Bioinformatics Lab4

Milda Poceviciute, Fanny Karelius, Rab Nawaz Jan Sher and Saman Zahid

6 December 2018

```
library(GEOquery)
library(simpleaffy)
library(RColorBrewer)
library(affyPLM)
library(limma)
library(hgu133plus2.db)
library(annotate)
library(ggplot2)
```

Question 1

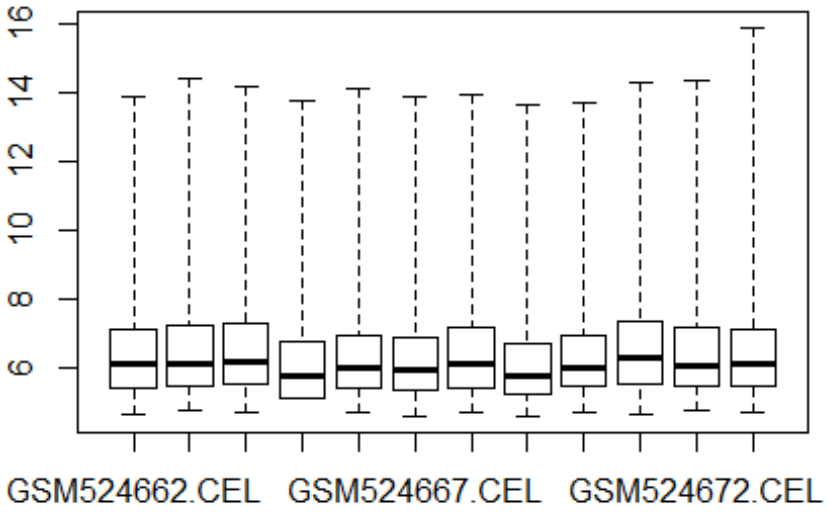
At first they import the data file of the Comparative Gene Expression Profiling of HUVEC and Ocular Vascular Endothelial Cells (Homo Sapiens) as a dataframe *phenodata*. They also include which target variables for which data set they will attempt to model.

```
## ---- cache=TRUE-------------------------------------------------------------
x = getGEOSuppFiles("GSE20986")
x
## ----------------------------------------------------------------------------
untar("GSE20986/GSE20986_RAW.tar", exdir = "data")

## ----------------------------------------------------------------------------
cels = list.files("data/", pattern = "[gz]")
sapply(paste("data", cels, sep = "/"), gunzip)

## ----------------------------------------------------------------------------
phenodata = matrix(rep(list.files("data"), 2), ncol =2)
class(phenodata)
phenodata <- as.data.frame(phenodata)
colnames(phenodata) <- c("Name", "FileName")
phenodata$Targets <- c("iris",
                       "retina",
                       "retina",
                       "iris",
                       "retina",
                       "iris",
                       "choroid",
                       "choroid",
                       "choroid",
                       "huvec",
                       "huvec",
                       "huvec")
write.table(phenodata, "data/phenodata.txt", quote = F, sep = "\t", row.names = F)
```

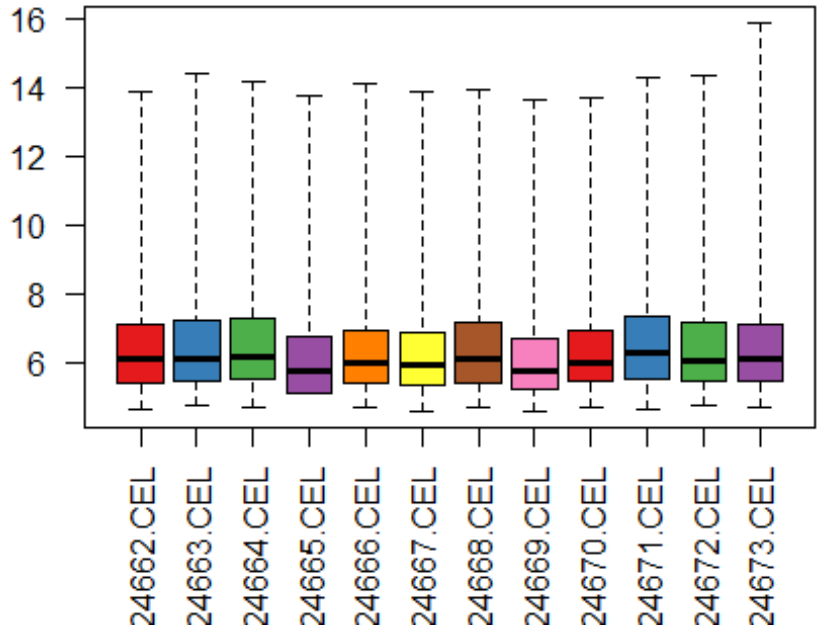
Then the data is read again as AffyBatch object into variable *celfiles*, the data comes from a MicroArray of a company Affymetrix. Afterwards the box plots are produced which show the statistics of the 12 samples (minimum, first quantile, median, third quantile, and maximum). One of them is coloured by the samples.

```
library(simpleaffy)
library(affy)
```

```
celfiles <- read.affy(covdesc = "phenodata.txt", path = "data")
boxplot(celfiles)

## Warning: replacing previous import 'AnnotationDbi::tail' by 'utils::tail'
## when loading 'hgu133plus2cdf'

## Warning: replacing previous import 'AnnotationDbi::head' by 'utils::head'
## when loading 'hgu133plus2cdf'

##
```

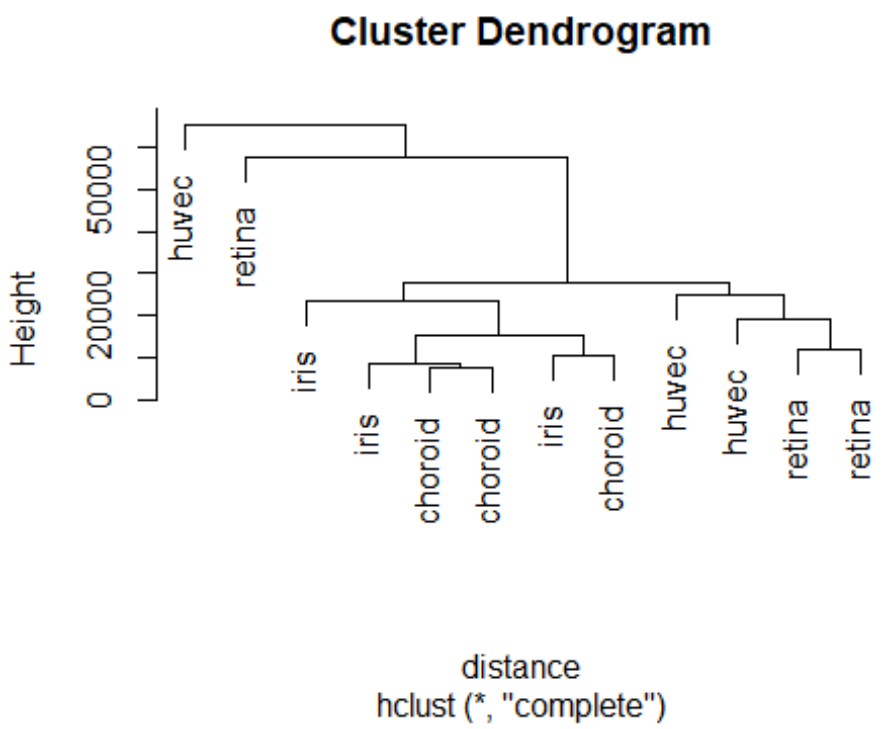



```
library(RColorBrewer)
cols = brewer.pal(8, "Set1")
eset <- affy::exprs(celfiles)
samples <- celfiles$Targets
colnames(eset)

##  [1] "GSM524662.CEL" "GSM524663.CEL" "GSM524664.CEL" "GSM524665.CEL"
##  [5] "GSM524666.CEL" "GSM524667.CEL" "GSM524668.CEL" "GSM524669.CEL"
##  [9] "GSM524670.CEL" "GSM524671.CEL" "GSM524672.CEL" "GSM524673.CEL"

colnames(eset) <- samples

boxplot(celfiles, col = cols, las = 2)
```



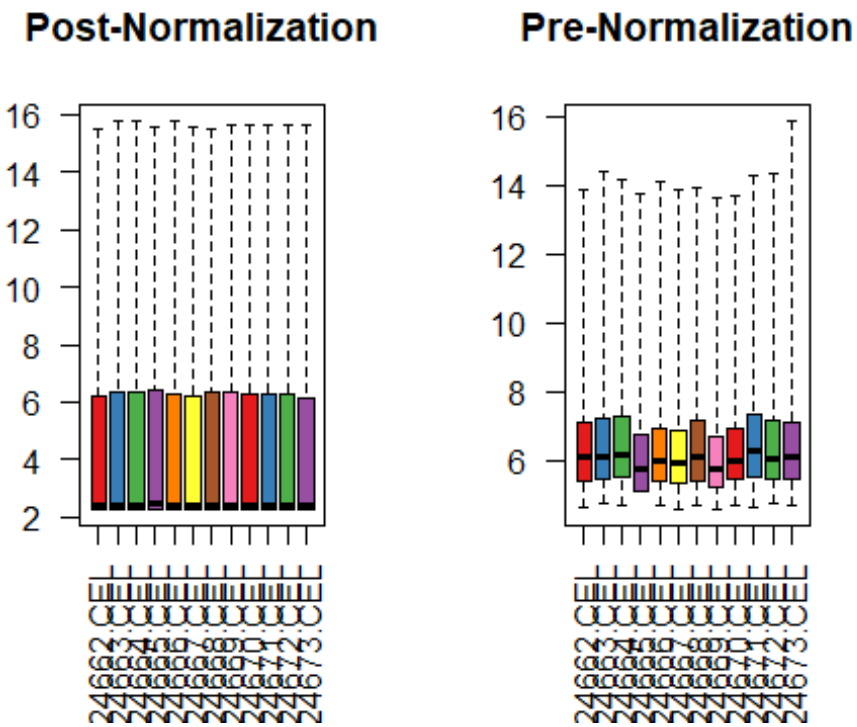
Then the distance matrix is created for the genome data from 12 samples, and the *hclus* function is used to create cluster dendogram of the samples:

```
distance <- dist(t(eset), method = "maximum")
clusters <- hclust(distance)
plot(clusters)
```



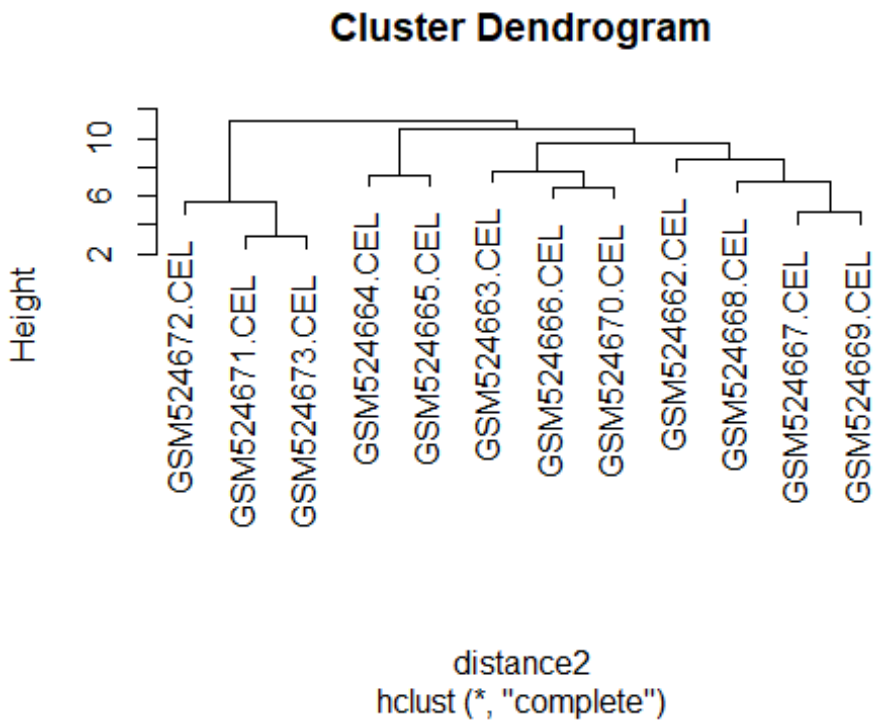
The genomic data is normalised and converted into ExpressionSet format, the bar plots of normalised and non-normalised data are plotted for comparison purposes:

```
library(simpleaffy)
library(affyPLM)
celfiles.gcrma = gcrma(celfiles)

## Adjusting for optical effect.....Done.
## Computing affinities.Done.
## Adjusting for non-specific binding.....Done.
## Normalizing
## Calculating Expression

par(mfrow=c(1,2))
boxplot(celfiles.gcrma, col = cols, las = 2, main = "Post-Normalization");
boxplot(celfiles, col = cols, las = 2, main = "Pre-Normalization")
```



And then the clustering is performed on a normalize data set:

```
distance2 <- dist(t(affy::exprs(celfiles.gcrma)), method = "maximum")
clusters2 <- hclust(distance2)
plot(clusters2)
```



Here the design matrix is created which indicates what explanatory variables are present in each of the 12 samples. This matrix is used to fit the linear regression model. Also a contrast matrix is created. It has three pairs of the explanatory variables, and each row indicates whether a single explanatory variable is present or not in each of the pairs. This is used in fitting a microarray linear model. The result shows the resulting statistics for the top 100000 genes for huvec_choroid variable.

```
library(limma)


samples <- as.factor(samples)
design <- model.matrix(~0+samples)
colnames(design)

## [1] "sampleschoroid" "sampleshuvec"   "samplesiris"    "samplesretina"

colnames(design) <- c("choroid", "huvec", "iris", "retina")
design

##    choroid huvec iris retina
## 1        0     0    1      0
## 2        0     0    0      1
## 3        0     0    0      1
## 4        0     0    1      0
## 5        0     0    0      1
## 6        0     0    1      0
## 7        1     0    0      0
## 8        1     0    0      0
## 9        1     0    0      0
## 10       0     1    0      0
## 11       0     1    0      0
## 12       0     1    0      0
## attr(,"assign")
```

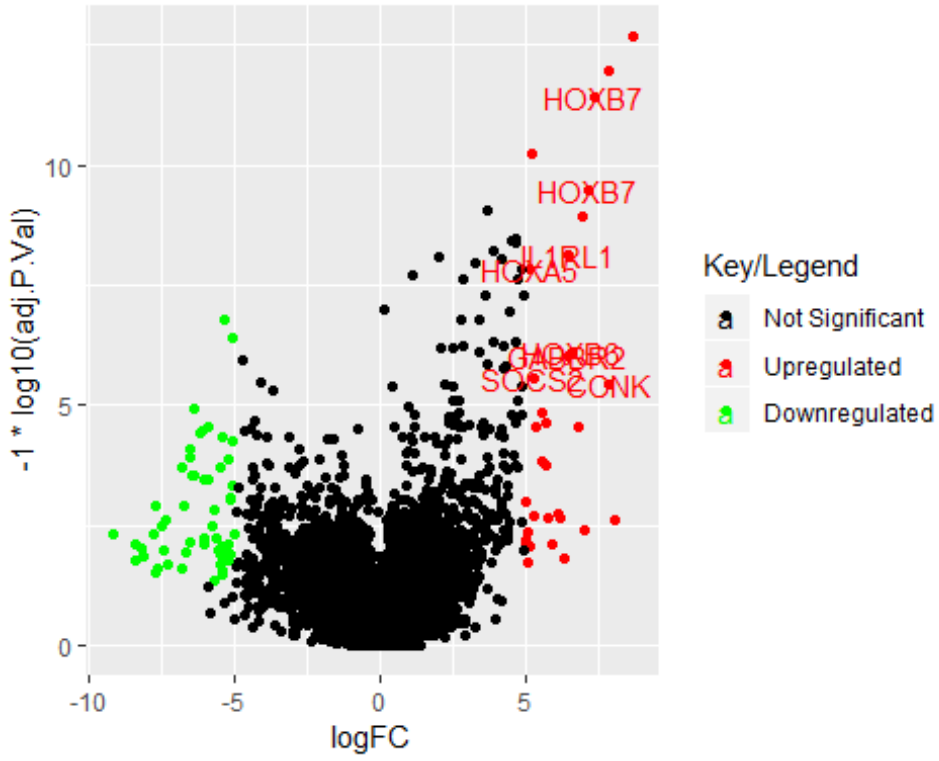
```
## [1] 1 1 1 1
## attr(,"contrasts")
## attr(,"contrasts")$samples
## [1] "contr.treatment"

contrast.matrix = makeContrasts(
    huvec_choroid = huvec - choroid,
    huvec_retina = huvec - retina,
    huvec_iris = huvec - iris,
    levels = design)

fit = lmFit(celfiles.gcrma, design)
huvec_fit <- contrasts.fit(fit, contrast.matrix)
huvec_ebay <- eBayes(huvec_fit)

library(hgu133plus2.db)
library(annotate)
probenames.list <- rownames(topTable(huvec_ebay, number = 100000))
getsymbols <- getSYMBOL(probenames.list, "hgu133plus2")
results <- topTable(huvec_ebay, number = 100000, coef = "huvec_choroid")
results <- cbind(results, getsymbols)
```

Here the summary of the results are printed out, and the genes are grouped into three groups (a new variable "threshold" is introduced for that). Threshold is 1 if p-value is above $\alpha = 0.05$ significant level, the genes that have p-values below $\alpha = 0.05$ and logFC >5, and those that have p-values below $\alpha = 0.05$, and logFC >-5. logFC is a log2-fold-change corresponding to a contrast.

```
summary(results)

##      logFC             AveExpr             t                P.Value
##  Min.   :-9.19111   Min.   : 2.279   Min.   :-39.77473   Min.   :0.0000
##  1st Qu.:-0.05967   1st Qu.: 2.281   1st Qu.: -0.70649   1st Qu.:0.1523
##  Median : 0.00000   Median : 2.480   Median :  0.00000   Median :0.5079
##  Mean   :-0.02353   Mean   : 4.375   Mean   :  0.07441   Mean   :0.5346
##  3rd Qu.: 0.03986   3rd Qu.: 6.241   3rd Qu.:  0.67455   3rd Qu.:1.0000
##  Max.   : 8.67086   Max.   :15.541   Max.   :296.84201   Max.   :1.0000
##
##    adj.P.Val            B             getsymbols
##  Min.   :0.0000   Min.   :-7.710   YME1L1  :   22
##  1st Qu.:0.6036   1st Qu.:-7.710   HFE     :   15
##  Median :1.0000   Median :-7.451   CFLAR   :   14
##  Mean   :0.7436   Mean   :-6.582   NRP2    :   14
##  3rd Qu.:1.0000   3rd Qu.:-6.498   ARHGEF12:   13
##  Max.   :1.0000   Max.   :21.290   (Other) :41857
##                                    NA's    :12740

results$threshold <- "1"
a <- subset(results, adj.P.Val < 0.05 & logFC > 5)
results[rownames(a), "threshold"] <- "2"
b <- subset(results, adj.P.Val < 0.05 & logFC < -5)
results[rownames(b), "threshold"] <- "3"
table(results$threshold)

##
##     1     2     3
## 54587    33    55
```

The plot below illustrates the results. The down-regulated genes are created by the process when a cell decreases the quantity of a cellular component (such as RNA or protein) in response to an external stimulus. The up-regulated genes are created by a reversed process.

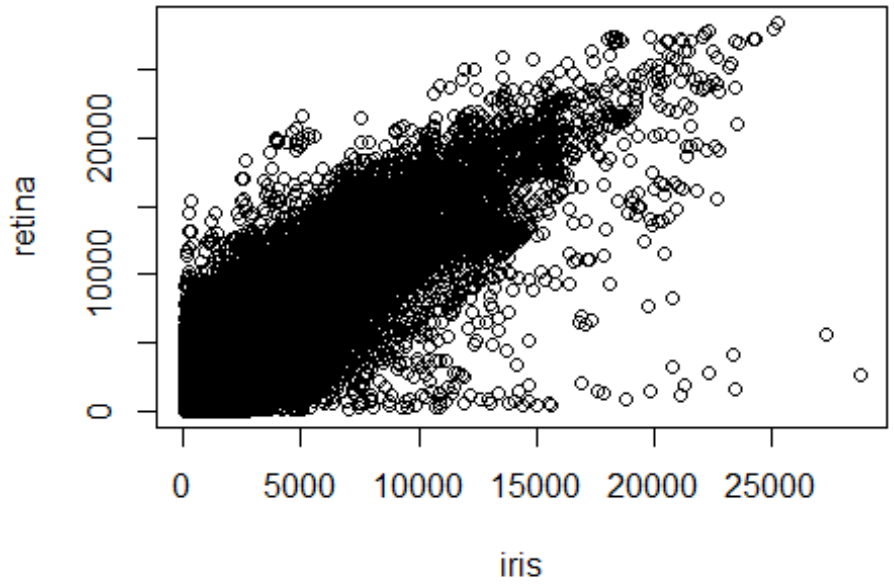
```
library(ggplot2)
volcano <- ggplot(data = results,
                  aes(x = logFC, y = -1*log10(adj.P.Val),
                      colour = threshold,
                      label = getsymbols))

volcano <- volcano +
    geom_point() +
    scale_color_manual(values = c("black", "red", "green"),
                       labels = c("Not Significant", "Upregulated", "Downregulated"),
                       name = "Key/Legend")

volcano +
    geom_text(data = subset(results, logFC > 5 & -1*log10(adj.P.Val) > 5), aes(x = logFC, y = -
1*log10(adj.P.Val), colour = threshold, label = getsymbols)  )

## Warning: Removed 4 rows containing missing values (geom_text).
```



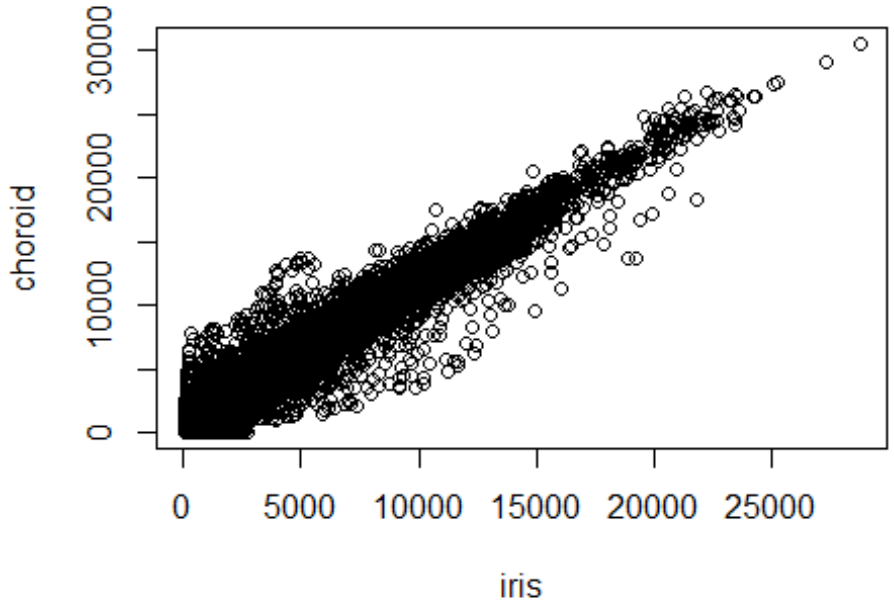
Question 2

Contrast matrices for the remaining pairs (DO WE NEED TO DO THIS???)

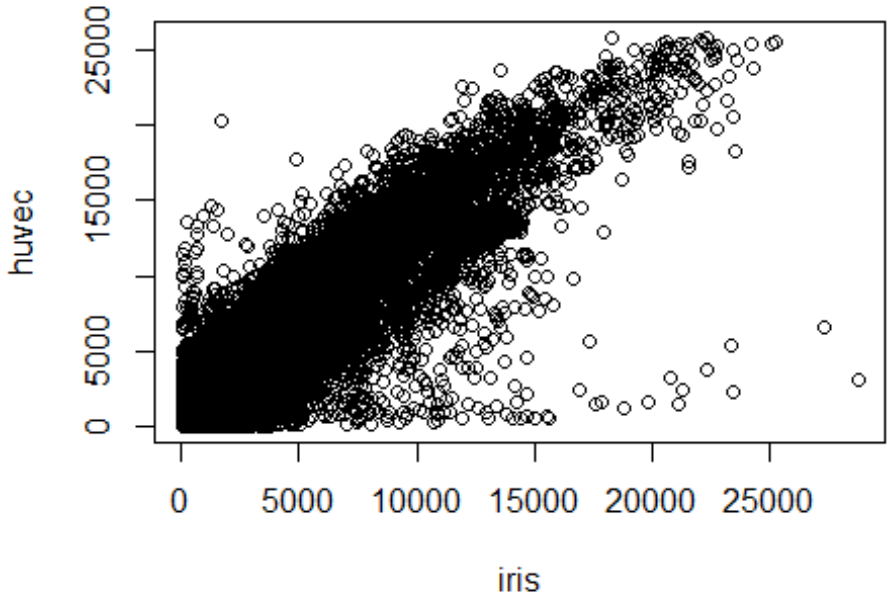
```
contrast.matrix = makeContrasts(
    huvec_choroid = huvec - choroid,
    huvec_retina = huvec - retina,
    huvec_iris = huvec - iris,
    levels = design)

# Not normalised data
data <- eset[,c(1,2,7,11)]
colnames(data) <- c("iris","retina","choroid","huvec")
plot(x=data[,1],y=data[,2],xlab="iris",ylab="retina")
```



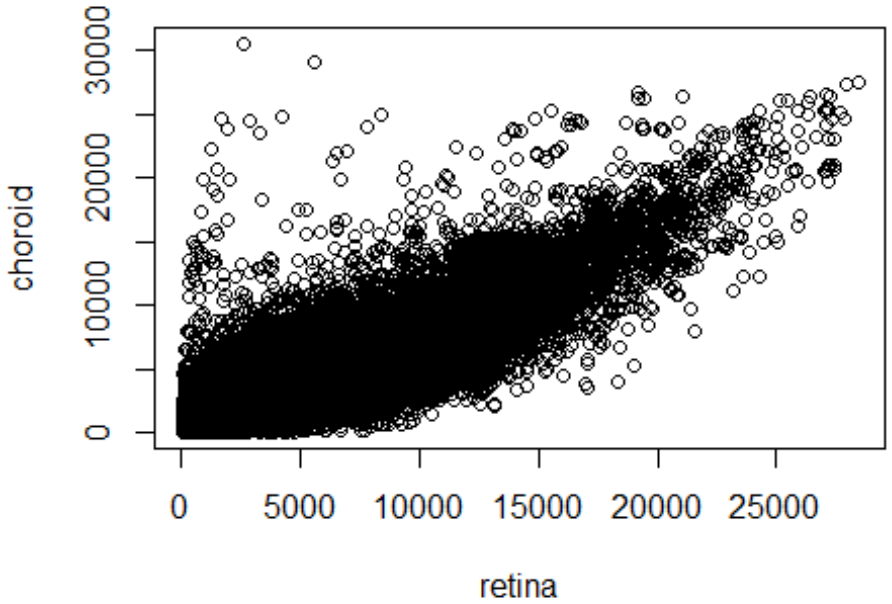

```r
plot(x=data[,1],y=data[,3],xlab="iris",ylab="choroid")
```



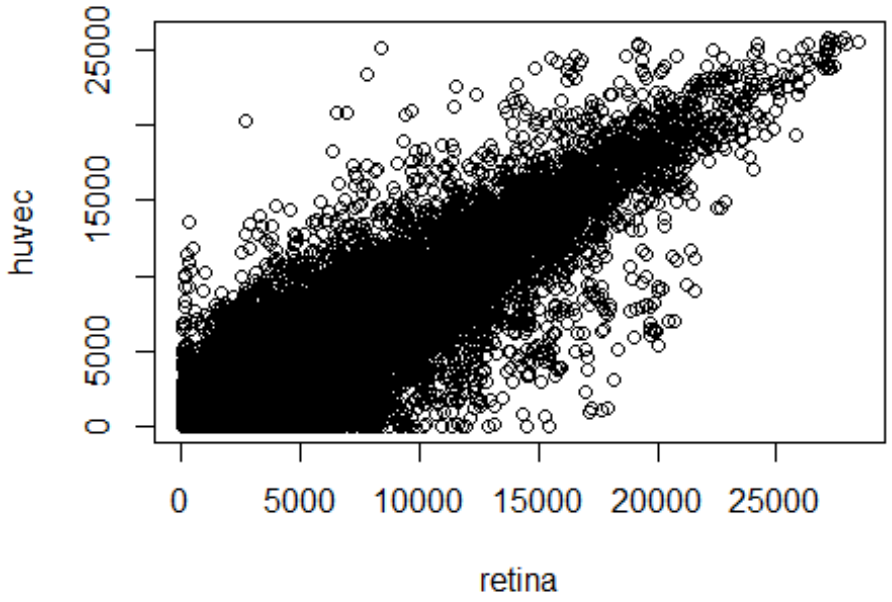

```r
plot(x=data[,1],y=data[,4],xlab="iris",ylab="huvec")
```



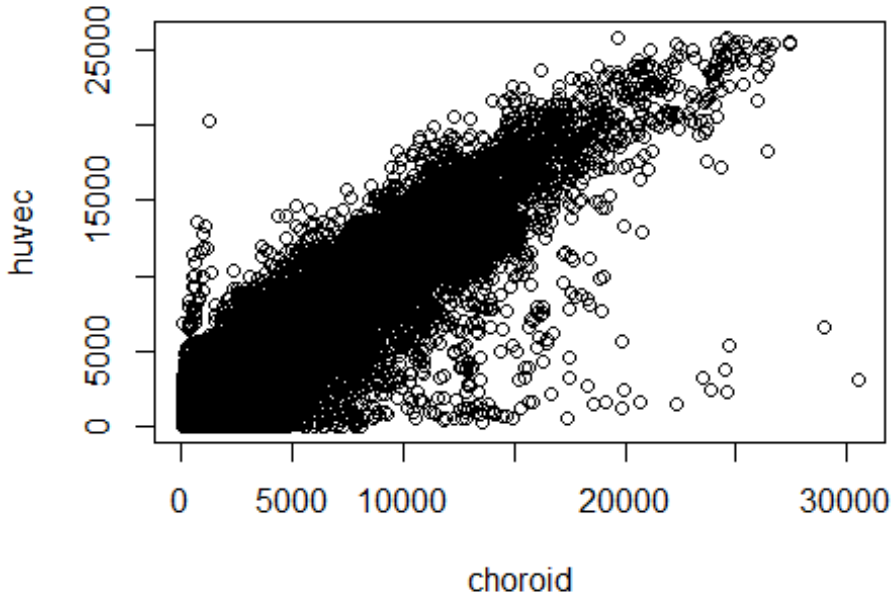

```
plot(x=data[,2],y=data[,3],xlab="retina",ylab="choroid")
```




```
plot(x=data[,2],y=data[,4],xlab="retina",ylab="huvec")
```

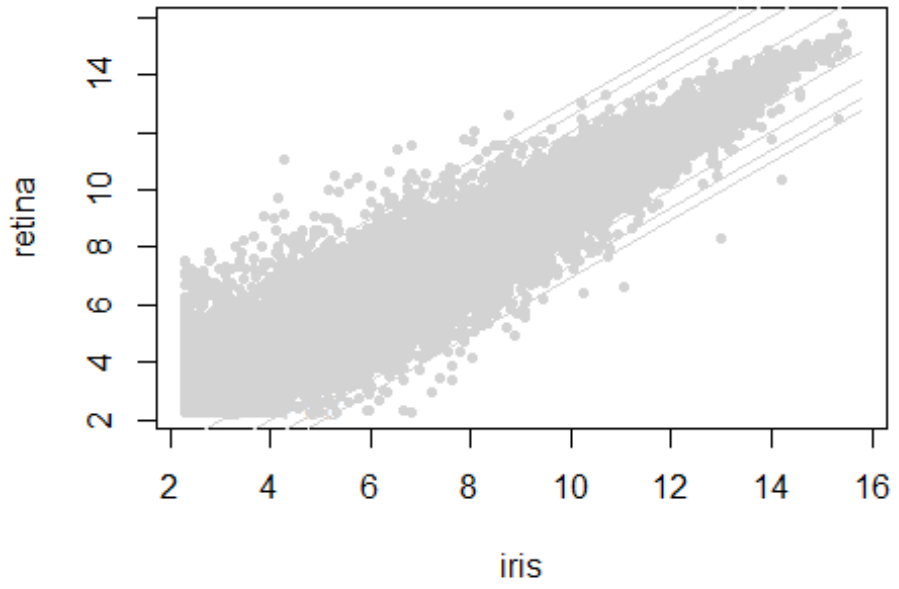



```
plot(x=data[,3],y=data[,4],xlab="choroid",ylab="huvec")
```

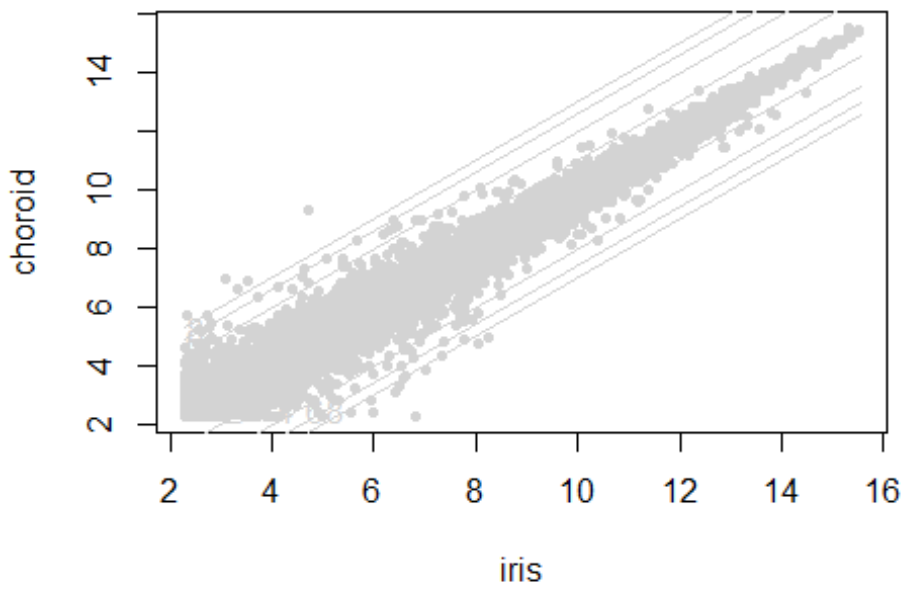



```
# Normalised data
comparison1 <- pairwise.comparison(celfiles.gcrma,"Targets",c("iris","retina"))
```

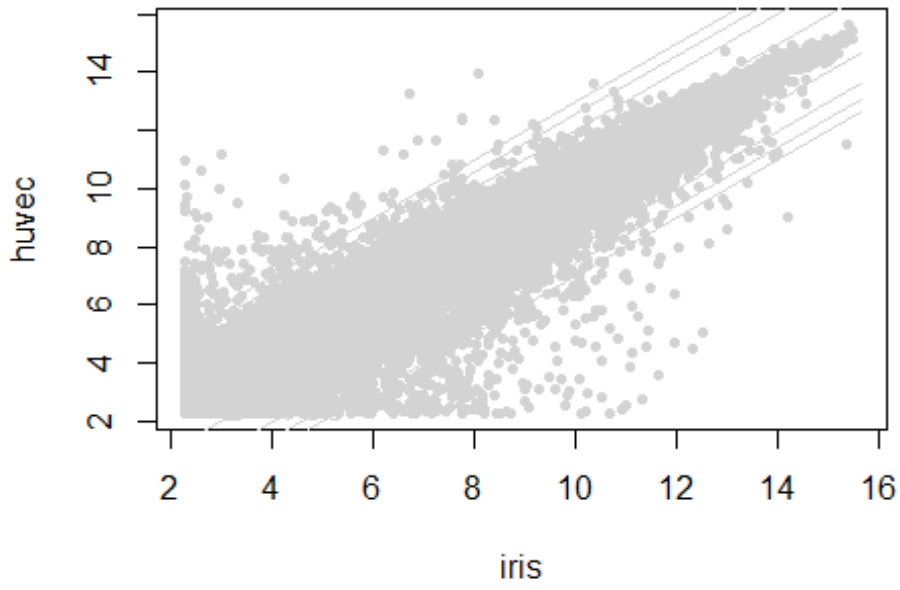
```
comparison2 <- pairwise.comparison(celfiles.gcrma,"Targets",c("iris","choroid"))
comparison3 <- pairwise.comparison(celfiles.gcrma,"Targets",c("iris","huvec"))
comparison4 <- pairwise.comparison(celfiles.gcrma,"Targets",c("retina","huvec"))
comparison5 <- pairwise.comparison(celfiles.gcrma,"Targets",c("retina","choroid"))
comparison6 <- pairwise.comparison(celfiles.gcrma,"Targets",c("choroid","huvec"))

plot(comparison1)
```



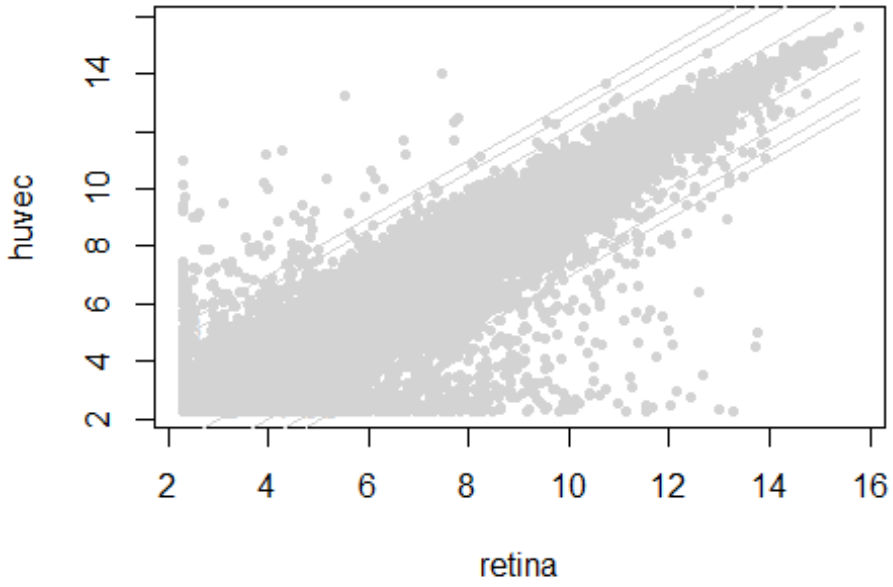

```
plot(comparison2)
```



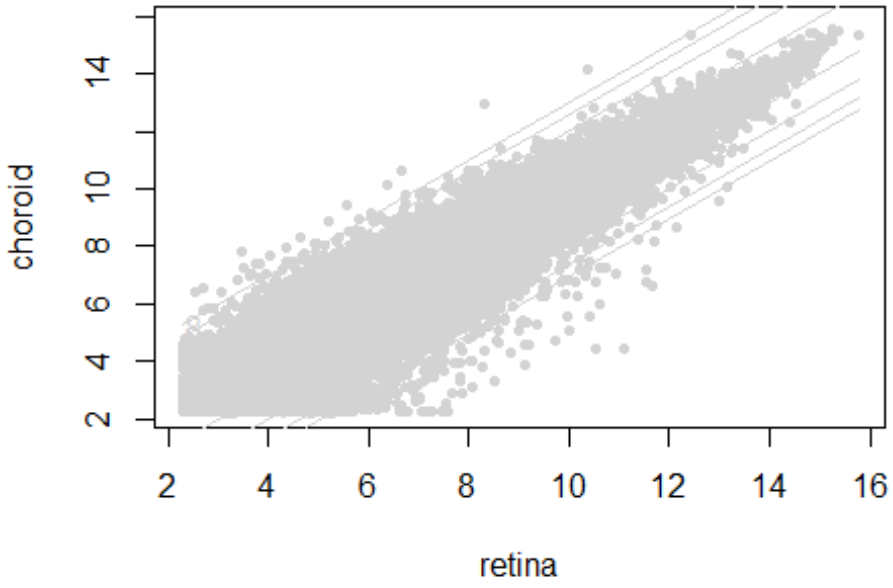

```
plot(comparison3)
```



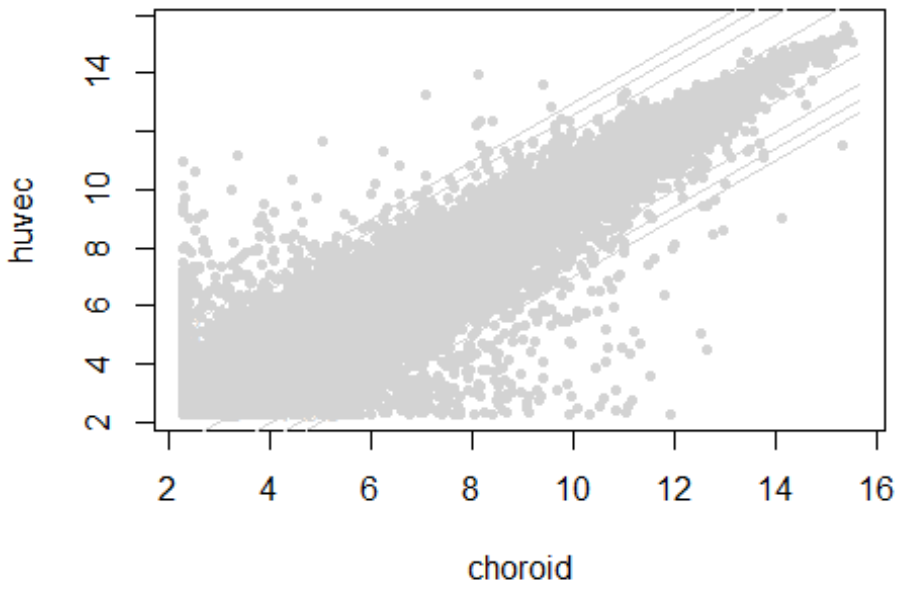

```
plot(comparison4)
```




```
plot(comparison5)
```

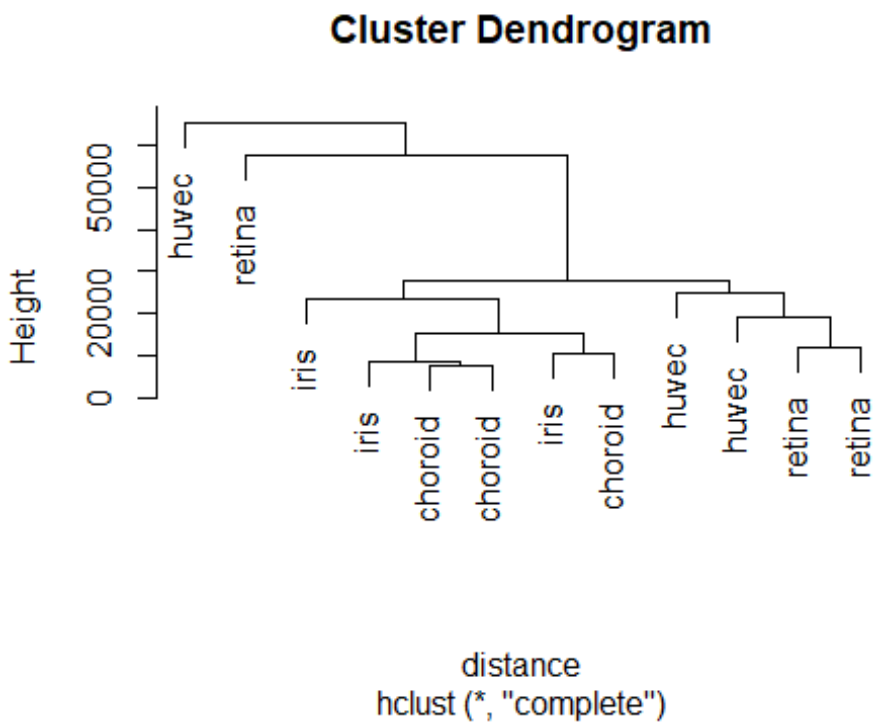



```
plot(comparison6)
```

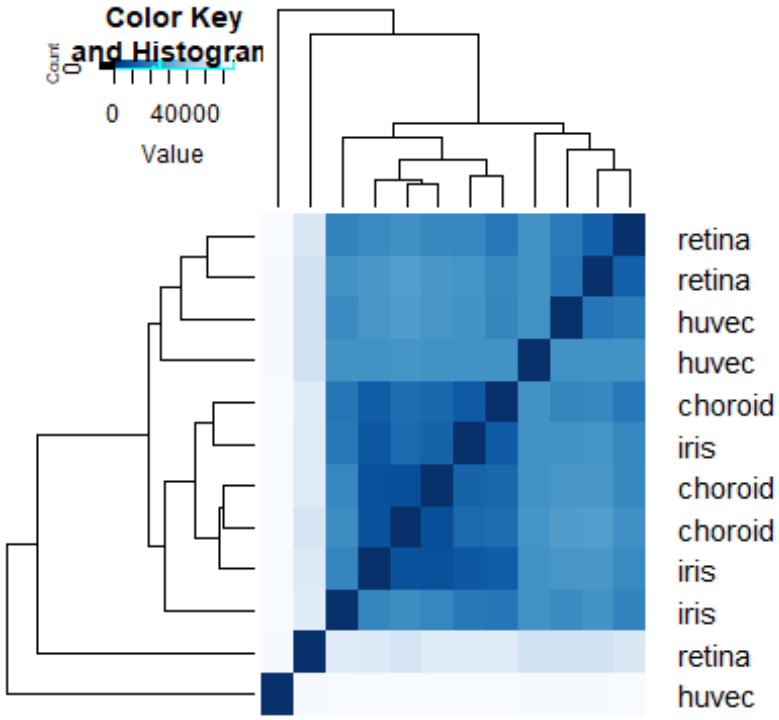


The dendograms and heatmaps are produced for both: raw data cluster analysis and the normalised data cluster analysis.

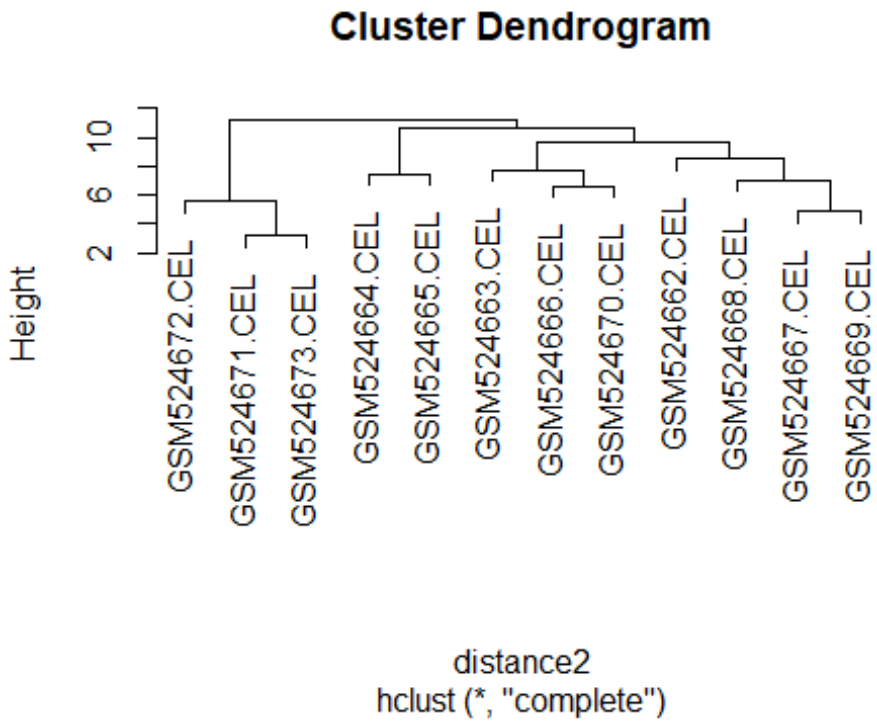
```
library("RColorBrewer")
library("gplots")

##
## Attaching package: 'gplots'

## The following object is masked from 'package:IRanges':
##
##     space

## The following object is masked from 'package:S4Vectors':
##
##     space

## The following object is masked from 'package:stats':
##
##     lowess

plot(clusters)
```




```
colors <- colorRampPalette( rev(brewer.pal(9, "Blues")) )(255)
dist_mat <- as.matrix(distance)
# Heatmap of the raw data
heatmap.2( dist_mat, Rowv=as.dendrogram(clusters),
          symm=TRUE, trace="none", col=colors,
          margins=c(2,10), labCol=FALSE )
```

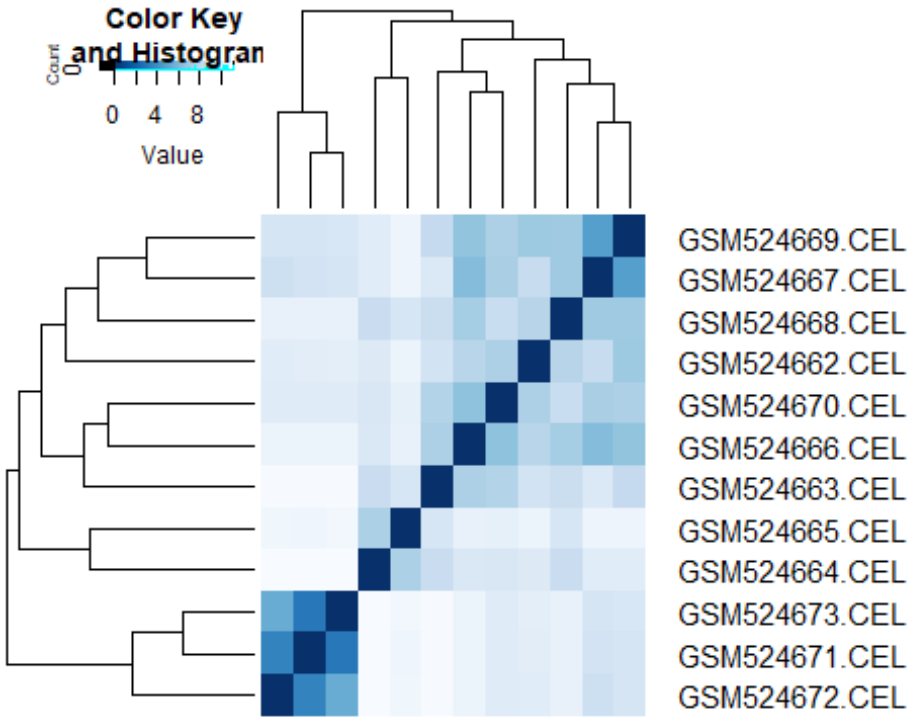



```
plot(clusters2)
```



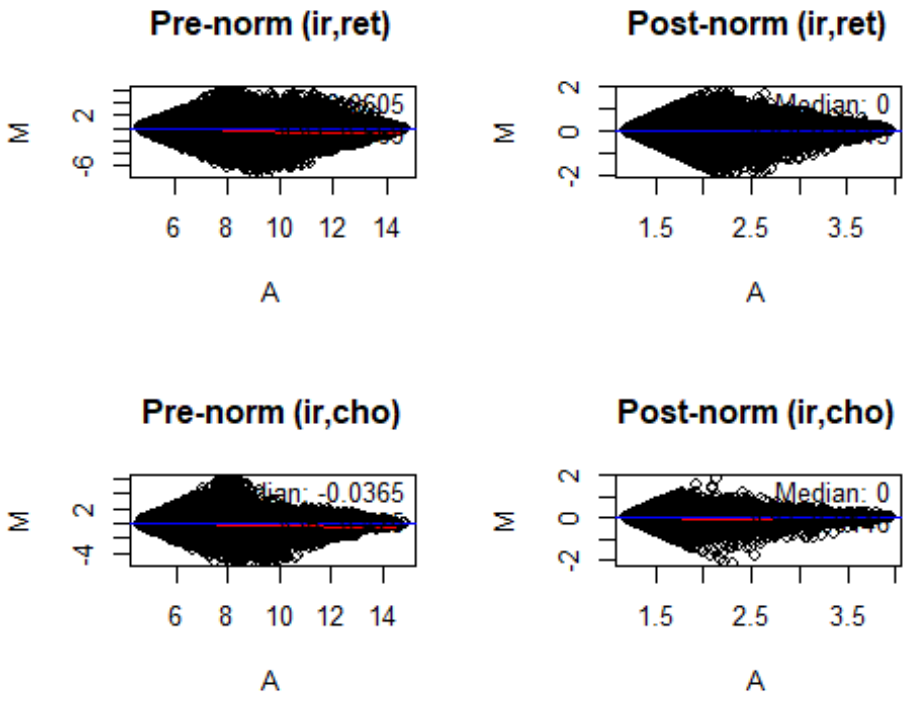

```
# Heatmap of the normalised data
dist2_mat <- as.matrix(distance2)
```

```
heatmap.2( dist2_mat, Rowv=as.dendrogram(clusters2),
          symm=TRUE, trace="none", col=colors,
          margins=c(2,10), labCol=FALSE )
```

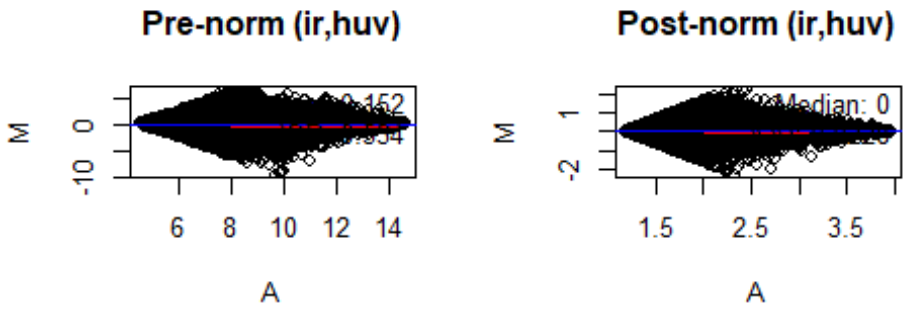


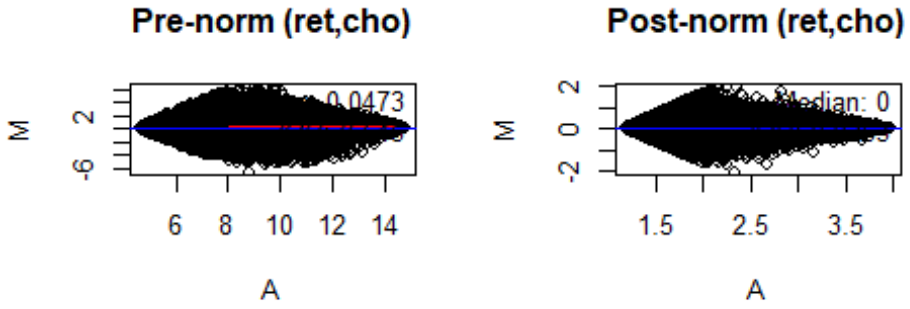
MA-plots were made for each pair both before and after normalization.

```
library(preprocessCore)
#1=iris, 2=retina
#7=choroid, huvec=10
ir <- which(colnames(eset)=="iris")[1]
ret <- which(colnames(eset)=="retina")[1]
cho <- which(colnames(eset)=="choroid")[1]
huv <- which(colnames(eset)=="huvec")[1]
log2eset <- log2(eset)
# Normalization
eset_norm<- exprs(celfiles.gcrma)
colnames(eset_norm) <- samples
log2eset_norm <- log2(eset_norm)

par(mfrow=c(2,2))
#ir, ret
ma.plot(rowMeans(log2eset[,c(ir,ret)]), log2eset[,ir]-log2eset[,ret], main="Pre-norm (ir,ret)",
cex=1)
ma.plot(rowMeans(log2eset_norm[,c(ir, ret)]), log2eset_norm[,ir]-log2eset_norm[,ret],
main="Post-norm (ir,ret)",cex=1)

#ir, cho
ma.plot(rowMeans(log2eset[,c(ir,cho)]), log2eset[,ir]-log2eset[,cho], main="Pre-norm (ir,cho)",
cex=1)
ma.plot(rowMeans(log2eset_norm[,c(ir, cho)]), log2eset_norm[,ir]-log2eset_norm[,cho],
main="Post-norm (ir,cho)",cex=1)
```

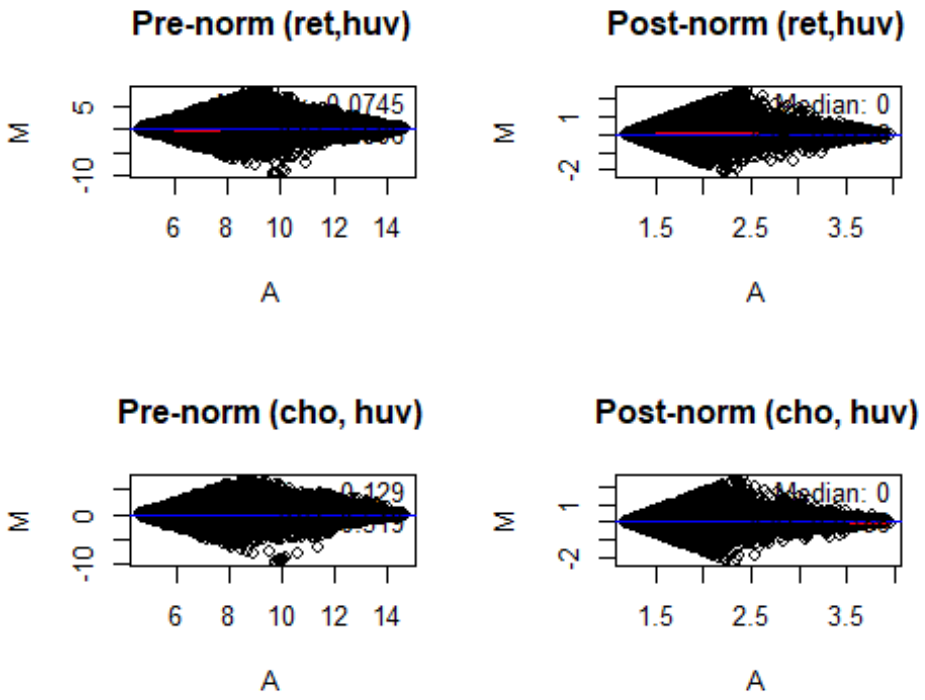



```
#ir, huv
ma.plot(rowMeans(log2eset[,c(ir,huv)]), log2eset[,ir]-log2eset[,huv], main="Pre-norm (ir,huv)",
cex=1)
ma.plot(rowMeans(log2eset_norm[,c(ir, huv)]), log2eset_norm[,ir]-log2eset_norm[,huv],
main="Post-norm (ir,huv)",cex=1)

#ret, cho
ma.plot(rowMeans(log2eset[,c(ret,cho)]), log2eset[,ret]-log2eset[,cho], main="Pre-norm
(ret,cho)", cex=1)
ma.plot(rowMeans(log2eset_norm[,c(ret, cho)]), log2eset_norm[,ret]-log2eset_norm[,cho],
main="Post-norm (ret,cho)",cex=1)
```






```
#ret, huv
ma.plot(rowMeans(log2eset[,c(ret,huv)]), log2eset[,ret]-log2eset[,huv], main="Pre-norm
(ret,huv)", cex=1)
ma.plot(rowMeans(log2eset_norm[,c(ret, huv)]), log2eset_norm[,ret]-log2eset_norm[,huv],
main="Post-norm (ret,huv)",cex=1)

#cho, huv
ma.plot(rowMeans(log2eset[,c(cho,huv)]), log2eset[,cho]-log2eset[,huv], main="Pre-norm (cho,
huv)", cex=1)
ma.plot(rowMeans(log2eset_norm[,c(cho, huv)]), log2eset_norm[,cho]-log2eset_norm[,huv],
main="Post-norm (cho, huv)",cex=1)
```



QUestion 3

The genes are grouped into three groups that contains genes that have: * p-value above $\alpha = 0.05$ significant level (insignificant) * p-values below $\alpha = 0.05$ and logFC >5 (Up-regulated) * p-values below $\alpha = 0.05$, and logFC >-5 (Down-regulated)

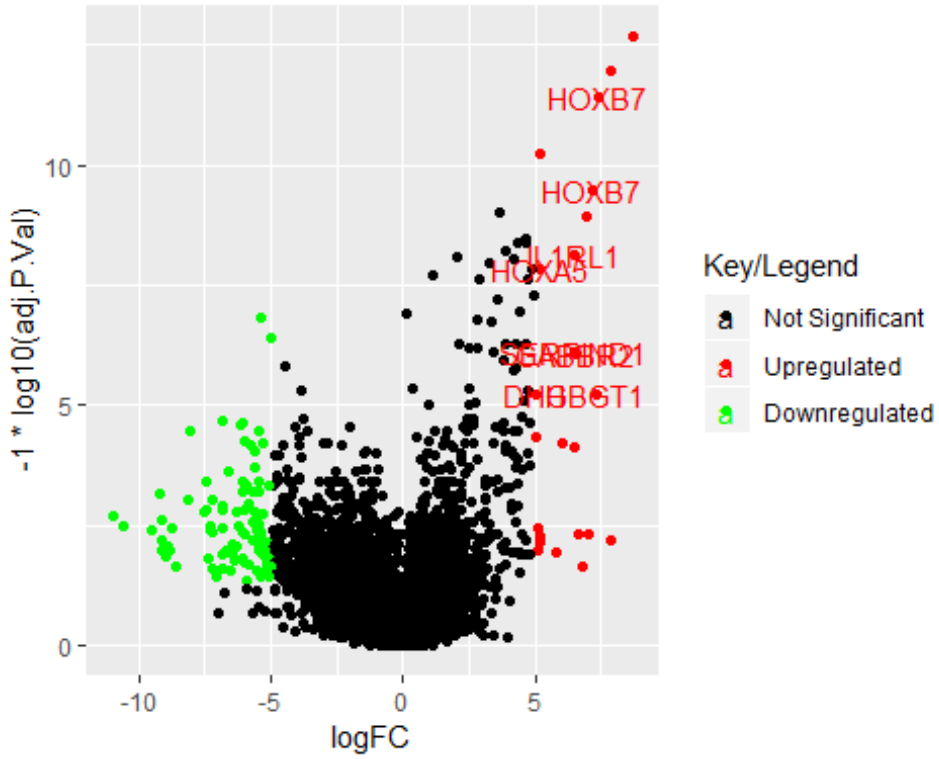
logFC is a log2-fold-change corresponding to a contrast.

```
## huvec-retina
results2 <- topTable(huvec_ebay, number = 100000, coef = "huvec_retina")
results2 <- cbind(results2, getsymbols)

results2$threshold <- "1"
a <- subset(results2, adj.P.Val < 0.05 & logFC > 5)
results2[rownames(a), "threshold"] <- "2"
b <- subset(results2, adj.P.Val < 0.05 & logFC < -5)
results2[rownames(b), "threshold"] <- "3"
table(results2$threshold)

##
##     1     2     3
## 54557    24    94

## huvec-iris
results3 <- topTable(huvec_ebay, number = 100000, coef = "huvec_iris")
results3 <- cbind(results3, getsymbols)

results3$threshold <- "1"
a <- subset(results3, adj.P.Val < 0.05 & logFC > 5)
results3[rownames(a), "threshold"] <- "2"
b <- subset(results3, adj.P.Val < 0.05 & logFC < -5)
```

```
results3[rownames(b), "threshold"] <- "3"
table(results3$threshold)

##
##     1     2     3
## 54601    25    49
```

Vulcano plot for the huvec-retina pair:

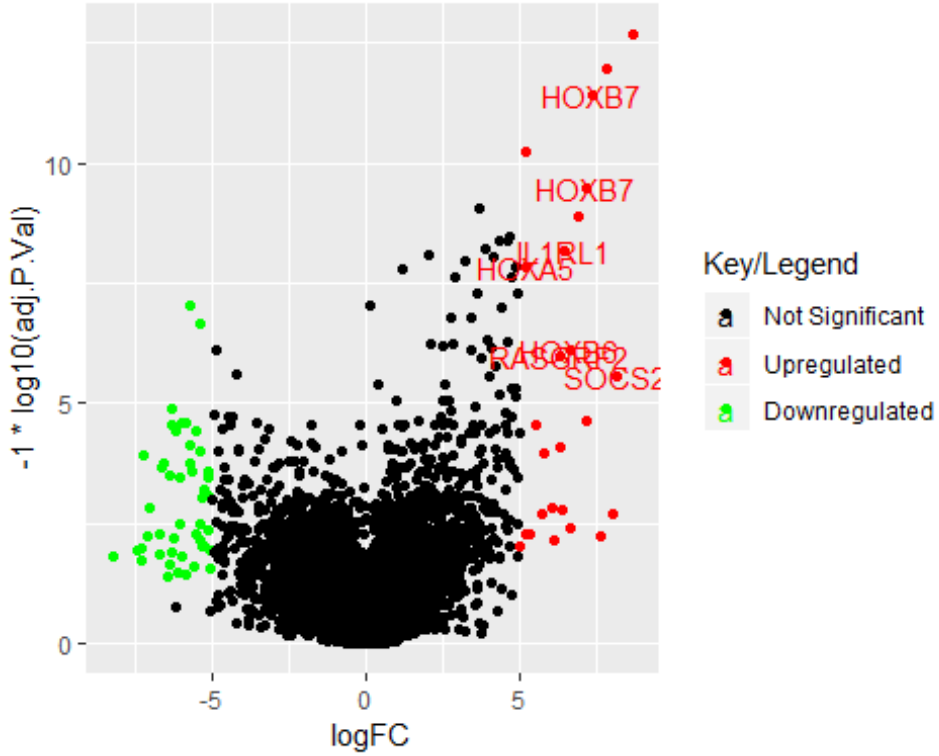
```
library(ggplot2)
volcano <- ggplot(data = results2,
                  aes(x = logFC, y = -1*log10(adj.P.Val),
                      colour = threshold,
                      label = getsymbols))

volcano <- volcano +
    geom_point() +
    scale_color_manual(values = c("black", "red", "green"),
                       labels = c("Not Significant", "Upregulated", "Downregulated"),
                       name = "Key/Legend")

volcano +
    geom_text(data = subset(results2, logFC > 5 & -1*log10(adj.P.Val) > 5), aes(x = logFC, y =
-1*log10(adj.P.Val), colour = threshold, label = getsymbols)  )

## Warning: Removed 4 rows containing missing values (geom_text).
```



Vulcano plot for the huvec-iris pair:

```
volcano <- ggplot(data = results3,
                  aes(x = logFC, y = -1*log10(adj.P.Val),
                      colour = threshold,
                      label = getsymbols))
```

```
volcano <- volcano +
    geom_point() +
    scale_color_manual(values = c("black", "red", "green"),
                       labels = c("Not Significant", "Upregulated", "Downregulated"),
                       name = "Key/Legend")

volcano +
    geom_text(data = subset(results3, logFC > 5 & -1*log10(adj.P.Val) > 5), aes(x = logFC, y =
-1*log10(adj.P.Val), colour = threshold, label = getsymbols)  )

## Warning: Removed 4 rows containing missing values (geom_text).
```



Question 4

```
# significant genes

sig1 <- unique (subset(results, logFC > 5 & -1*log10(adj.P.Val) > 5)$getsymbols)
sig2 <- unique (subset(results2, logFC > 5 & -1*log10(adj.P.Val) > 5)$getsymbols)
sig3 <- unique (subset(results3, logFC > 5 & -1*log10(adj.P.Val) > 5)$getsymbols)
# all_significant = unique(c( as.character(sig1), as.character(sig2), as.character(sig3)))
common_genes <- intersect(intersect(sig1,sig2),sig3)
cat("The significant genes that have a substential effect are: ")

## The significant genes that have a substential effect are:

as.character(common_genes)[2:length(common_genes)]

## [1] "HOXB7"  "IL1RL1" "HOXA5"
```

- HOXB7: The HOXB7 gene encodes a protein with a homeobox DNA-binding domain. It is a member of the Antp homeobox family and is included in a cluster of homeobox B genes that are located on chromosome 17. The encoded protein works as a transcription factor and is involved in cell proliferation and

differentiation. Increased expression of HOXB7 has been associated with cases of melanoma and ovarian cancer. The gene is expressed in several tissues, including (but not limited to) kidney, colon, small intestine, urinary bladder, and fat tissue. (source: https://www.ncbi.nlm.nih.gov/gene/3217)

- IL1RL1: This gene, that is found on human Chromosome 2, is encoding a protein that is a member of the interleukin 1 receptor family. The gene itself is part of the Toll-like receptors super family which is known to play an important role in the human innate immune system. Mutations in this gene have been linked to atopic dermatitis and asthma, while it is also involved in the progression of some cardiac diseases. The protein encoded by this gene is an indicator of the severity or presence of cardiatic diseases. The gene is expressed in placenta, kidney, lung, gal bladder and adrenal tissues. (source https://en.wikipedia.org/wiki/IL1RL1)

- HOXA5:

HOXA5 is a gene that is part of the A cluster on chromosome 7 and encodes the Homeobox Hox-A5 protein. It encodes a DNA-binding transcription factor that regulates gene expression, morphogenesis, and differentiation. Since the protein that HOXA5 encodes up-regulates a tumor suppressor, the protein may be important for tumor formation. HoxA5 is suppressed in acute myeloid leukemia. The gene is expressed in several tissues, for example, adrenal, lung, kidney, and fat. (source: https://en.wikipedia.org/wiki/HOXA5)

Gene Ontology terms associated with each gene:

HOXB7:

Molecular function * DNA binding * sequence-specific DNA binding * transcription factor activity, sequence-specific DNA binding * transcriptional activator activity, RNA polymerase II core promoter proximal region sequence-specific binding * RNA polymerase II core promoter proximal region sequence-specific DNA binding * protein binding * RNA polymerase II transcription factor activity, sequence-specific DNA binding * RNA polymerase II distal enhancer sequence-specific DNA binding * transcription factor activity, RNA polymerase II distal enhancer sequence-specific binding * transcriptional activator activity, RNA polymerase II transcription regulatory region sequence-specific binding *Cellular component*
* cell nucleus *Biological process*
* embryonic skeletal system morphogenesis * pattern specification process * skeletal system development * respiratory system process * regulation of transcription, DNA-templated * multicellular organism growth * lung development * trachea morphogenesis * positive regulation of receptor biosynthetic process * bronchiole development * cell-cell signaling involved in mammary gland development * regulation of mammary gland epithelial cell proliferation * mammary gland alveolus development * lung goblet cell differentiation * transcription, DNA-templated * mammary gland epithelial cell differentiation * negative regulation of erythrocyte differentiation * respiratory gaseous exchange * trachea cartilage morphogenesis * embryonic skeletal system development * morphogenesis of an epithelium * multicellular organism development * thyroid gland development * lobar bronchus epithelium development * mesenchymal-epithelial cell signaling * positive regulation of apoptotic process * intestinal epithelial cell maturation * negative regulation of angiogenesis * cartilage morphogenesis * lung alveolus development * lung-associated mesenchyme development * epithelial tube branching involved in lung morphogenesis * cell migration * positive regulation of myeloid cell differentiation * anterior/posterior pattern specification * positive regulation of transcription from RNA polymerase II promoter * transcription from RNA polymerase II promoter

source (https://en*wikipedia*org/wiki/HOXA5)

IL1RL1:

Molecular function
* interleukin-33 receptor activity * interleukin-1 receptor activity * cytokine receptor activity * protein binding * signal transducer activity, downstream of receptor * interleukin-33 binding *Cellular component*
* integral component of membrane * external side of plasma membrane * membrane * extracellular region * plasma membrane * extracellular matrix *Biological process*
* negative regulation of T-helper 1 type immune response * positive regulation of inflammatory response *

negative regulation of interferon-gamma production * positive regulation of macrophage activation * negative regulation of I-kappaB kinase/NF-kappaB signaling * positive regulation of chemokine secretion * signal transduction * immune response * positive regulation of interleukin-5 production * interleukin-33-mediated signaling pathway

source (https://en.wikipedia.org/wiki/IL1RL1)

HOXA5:

Molecular function
* DNA binding * sequence-specific DNA binding * transcription factor activity, sequence-specific DNA binding * transcriptional activator activity, RNA polymerase II core promoter proximal region sequence-specific binding * RNA polymerase II core promoter proximal region sequence-specific DNA binding * protein binding * RNA polymerase II transcription factor activity, sequence-specific DNA binding * RNA polymerase II distal enhancer sequence-specific DNA binding * transcription factor activity, RNA polymerase II distal enhancer sequence-specific binding * transcriptional activator activity, RNA polymerase II transcription regulatory region sequence-specific binding *Cellular component*
* cell nucleus *Biological process*
* embryonic skeletal system morphogenesis * pattern specification process * skeletal system development * respiratory system process * regulation of transcription, DNA-templated * multicellular organism growth * lung development * trachea morphogenesis * positive regulation of receptor biosynthetic process * bronchiole development * cell-cell signaling involved in mammary gland development * regulation of mammary gland epithelial cell proliferation * mammary gland alveolus development * lung goblet cell differentiation * transcription, DNA-templated * mammary gland epithelial cell differentiation * negative regulation of erythrocyte differentiation * respiratory gaseous exchange * trachea cartilage morphogenesis * embryonic skeletal system development * morphogenesis of an epithelium * multicellular organism development * thyroid gland development * lobar bronchus epithelium development * mesenchymal-epithelial cell signaling * positive regulation of apoptotic process * intestinal epithelial cell maturation * negative regulation of angiogenesis * cartilage morphogenesis * lung alveolus development * lung-associated mesenchyme development * epithelial tube branching involved in lung morphogenesis * cell migration * positive regulation of myeloid cell differentiation * anterior/posterior pattern specification * positive regulation of transcription from RNA polymerase II promoter * transcription from RNA polymerase II promoter

As we can see from the lists of the GO above, all three genes are involved in protein binding. Protein binding is the ability of the proteins to form bonds with other substances, commonly this refers to the drugs binding to the proteins in blood. Hence, all of them play a role into how drugs and human bodies interact (they can enhance or detract the drug performances).