05  3.888470 5.407952e-05
## GO:0002009 morphogenesis of an epithelium  5.727599e-05  3.878706 5.727599e-05
## GO:0030855 epithelial cell differentiation 2.053700e-04  3.554776 2.053700e-04
## GO:0060562 epithelial tube morphogenesis   2.927804e-04  3.458463 2.927804e-04
## GO:0048598 embryonic morphogenesis         2.959270e-04  3.446527 2.959270e-04
##                                                 q.val set.size         exp1
## GO:0007156 homophilic cell adhesion        0.07103646      138 1.624062e-05
## GO:0048729 tissue morphogenesis            0.08350839      483 5.407952e-05
## GO:0002009 morphogenesis of an epithelium  0.08350839      382 5.727599e-05
## GO:0030855 epithelial cell differentiation 0.14646752      299 2.053700e-04
## GO:0060562 epithelial tube morphogenesis   0.14646752      289 2.927804e-04
## GO:0048598 embryonic morphogenesis         0.14646752      498 2.959270e-04
##
## $less
##                                             p.geomean stat.mean        p.val
## GO:0048285 organelle fission             6.386337e-16 -8.175381 6.386337e-16
## GO:0000280 nuclear division              1.726380e-15 -8.056666 1.726380e-15
## GO:0007067 mitosis                       1.726380e-15 -8.056666 1.726380e-15
## GO:0000087 M phase of mitotic cell cycle 4.593581e-15 -7.919909 4.593581e-15
## GO:0007059 chromosome segregation        9.576332e-12 -6.994852 9.576332e-12
## GO:0051301 cell division                 8.718528e-11 -6.455491 8.718528e-11
##                                                 q.val set.size         exp1
## GO:0048285 organelle fission             2.517062e-12      386 6.386337e-16
## GO:0000280 nuclear division              2.517062e-12      362 1.726380e-15
## GO:0007067 mitosis                       2.517062e-12      362 1.726380e-15
## GO:0000087 M phase of mitotic cell cycle 5.023080e-12      373 4.593581e-15
## GO:0007059 chromosome segregation        8.377375e-09      146 9.576332e-12
## GO:0051301 cell division                 6.355807e-08      479 8.718528e-11
##
## $stats
##                                            stat.mean     exp1
## GO:0007156 homophilic cell adhesion         4.226117 4.226117
## GO:0048729 tissue morphogenesis             3.888470 3.888470
## GO:0002009 morphogenesis of an epithelium   3.878706 3.878706
## GO:0030855 epithelial cell differentiation  3.554776 3.554776
## GO:0060562 epithelial tube morphogenesis    3.458463 3.458463
## GO:0048598 embryonic morphogenesis          3.446527 3.446527
```