

thema07

Keimpe Dijkstra

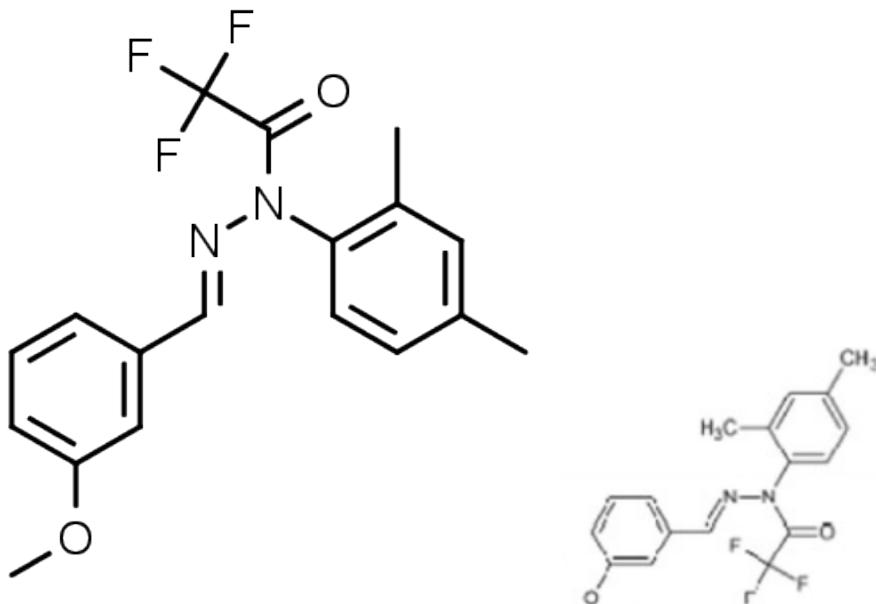
8-3-2022

Contents

1	Intrdoduction	3
2	Exploratory data analysis(EDA)	3
2.1	Data preperation	3
2.1.1	Downloading neccesary packages	3
2.1.2	loading data into R	4
2.2	Visualizations	4
2.2.1	Basic statistics	4
2.2.2	More preperation	5
2.2.3	Boxplot	5
2.2.4	Density plot	6
2.2.5	Barplot sequence depth	7
2.2.6	Normalization	8
2.2.7	Distance Calculation	9
2.2.8	heatmap(both datasets)	9
2.2.9	heatmap j147	10
2.2.10	heatmap cad-31	11
2.2.11	multi dimensional scaling plot	12
2.2.12	conclusions EDA	13
3	Discovering Differentialy Expressed Genes (DEGs)	14
3.1	Pre-processing	14
3.1.1	FPM normalization	14
3.1.2	14
3.2	Fold Change Value	15
3.2.1	j47 old vs old	15
3.2.2	AD.cad31 vs AD	16
3.3	Bio conductor	17
3.3.1	design matrix	17
3.3.2	DESeq2	18
3.3.3	EdgeR	21
4	Data Analysis and Visualization	25
4.1	Volcano Plot	25
4.2	Venn diagram	27
4.3	Signaling Pathway Impact Analysis	27
4.3.1	Collecting deg's	27
4.3.2	Data preperation	27
4.3.3	Pathway analysis	28
5	Quick overview	38

1 Intrdroduction

In this project we compared two different rna-seq datasets, both datasets came from studies that researched novel drugs for alzheimers disease(AD). The two drugs of choice are called J147 and CAD31, the CAD-31 drug is a derivative of the J147 drug. Their structures are shown below, we can see their structures are very similar weher the noticeable difference lies in the second ring structure.



The organism used for this experiment are so called SAMP8 mice, this strain of mice exhibits a lot of the syptomes that a AD patient would. The sequenced samples were taken from the hippocampal region in the brain. The J147 study had three groups, two old groups with alzheimers with one on and the other group off the drug, and a young control group also with induced AD. The CAD-31 study consisted of four groups, two samp8 groups and two wild type(WT) groups. Each of these groups had mice on and off the drug.

2 Exploratory data analysis(EDA)

2.1 Data preperation

2.1.1 Downloading neccesary packages

```
library(affy)
library(scales)
library(DESeq2)
library(pheatmap)
library(PoClaClu)
library(ggplot2)
library(edgeR)
library(EnhancedVolcano)
library(gplots)
library(SPIA)
library(dplyr)
```

2.1.2 loading data into R

```
#load in the data using the read.table method
j147 <- read.table("./data/j147.csv", header = T, sep = ",", quote="")
cad31 <- read.table("./data/cad-31.txt", header = T, sep = "\t", fill = T, quote="")

#create a new column for te gene annotation
cad31$Gene <- strsplit(cad31$Annotation.Divergence, "|", 1)

#filter the gene annotation out of the string
counter <- 1
while( counter < length(cad31$Gene)) {

  cad31$Gene[counter] <- cad31$Gene[[counter]][1]
  counter <- counter + 1
}

#drop some unnecesary columns
cad31 <- cad31[c(22, 9:21)]

#merge all the data into one dataframe
data <- merge(j147, cad31)

#replace na's with zero's
data[is.na(data)] <- 0

#rename the columns
names(data) <- c(
  paste0('Gene'),
  paste0('AD.old.j147.', 1:3),
  paste0('AD.old.', 1:3),
  paste0('AD.young.', 1:4),
  paste0('AD.cad31.', 1:3),
  paste0('AD.', 1:3),
  paste0('WT.CAD31.', 1:3),
  paste0('WT.', 1:4)
)

#set some indices to help future work
AD.old.j147 <- 2:4
AD.old <- 5:7
AD.young <- 8:11
AD.cad31 <- 12:14
AD <- 15:17
WT.CAD31 <- 18:20
WT <- 21:24
```

2.2 Visualizations

2.2.1 Basic statistics

```
#Create summary's for all groups
summary(rowSums(data[AD.old.j147])/3)
```

```

##      Min. 1st Qu. Median   Mean 3rd Qu.   Max.
##      0.00    0.33  56.67 580.69 498.33 87522.00
summary(rowSums(data[AD.old])/3)

##      Min. 1st Qu. Median   Mean 3rd Qu.   Max.
##      0.00    0.33  53.00 540.94 464.67 76552.67
summary(rowSums(data[AD.young])/4)

##      Min. 1st Qu. Median   Mean 3rd Qu.   Max.
##      0.00    0.00  41.25 446.74 382.00 76737.25
summary(rowSums(data[AD.cad31])/3)

##      Min. 1st Qu. Median   Mean 3rd Qu.   Max.
##      0.00    0.33  62.67 585.45 472.00 69317.33
summary(rowSums(data[AD])/3)

##      Min. 1st Qu. Median   Mean 3rd Qu.   Max.
##      0.00    0.33  55.67 499.38 410.00 63741.33
summary(rowSums(data[WT.CAD31])/3)

##      Min. 1st Qu. Median   Mean 3rd Qu.   Max.
##      0.00    0.33  44.00 412.07 342.00 43685.33
summary(rowSums(data[WT])/4)

##      Min. 1st Qu. Median   Mean 3rd Qu.   Max.
##      0.0     0.5    63.0   633.5  508.2  71066.2

```

The min and first quantile all show similar or very similar results. The wildtype mouse shows the most expression with an almost double maximum than its drugged counterpart.

The young mouse and the mouse on CAD-31 have the least sequences read whereas the mouse with AD and on drugs seem to be upregulated.

This might be explained because of regulation but also the testing can have an influence on the amount of sequences that are read, therefore we need to normalize the data.

2.2.2 More preparation

Add 1 to the whole dataframe so it can logscale

```

#set the gene column as row names
row.names(data) <- data$Gene
data[1] <- NULL

data <- data + 1

```

2.2.3 Boxplot

```

# set some colors to identify groups
myColors <- hue_pal()(7)

#create the boxplot
boxplot(log2(data), las=2, xlab = "sample mice",
        ylab = "log2 raw gene count", cex.axis=0.5,

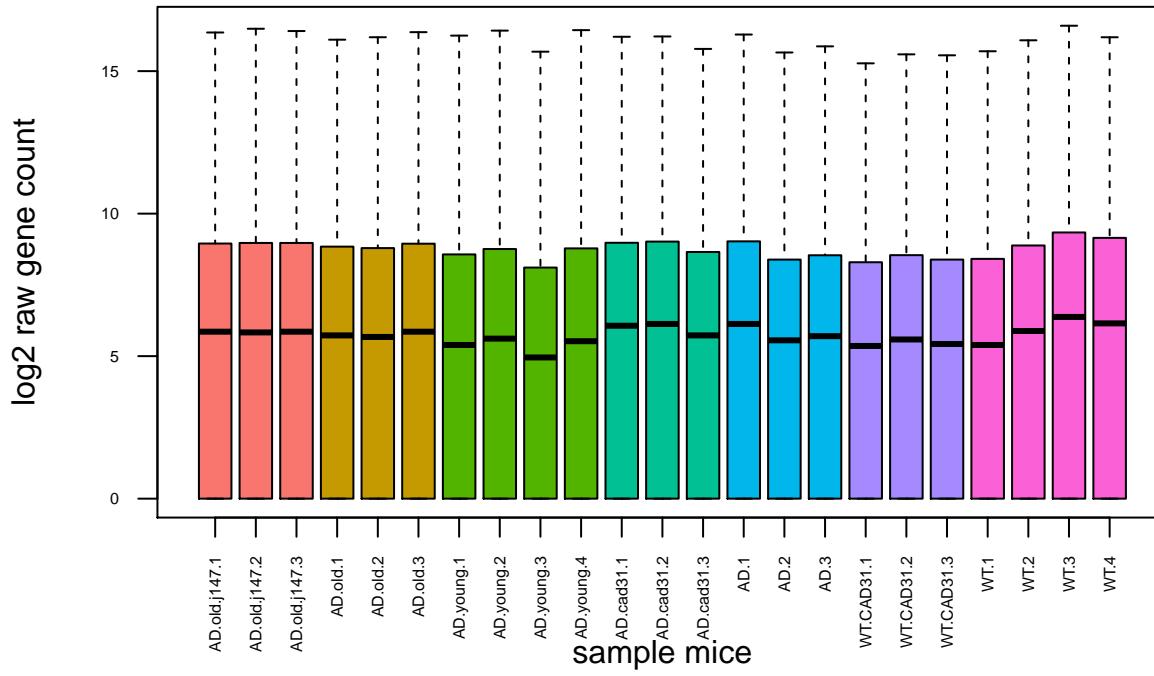
```

```

  col=rep(myColors,c(3,3,4,3,3,3,4)))
title("Boxplot of read counts")

```

Boxplot of read counts



This

boxplot shows the read counts for all groups on a log2 scale. We see no outliers, this is a good sign and means we can continue.

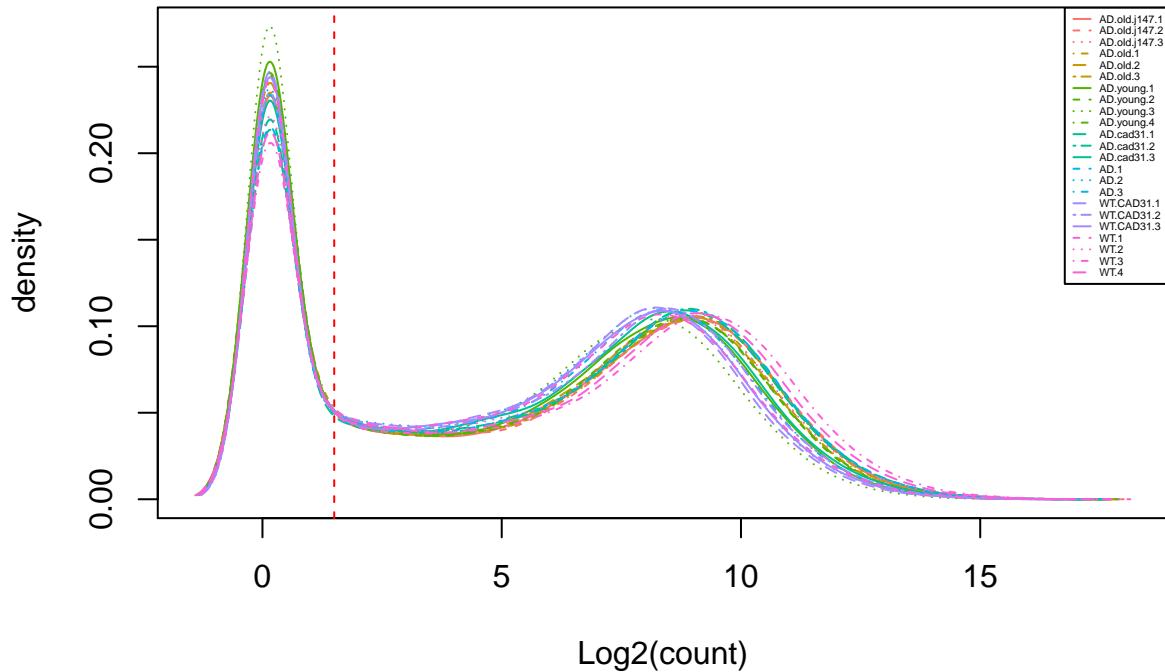
2.2.4 Density plot

```

## Plot the log2-transformed data with a 0.1 pseudocount
plotDensity(log2(data + 0.1),
            col=rep(myColors,c(3,3,4,3,3,3,4)),
            lty=c(1:ncol(data)), xlab='Log2(count)',
            main='Expression Distribution')
legend('topright', names(data), lty=c(1:ncol(data)),
       col=rep(myColors,c(3,3,4,3,3,3,4)), cex = 0.3)
abline(v=1.5, lwd=1, col='red', lty=2)

```

Expression Distribution



Log2(count)

In this

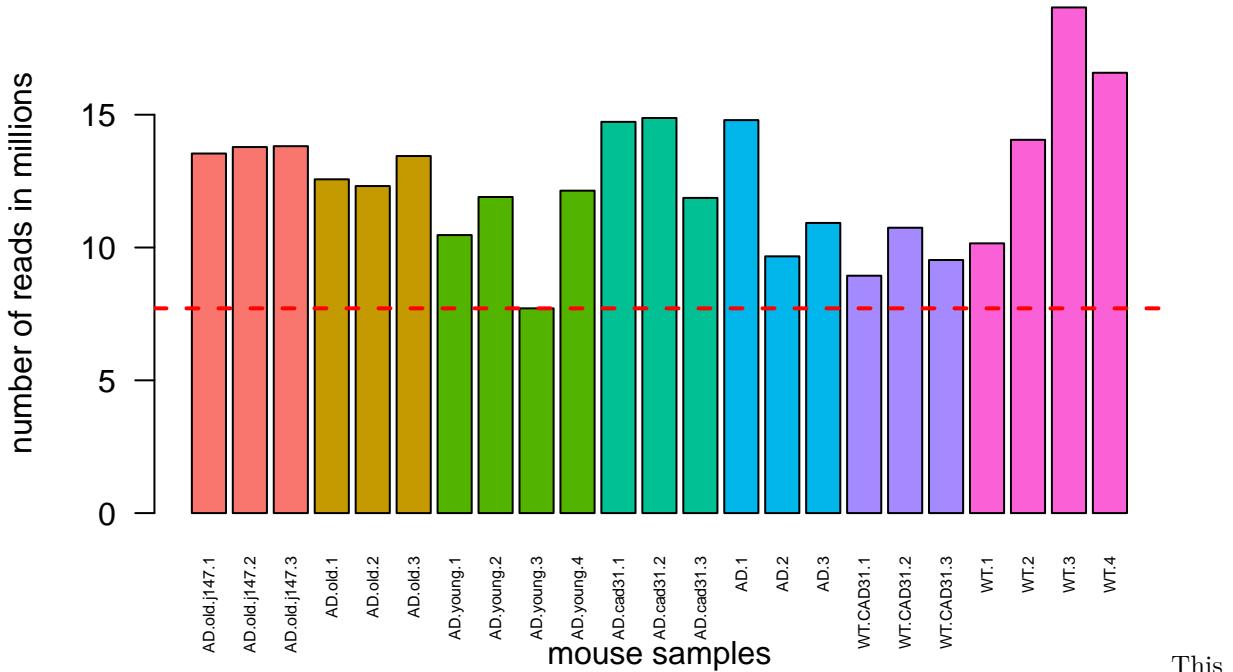
density plot of the expression distribution we see two distinct peaks. The peak left of the dotted line is of little interest because we logscaled the data. The second peak shows a lot of uniformity, exactly what we are looking for, the samples grouped together like this shows the number of reads compared over all the groups is very similar. The height of the peak is of no interest.

2.2.5 Barplot sequence depth

```
barplot(colSums(data)/1000000,      col=rep(myColors,c(3,3,4,3,3,3,4)),
       xlab = "mouse samples", ylab = "number of reads in millions", cex.names= 0.5, las=2)
title("Barplot of sequence depth")

abline(h = min(colSums((data)/1000000)), col = "red",
       lty=2, lwd=2)
```

Barplot of sequence depth



This

barplot shows us the number of reads (in millions) for each sample, an abline is drawn at the sample with the lowest read depth. We can see that there is a more than two fold difference between samples, this isn't a huge problem it just means we have to normalize the data quite rigorously.

2.2.6 Normalization

We first make a summarized experiment object using the deseq package.

```
(ddsMat <- DESeqDataSetFromMatrix(countData = data,
                                    colData = data.frame(samples = names(data)),
                                    design = ~ 1))
```

```
## converting counts to integer mode
## class: DESeqDataSet
## dim: 23573 23
## metadata(1): version
## assays(1): counts
## rownames(23573): 0610005C13Rik 0610007P14Rik ... a 17Rn6
## rowData names(0):
## colnames(23): AD.old.j147.1 AD.old.j147.2 ... WT.3 WT.4
## colData names(1): samples
```

After that we perform the actual normalization and extract its results.

```
#Perform normalization
rld.dds <- vst(ddsMat)
#extract results
rld <- assay(rld.dds)

j.av <- mean(rld[,1:10])
c.av <- mean(rld[,11:23])
```

```
j.bi <- mean(rld[,1:10]) +qt(c(0.025, 0.975), length(rld[,1:10])-1)*sd(rld[,1:10])/sqrt(length(rld[,1:10]))

c.bi <- mean(rld[,11:23]) +qt(c(0.025, 0.975), length(rld[,11:23])-1)*sd(rld[,11:23])/sqrt(length(rld[,11:23]))

#rld[,11:23] <- rld[,11:23] * (c.bi/j.bi)
```

2.2.7 Distance Calculation

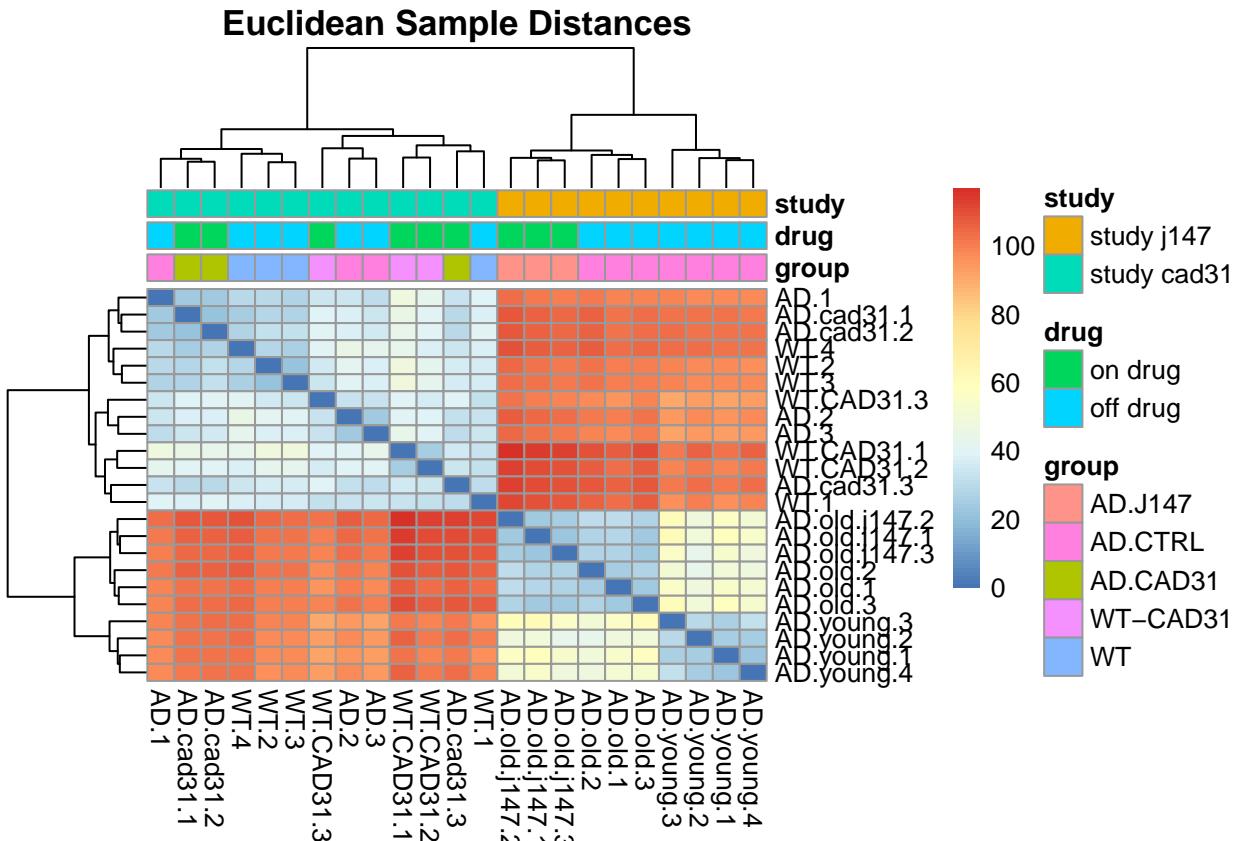
Now the data has been normalized we can calculate the distance metrics for all possible combination of samples

```
#Calculate the distances using the euclidean method
sampledist <- dist(t(rld), method = "euclidean", diag = T, upper = T)
#Convert the results to a matrix
sampledistmatrix <- as.matrix(sampledist)
```

2.2.8 heatmap(both datasets)

```
#Construction of annotation data
ann <- data.frame(group = factor(c(1,1,1,rep(2,each=7),3,3,3,4,4,4,5,5,5,6,6,6,6),
                                labels = c("AD.J147", "AD.CTRL", "AD.CAD31", "AD.CTRL","WT-CAD31", "WT"),
                                drug = factor(c(rep.int(1,3), rep.int(2,7), rep.int(1,3), rep.int(2,3), rep.int(1,3),
                                study = factor(c(rep.int(1,10), rep.int(2,13)),
                                labels = c("study j147", "study cad31"))))

row.names(ann) <- names(data)
#construct heatmap
pheatmap(sampledistmatrix, show_colnames = T,
        annotation_col = ann,
        clustering_distance_rows = sampledist,
        clustering_distance_cols = sampledist,
        main = "Euclidean Sample Distances")
```



In this heatmap the first thing that stands out is the big difference between the two studies, therefore we decided to also make two separate heatmaps for each study.

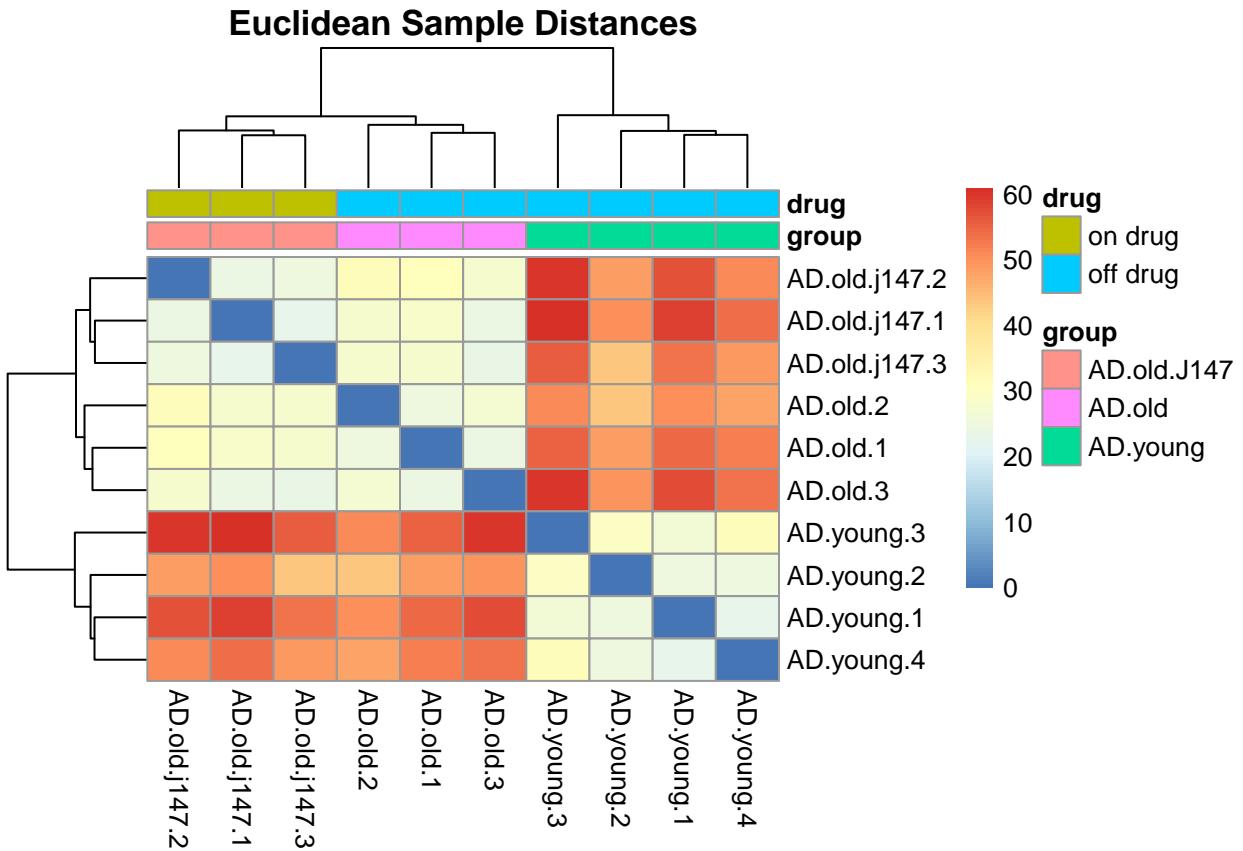
2.2.9 heatmap j147

```
ann <- data.frame(group = factor(c(1,1,1,rep(2,each=3), rep(3, each=4)), labels = c("AD.old.J147", "AD.old.J147", "AD.old.J147), drug = factor(c(rep.int(1,3), rep.int(2,7)), labels = c("on drug", "off drug")))

row.names(ann) <- names(data[1:10])

sampledist <- dist(t(rld[,1:10]), method = "euclidean", diag = T, upper = T)
sampledistmatrix <- as.matrix(sampledist)

pheatmap(sampledistmatrix, show_colnames = T,
         annotation_col = ann,
         clustering_distance_rows = sampledist,
         clustering_distance_cols = sampledist,
         main = "Euclidean Sample Distances")
```



This heatmap is much clearer, it shows a significant difference between the young and old mice. This is not a surprise since age is the biggest risk factor for alzheimer's disease. Further more we see the other two groups grouped together as well, although they don't show that much of a difference between eachother. Notice that the scale is relatively small meaning that the differences are very distinct but not that big in literal terms.

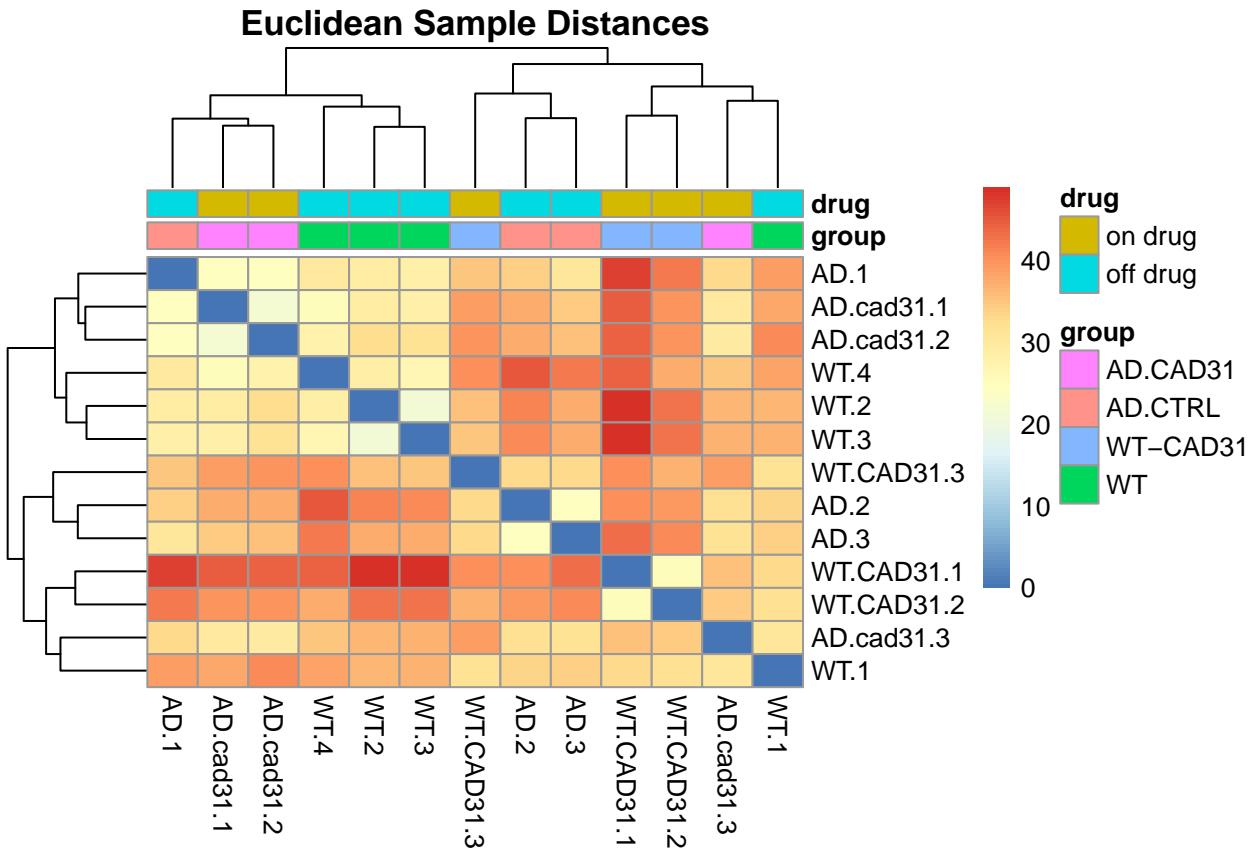
2.2.10 heatmap cad-31

```
ann <- data.frame(group = factor(c(rep(1,each=3),rep(2,each=3), rep(3, each=3), rep(4,each=4)), labels = c("AD.old.J147.1", "AD.old.J147.2", "AD.old.J147.3", "AD.old.1", "AD.old.2", "AD.old.3", "AD.young.1", "AD.young.2", "AD.young.3", "AD.young.4", "AD.old.j147.1", "AD.old.j147.2", "AD.old.j147.3"), drug = factor(c(rep.int(1,3), rep.int(2,3), rep.int(1,3), rep.int(2,4))), labels = c("on drug", "off drug"))

row.names(ann) <- names(data[11:23])

sampledist <- dist(t(rld[,11:23]), method = "euclidean", diag = T, upper = T)
sampledistmatrix <- as.matrix(sampledist)

pheatmap(sampledistmatrix, show_colnames = T,
         annotation_col = ann,
         clustering_distance_rows = sampledist,
         clustering_distance_cols = sampledist,
         main = "Euclidean Sample Distances")
```



This heatmap of the CAD-31 drug is all over the place, we see an even smaller scale and on first sight no clear forming of groups. But if we assume each group somehow has an outlier it all makes a lot more sense. Unfortunately we can't drop any samples since we at most only have in a group.

2.2.11 multi dimensional scaling plot

```

dds <- assay(ddsMat)
poisD <- PoissonDistance( t(dds), type = "deseq")
# Extract the matrix with distances
samplePoisDistMatrix <- as.matrix(poisD$dd)
# Calculate the MDS and get the X- and Y-coordinates
mdsPoisData <- data.frame( cmdscale(samplePoisDistMatrix) )

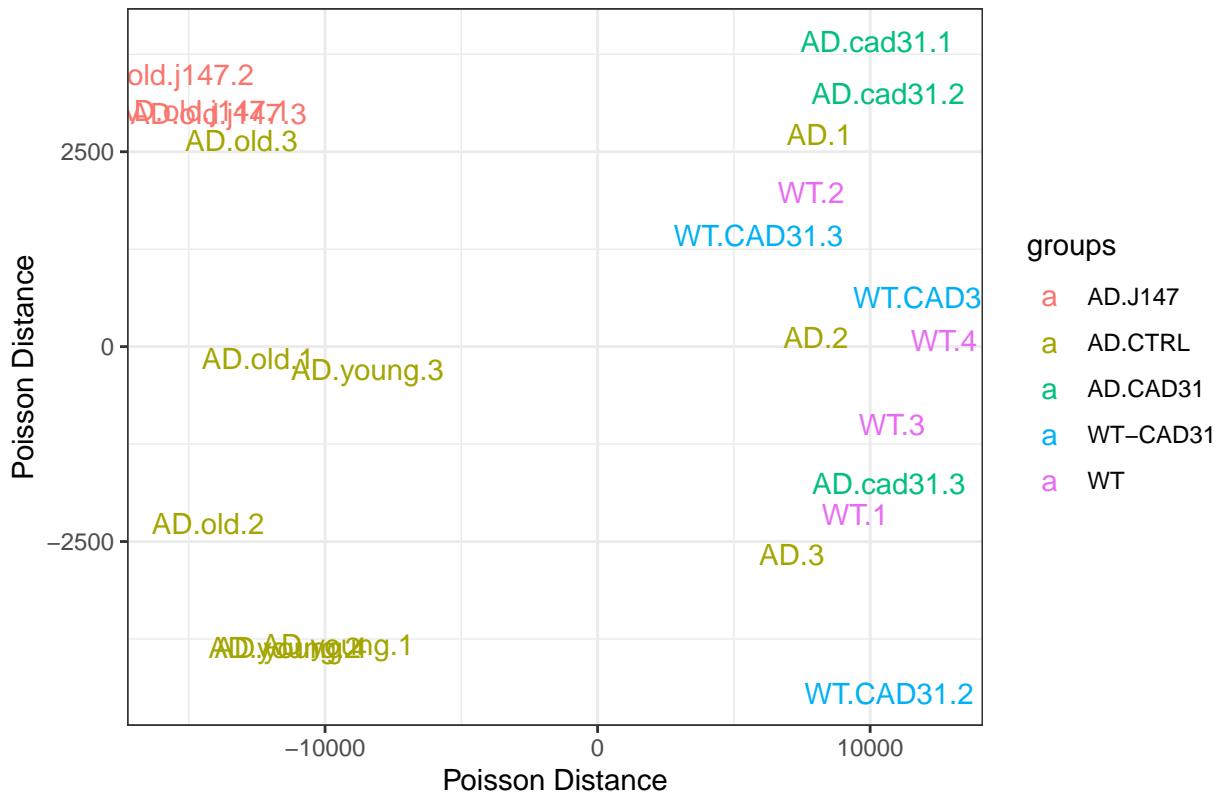
# And set some better readable names for the columns
names(mdsPoisData) <- c('x_coord', 'y_coord')
groups <- factor(c(1,1,1,rep(2,each=7),3,3,3,4,4,4,5,5,5,6,6,6,6),
                  labels = c("AD.J147", "AD.CTRL", "AD.CAD31", "AD.CTRL", "WT-CAD31", "WT"))

coldata <- names(data)

# Create the plot using ggplot
ggplot(mdsPoisData, aes(x_coord, y_coord, color = groups, label = coldata)) +
  geom_text(size = 4) +
  ggtitle('Multi Dimensional Scaling') +
  labs(x = "Poisson Distance", y = "Poisson Distance") +
  theme_bw()

```

Multi Dimensional Scaling



The multi dimensional scaling plot shows the distance on a 2d surface. Here we again see a noticeable difference between the two studies on the x-axis. But in contrast to the heatmap of J147 we see a very clear grouping of the AD group on the J157 drug. We also see that two of the three samp8 mice on CAD-31 reside on about the same y value as the AD mice on J147, suggesting similarity. Furthermore in the CAD-31 study we see that mice on and off CAD-31 seem to cluster together a bit, but this can't be said with certainty. Finally, apart from one outlier, we can see that in the J147 study the old and young AD mice show moderate but distinct difference in distance.

2.2.12 conclusions EDA

The first thing that pops out when looking at the images created is that there are big differences between the two datasets. This can be seen in the heatmap containing both studies and in the mds plot as well. This throws our comparison of gene expression a bit off, but it shouldn't be too big of a problem because we can just compare the DEG's later on.

A second thing that can be noticed in the heatmap of the j147 is that the age of the mouse seems of high influence on the gene expression but in the mds plot there also is a noticeable difference between mice on and off the drug.

The cad-31 study seems to produce some more mixed results, although harder to detect the same pattern as in the j147 seems to apply.

3 Discovering Differentially Expressed Genes (DEGs)

3.1 Pre-processing

3.1.1 FPM normalization

```
# Perform a naive FPM normalization
counts.fpm <- log2( (data / (colSums(data) / 1e6)) + 1 )
```

3.1.2

```
num.below.one <- 0
num.below.two <- 0

means <- rowMeans(counts.fpm)
difference <- c()

num.with.zero <- 0
num.with.one <- 0
num.with.zero.rows <- c()

i <- 1

while (i <= length(means)){
  if(means[i] < 1){
    num.below.one <- num.below.one + 1
  }
  if(means[i] < 2){
    num.below.two <- num.below.two + 1
  }
  j <- 1
  pass <- F

  while(j < length(counts.fpm[i])){
    if(counts.fpm[i,j] < 1 & pass == F){
      num.with.zero <- num.with.zero + 1
      pass <- T
      num.with.zero.rows <- append(num.with.zero.rows, i)
    }
    if(counts.fpm[i,j] < 2){
      num.with.one <- num.with.one + 1
      j <- length(counts.fpm[i])
    }
    j <- j+1
  }

  difference <- append(difference, max(counts.fpm[i])-min(counts.fpm[i]))

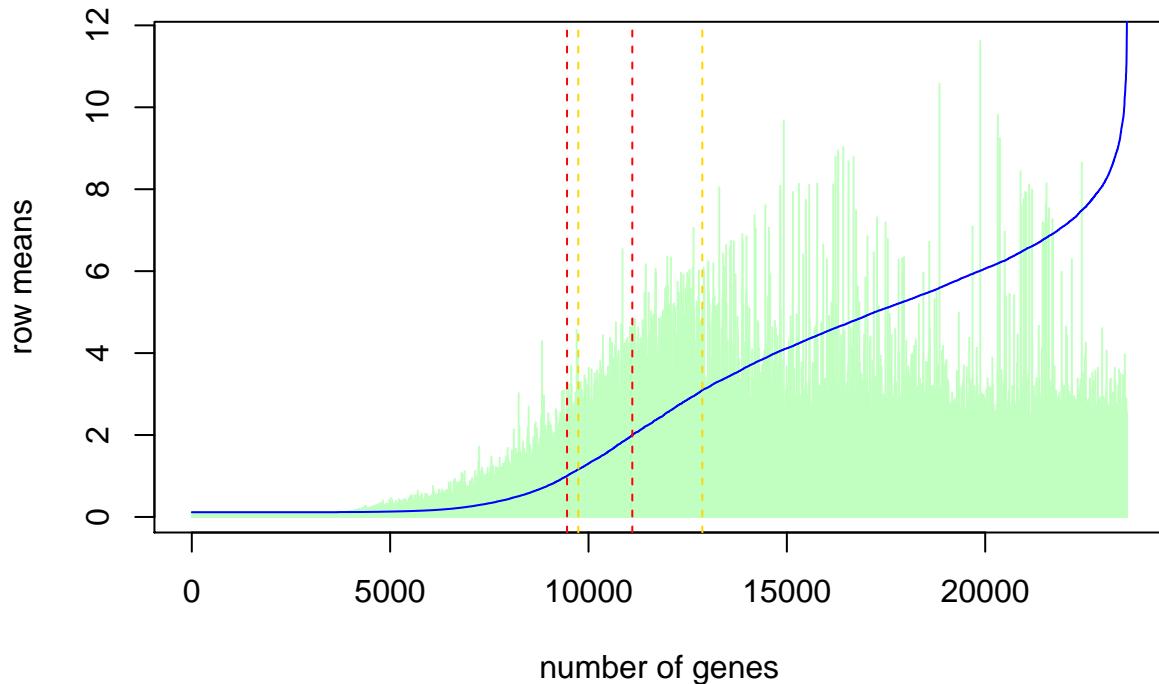
  i <- i+1}

means <- as.data.frame(means)
means <- cbind(means, difference)
means <- means[order(means[,1]),]
```

```

plot(means$difference,type="h",xlab="number of genes", ylab="row means", col="darkseagreen1")
lines(means$means,type="l" ,col="blue",)
abline(v=num.below.one, col="red", lty=2)
abline(v=num.below.two, col="red", lty=2)
abline(v=num.with.zero, col="gold", lty=2)
abline(v=num.with.one, col="gold", lty=2)

```



The plot above describes the means of the rows in the normalized data. The abline's in red are values where the means are below 1 and two respectively and the yellow abline's rows that contain values below 1 and 2. The small difference between the abline's below 1 suggests that there isn't much of a difference between the values, the bigger gap between the values of two suggest the opposite.

The difference is seen in the light green histogram in the back, it first shows no difference relative to the means value, but as the mean gets bigger the difference also starts to increase and eventually kinda settles.

Therefore i decide to set the threshold of filtering where none of the values in the row is below one since after that threshold there seems to be a incline of the row means.

```

original_length <- nrow(counts.fpm)
counts.fpm <- counts.fpm[-num.with.zero.rows,]
cat(original_length-nrow(counts.fpm), "rows are deleted")

```

9743 rows are deleted

3.2 Fold Change Value

Next up we will calculate and plot the fold change values using the earlier calculated log2 values in the fpm variable. The ablines on 1 and -1 indicate the threshold for significance.

3.2.1 j47 old vs old

```

two.groups <- as.data.frame(rowMeans(counts.fpm[AD.old.j147]))
two.groups["AD.old1"] <- rowMeans(counts.fpm[AD.old])

```

```

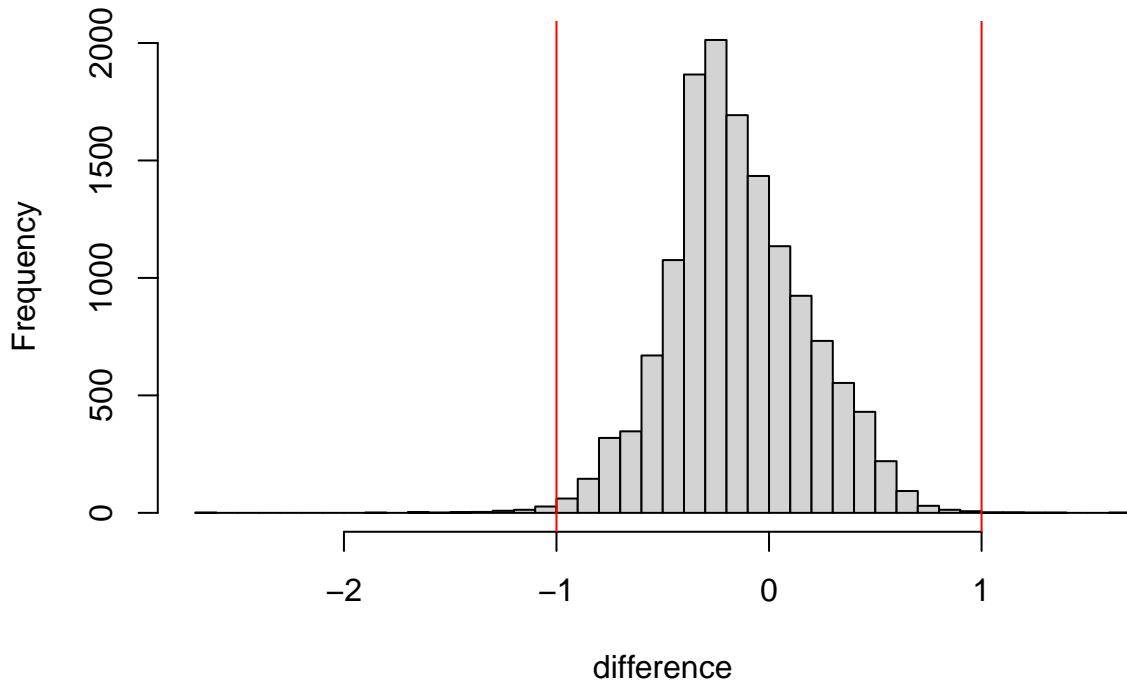
names(two.groups)[1] <- "AD.old.j147"

two.groups["difference"] <- two.groups$AD.old1 - two.groups$AD.old.j147

hist(two.groups$difference, breaks=60, main = "histogram of fold change difference", xlab="difference")
abline(v=-1, col="red")
abline(v=1, col="red")

```

histogram of fold change difference



see minimal number of values that are significant.

Here we

3.2.2 AD.cad31 vs AD

```

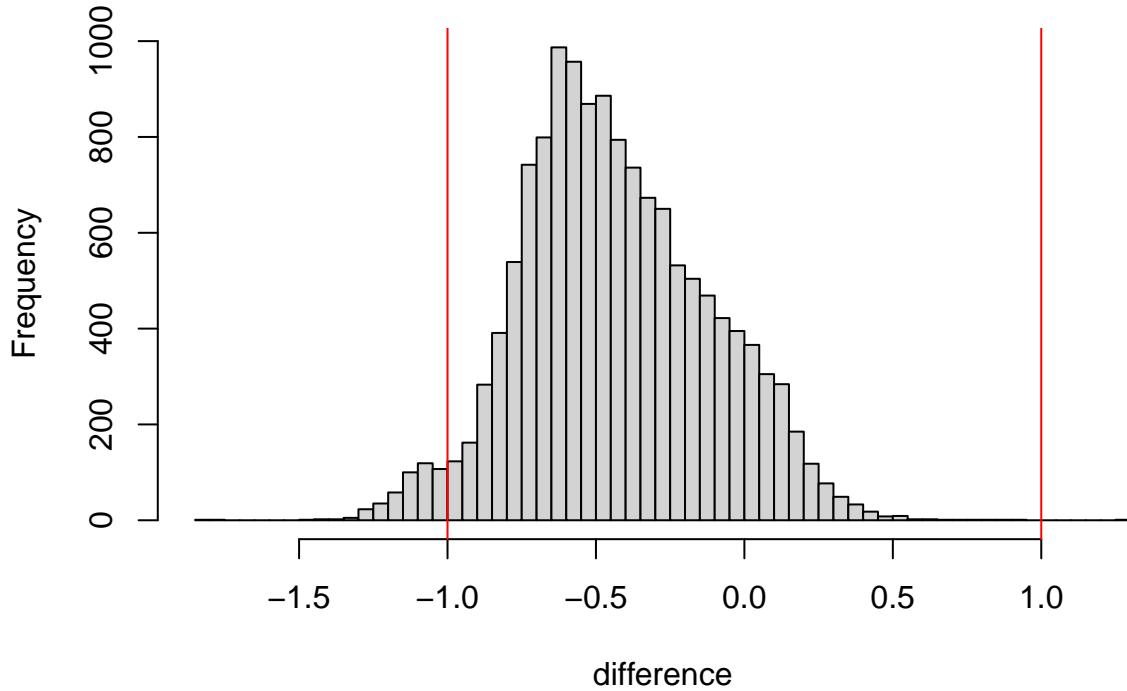
two.groups <- as.data.frame(rowMeans(counts.fpm[AD.cad31]))
two.groups["AD"] <- rowMeans(counts.fpm[AD])
names(two.groups)[1] <- "AD.cad31"

two.groups["difference"] <- two.groups$AD - two.groups$AD.cad31

hist(two.groups$difference, breaks=60, main = "histogram of fold change difference", xlab = "difference")
abline(v=-1, col="red")
abline(v=1, col="red")

```

histogram of fold change difference



plot of CAD-31 we see some more values below -1, meaning that they are downregulated.

In this

3.3 Bio conductor

In this portion of the analysis we will use bioconductor packages for further investigation, this means we will redo some of the previous steps such as normalization. Originally we wanted to work with DESeq only but edgeR gave us some better results, so we chose to continue on the edgeR package. The analysis with DESeq is still kept in the logbook.

3.3.1 design matrix

We first need to again construct a design matrix, similar to the one used with the heatmaps.

```
group <- factor(c(1,1,1,rep(2,each=3),3,3,3,3,4,4,4,5,5,5,6,6,6,rep(7, each=4)),
                 labels = c("AD.J147", "AD.CTRL.OLD","AD.CTRL.YOUNG", "AD.CAD31", "AD.CT"))
(design <- model.matrix(~ group))
```

	(Intercept)	groupAD.CTRL.OLD	groupAD.CTRL.YOUNG	groupAD.CAD31	groupAD.CTRL
## 1	1	0	0	0	0
## 2	1	0	0	0	0
## 3	1	0	0	0	0
## 4	1	1	0	0	0
## 5	1	1	0	0	0
## 6	1	1	0	0	0
## 7	1	0	1	0	0
## 8	1	0	1	0	0
## 9	1	0	1	0	0
## 10	1	0	1	0	0
## 11	1	0	0	1	0
## 12	1	0	0	1	0
## 13	1	0	0	1	0

```

## 14      1      0      0      0      1
## 15      1      0      0      0      1
## 16      1      0      0      0      1
## 17      1      0      0      0      0
## 18      1      0      0      0      0
## 19      1      0      0      0      0
## 20      1      0      0      0      0
## 21      1      0      0      0      0
## 22      1      0      0      0      0
## 23      1      0      0      0      0
##     groupWT-CAD31 groupWT
## 1      0      0
## 2      0      0
## 3      0      0
## 4      0      0
## 5      0      0
## 6      0      0
## 7      0      0
## 8      0      0
## 9      0      0
## 10     0      0
## 11     0      0
## 12     0      0
## 13     0      0
## 14     0      0
## 15     0      0
## 16     0      0
## 17     1      0
## 18     1      0
## 19     1      0
## 20     0      1
## 21     0      1
## 22     0      1
## 23     0      1
## attr("assign")
## [1] 0 1 1 1 1 1
## attr("contrasts")
## attr("contrasts")$group
## [1] "contr.treatment"

```

We also reset the count data back to its original values since the packages handle normalization for us.

```
data <- data -1
```

3.3.2 DESeq2

```

#Firstly we make a DESeqdataset object
dq.dataset <- DESeqDataSetFromMatrix(data, as.data.frame(group), ~0+group)

#Now we fetch the results
result <- DESeq(dq.dataset)
DESeqResults.J147 <- results(result, contrast = c("group", "AD.CTRL.OLD", "AD.J147"),
                               alpha = 0.05)
DESeqResults.CAD31 <- results(result, contrast = c("group", "AD.CTRL", "AD.CAD31"),
                               alpha = 0.05)

```

```

DESeqResults.CTRL <- results(result, contrast = c("group", "AD CTRL", "WT"),
                             alpha = 0.05)

#printing the summary of results
print(summary(DESeqResults.J147))

## 
## out of 19885 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)      : 2, 0.01%
## LFC < 0 (down)    : 22, 0.11%
## outliers [1]       : 2, 0.01%
## low counts [2]     : 7253, 36%
## (mean count < 35)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
##
## NULL

print(summary(DESeqResults.CAD31))

## 
## out of 19885 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)      : 55, 0.28%
## LFC < 0 (down)    : 14, 0.07%
## outliers [1]       : 2, 0.01%
## low counts [2]     : 3818, 19%
## (mean count < 2)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
##
## NULL

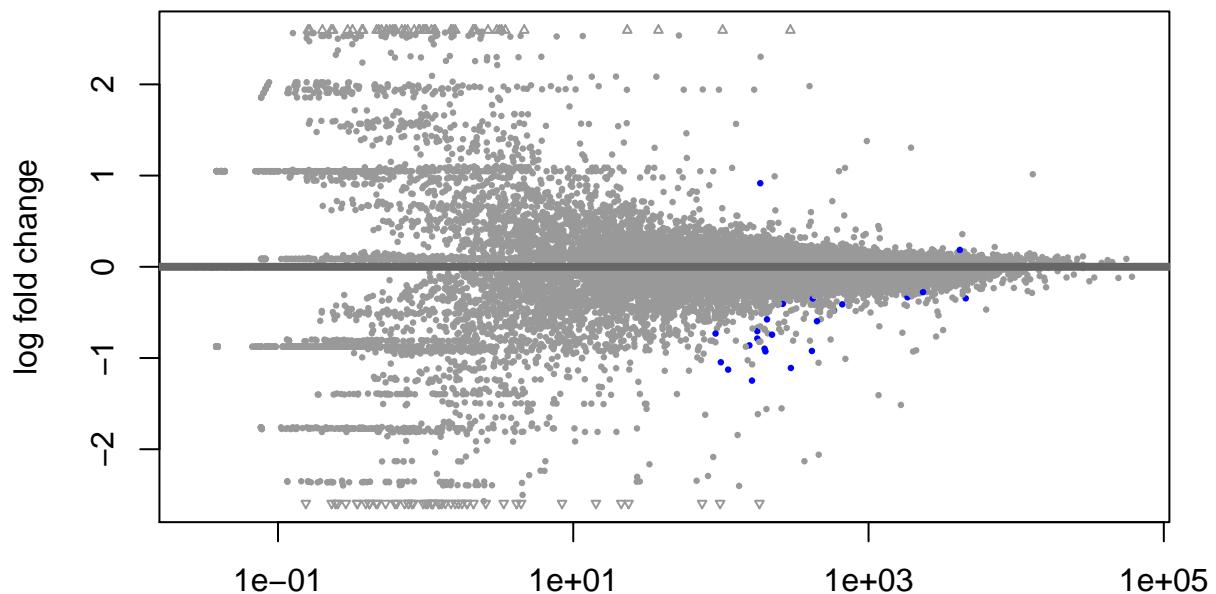
print(summary(DESeqResults.CTRL))

## 
## out of 19885 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)      : 487, 2.4%
## LFC < 0 (down)    : 148, 0.74%
## outliers [1]       : 2, 0.01%
## low counts [2]     : 3436, 17%
## (mean count < 2)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
##
## NULL

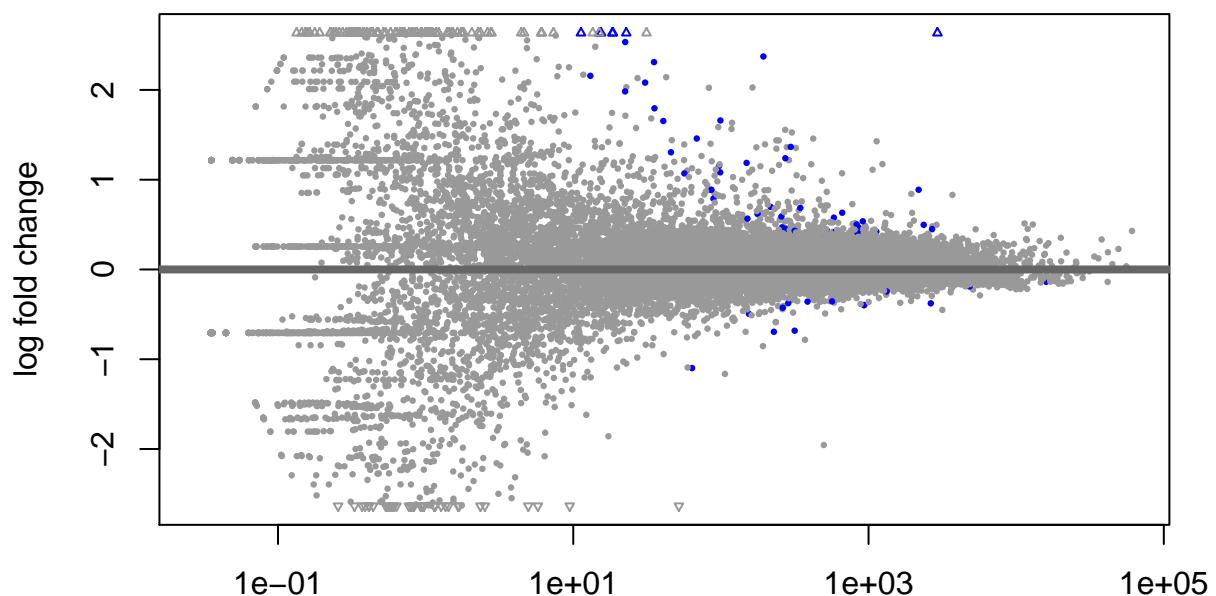
```

These results are quite shocking, the percentage log fold changes and outliers are really low. This falls somewhat in line with the results we saw in the log fold change section. These low results are the reason we reran the analysis in edgeR, hoping for better results.

J147 VS AD.CTRL

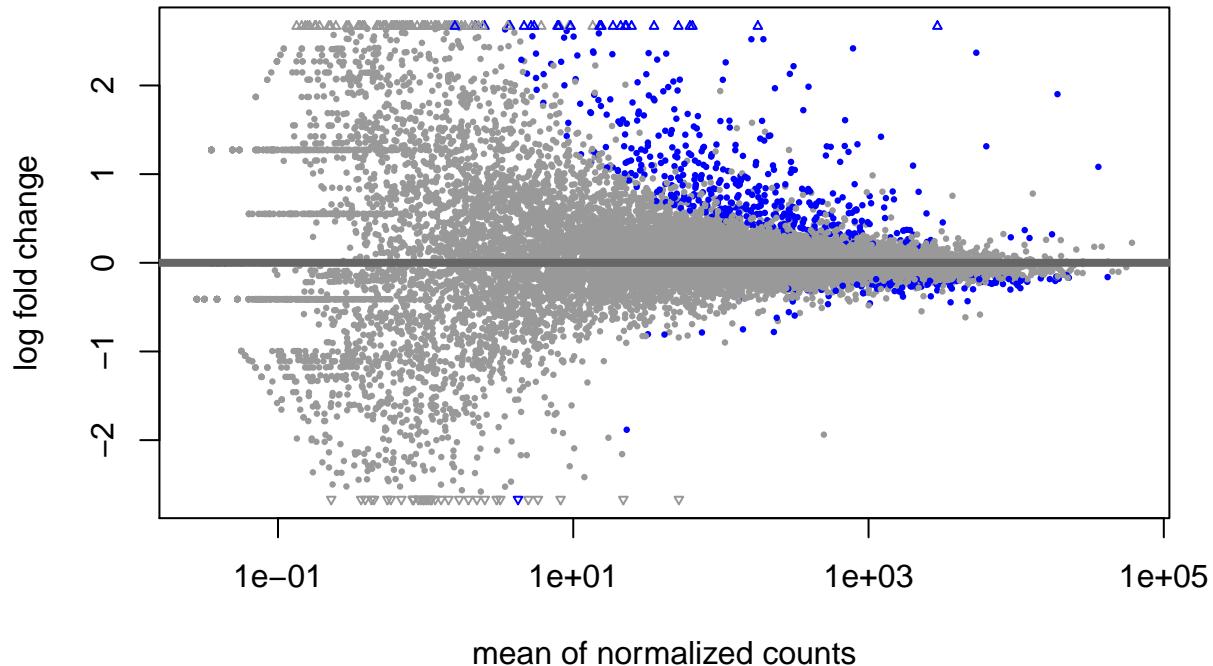


mean of normalized counts
CAD-31 VS AD.CTRL



mean of normalized counts

AD.CTRL(CAD-31 study) VS WT



With these plots the lack of significant is underlined once again. The blue dots represent significant values, in the plots with J147 and CAD-31 we hardly see any. We threw in the AD control vs wildtype just for good measure, expecting a lot more significant values. And yes we do see a lot more in the latter comparision.

3.3.3 EdgeR

```
#function
get_deg <- function(data, group, design, title){
  dl <- DGEList(as.matrix(data), lib.size = colSums(data), norm.factors = rep(1,ncol(data)), group = group)

  dl <- calcNormFactors(dl)
  plotMDS(dl, main = title)

  ed <- estimateDisp(dl, design = design)
  plotBCV(ed, main = title)

  et <- exactTest(ed)

  toptag<- topTags(et,n = 1000, p.value = 0.05)
  return(toptag)
}
```

3.3.3.1 Construction of analysis function

```
# Make the comparison data
J147.data <- data[1:6]
J147.group <- factor(c(1,1,1,rep(2,each=3)),
                      labels = c("AD.J147", "AD.CTRL.OLD"))
J147.design <- model.matrix(~ J147.group)
```

```

cad31.AD.cad31.data <- data[11:16]
cad31.AD.cad31.group <- factor(c(1,1,1,rep(2,each=3)),
                                labels = c("AD.CAD31", "AD"))
cad31.AD.cad31.design <- model.matrix( ~ cad31.AD.cad31.group)

cad31.wt.cad31.data <- data[17:23]
cad31.wt.cad31.group <- factor(c(1,1,1,rep(2,each=4)),
                                labels = c("WT.CAD31", "WT"))
cad31.wt.cad31.design <- model.matrix( ~ cad31.wt.cad31.group)

```

3.3.3.2 Subsetting data

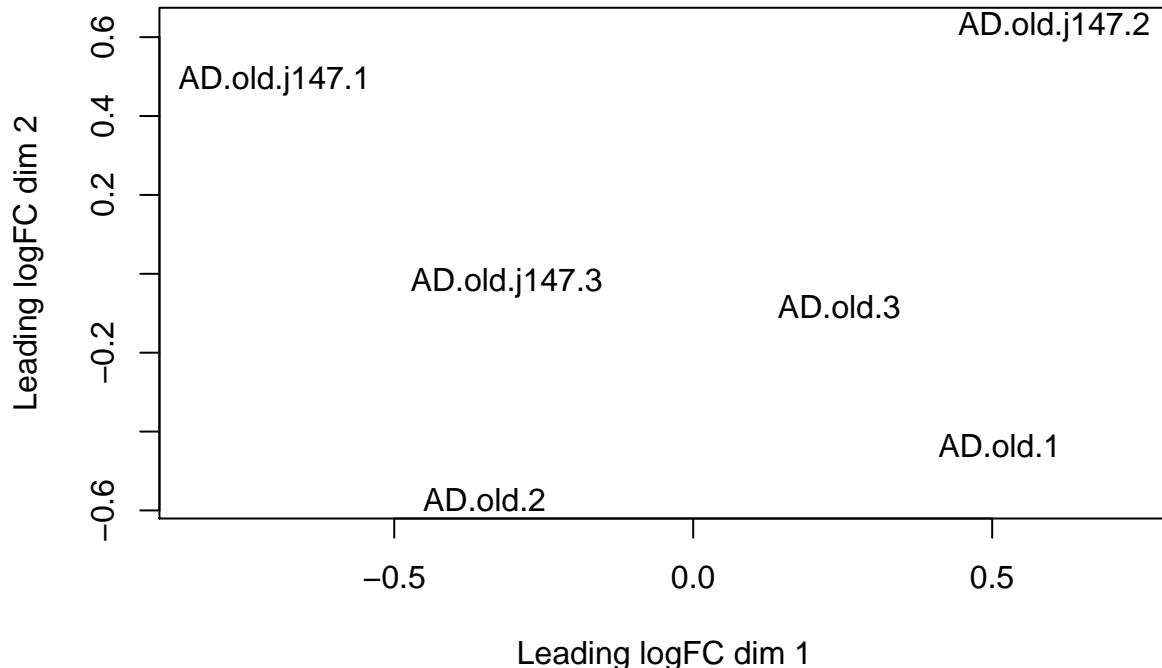
```

#do the analysis
j147.toptag <- get_deg(J147.data, J147.group, J147.design, "J147 VS AD.CTRL")

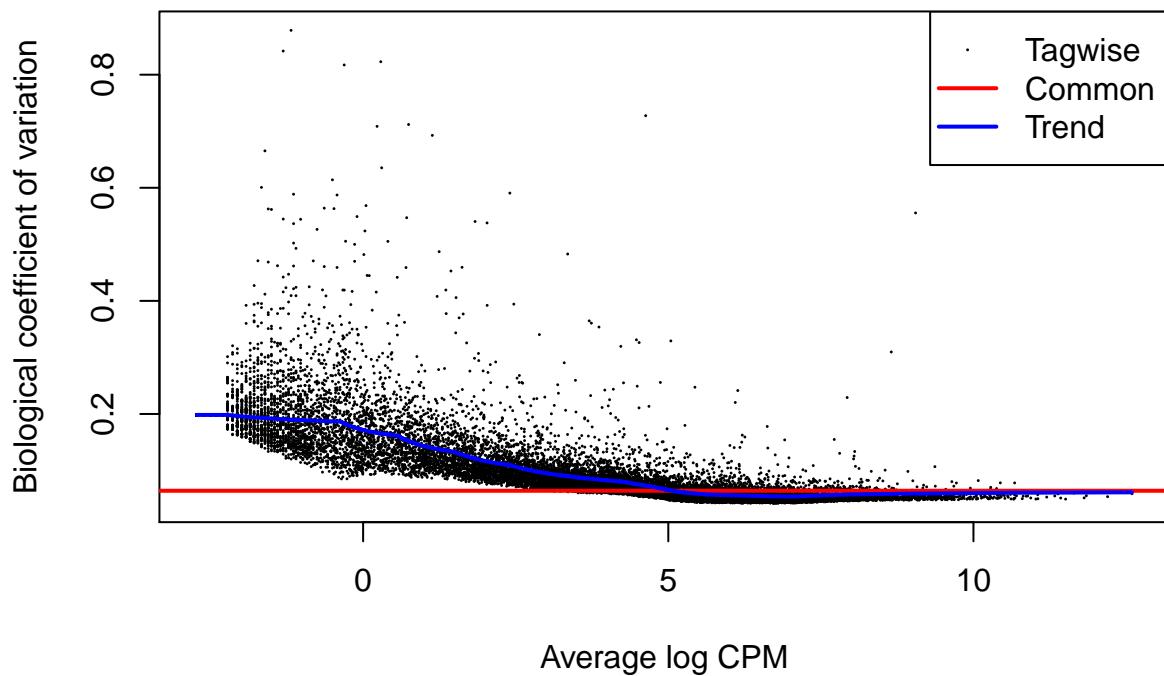
```

3.3.3.3 Performing the analysis

J147 VS AD.CTRL

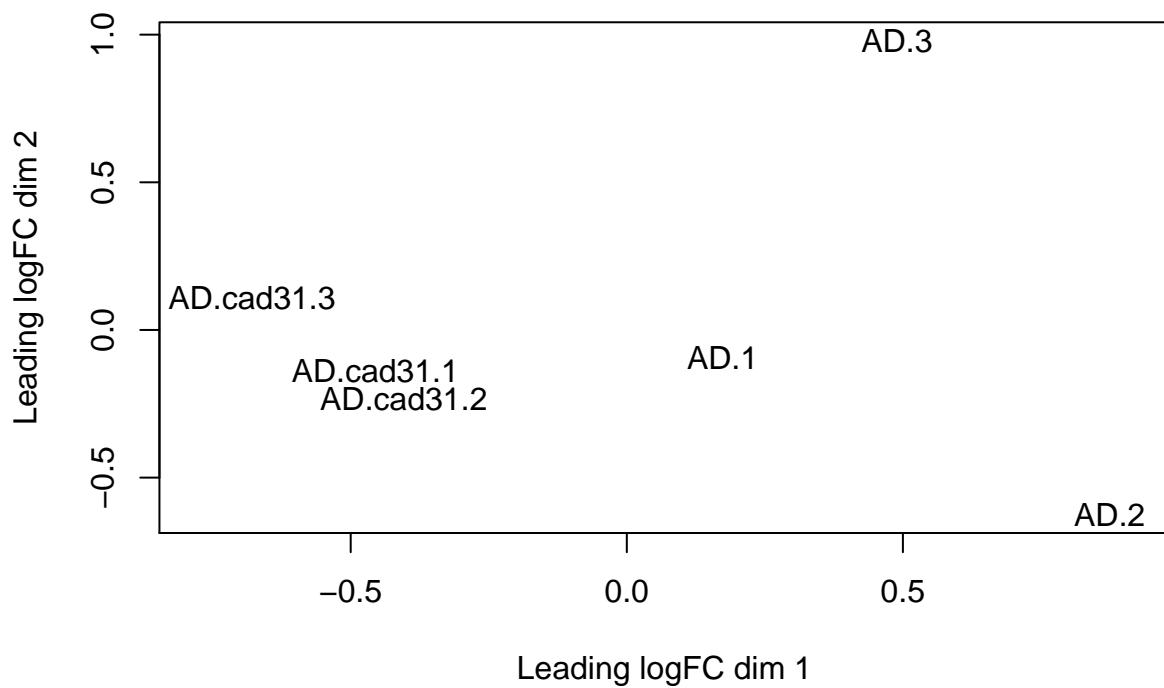


J147 VS AD.CTRL

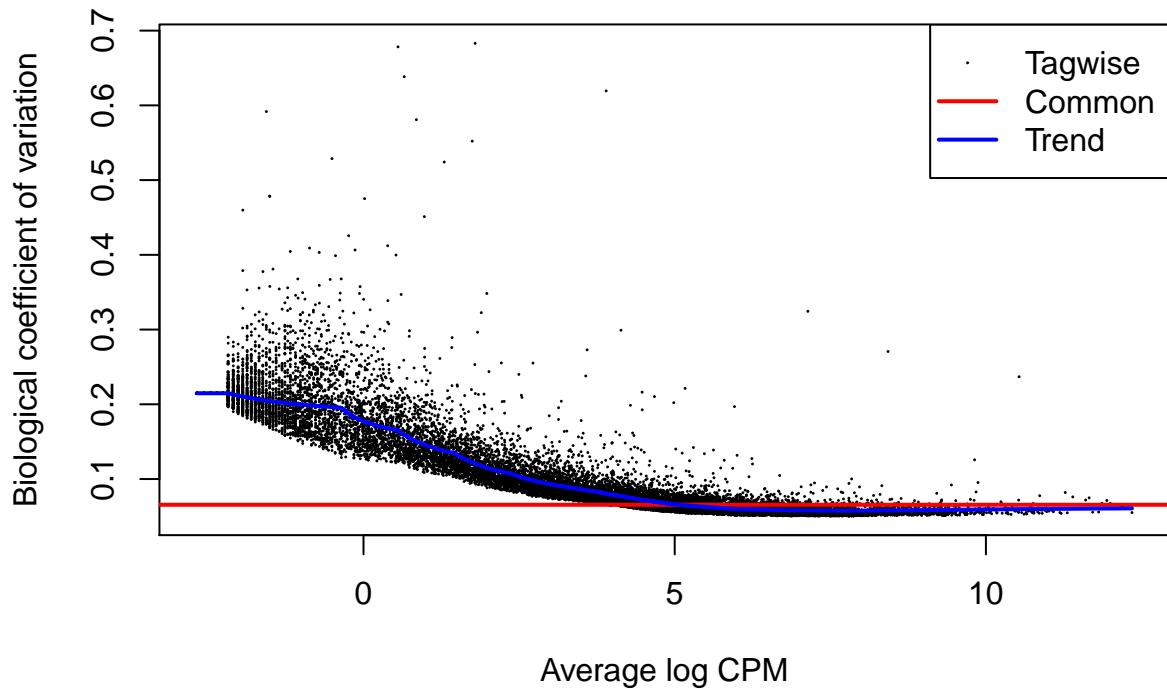


```
cad31.AD.cad31.toptag <- get_deg(cad31.AD.cad31.data, cad31.AD.cad31.group, cad31.AD.cad31.design, "CAD")
```

CAD-31 VS AD.CTRL

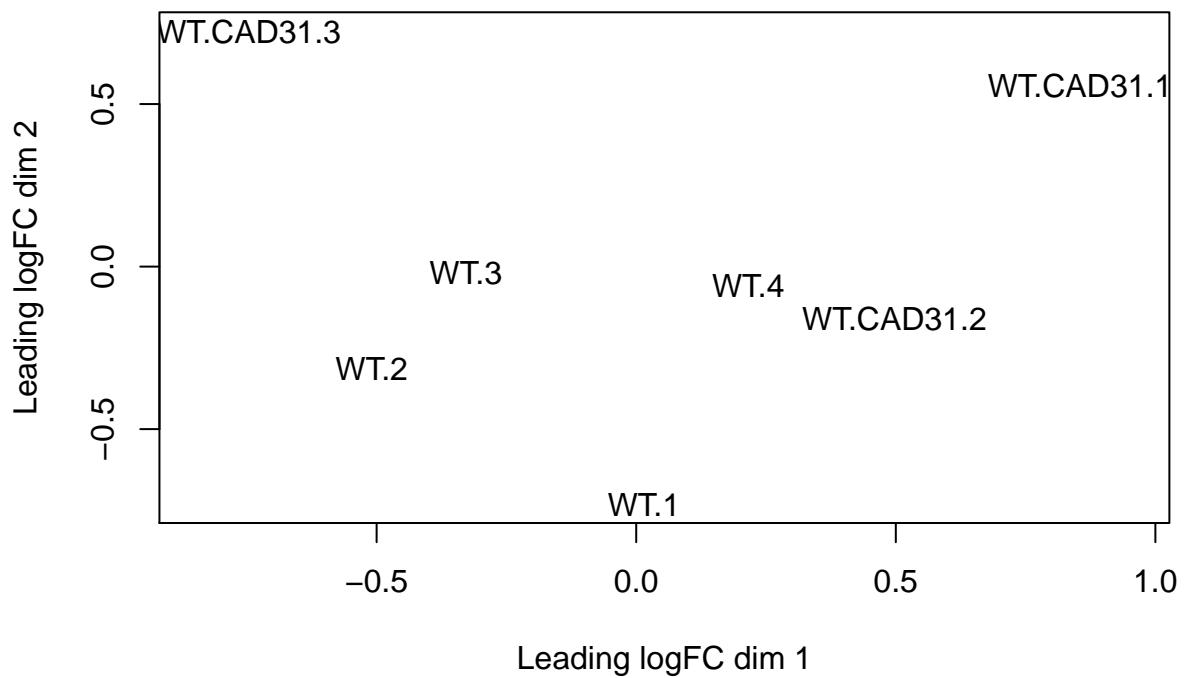


CAD-31 VS AD.CTRL

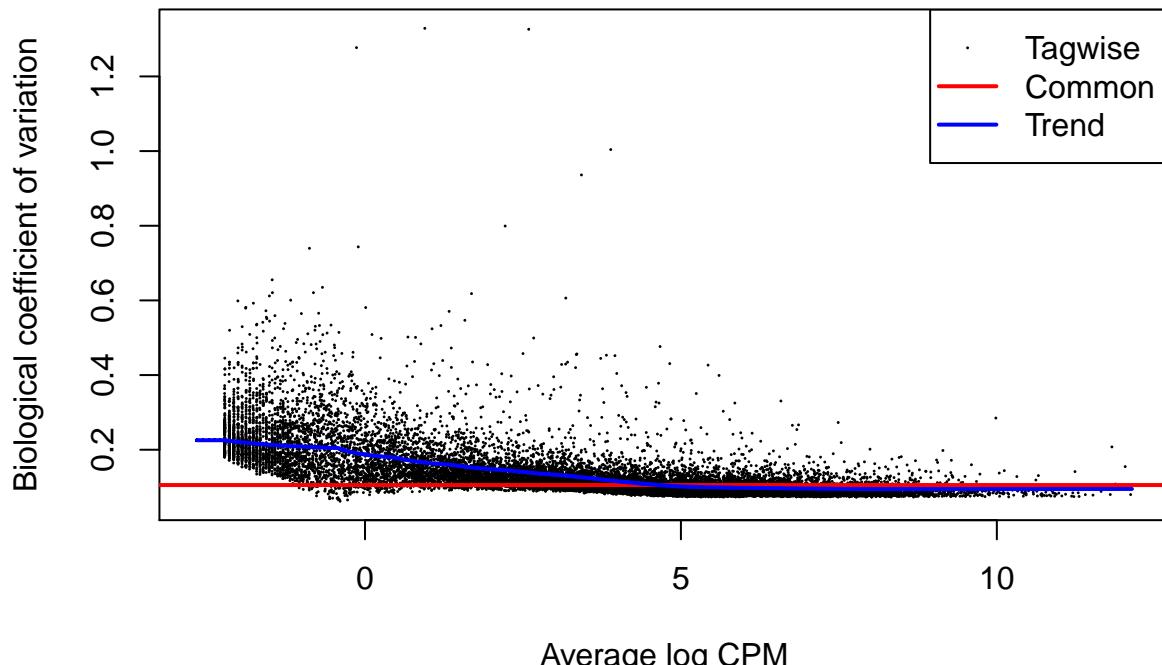


```
cad31.wt.cad31.toptag <- get_deg(cad31.wt.cad31.data, cad31.wt.cad31.group, cad31.wt.cad31.design, "AD.CTRL")
```

AD.CTRL(CAD-31 study) VS WT



AD.CTRL(CAD-31 study) VS WT



J147 VS AD.CTRL In the first plot of this comparison shows a clear difference between the samples where the mice on J147 seem to be upregulated versus the downregulated samp8 mice without medication. The second BCV plot seems relatively normal.

3.3.3.3.1 CAD-31 VS AD.CTRL In the first dimensional plot of the second comparison of CAD-31 vs AD we see a very clear grouping of the CAD-31 samples, and the AD samples being somewhat more dispersed. Here the BCV plot seems very similar to the first one all be it with slightly lower outliers.

3.3.3.3.2 AD.CTRL(CAD-31 study) VS WT We once again added this comparison just to have a reference, we can see a grouping of wildtype in the first plot and the wildtype on CAD-31 seems to have one sample that is an outlier. The BCV plot once again shows a very similar result, this is nice because it confirms that the first two bcv plots show good results.

4 Data Analysis and Visualization

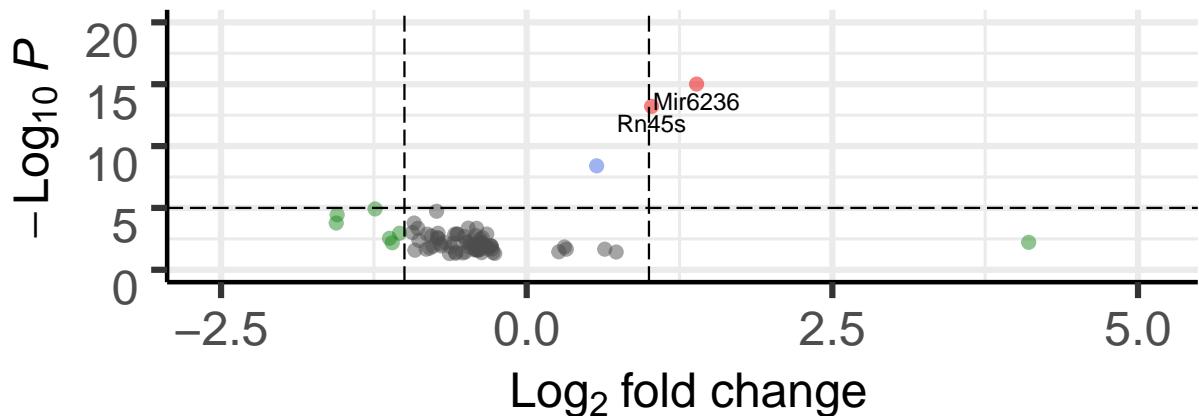
4.1 Volcano Plot

```
EnhancedVolcano(j147.toptag$table, x = 'logFC', y = 'FDR', lab = rownames(j147.toptag$table), title = "J147 VS AD.CTRL")
```

J147 VS AD.CTRL

EnhancedVolcano

● NS ● Log₂ FC ● p-value ● p – value and log₂ FC



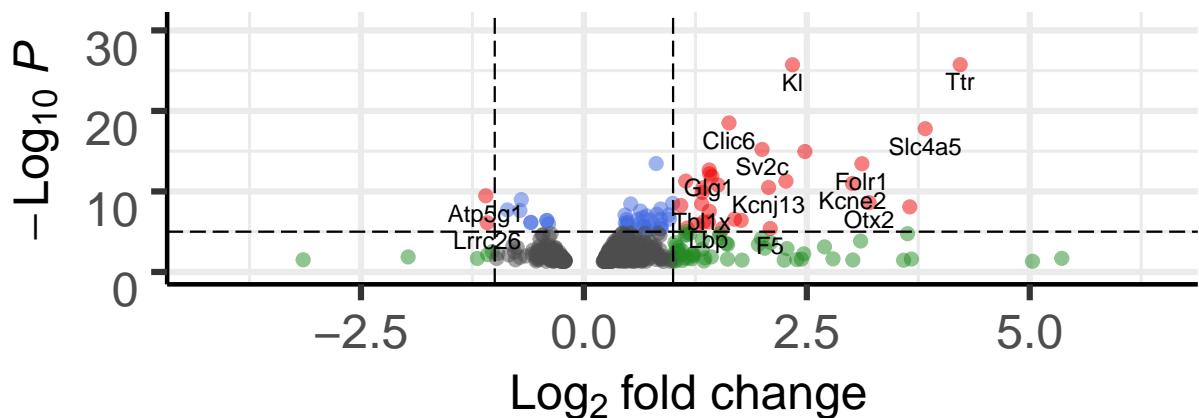
Total = 76 variables

```
EnhancedVolcano(cad31.AD.cad31.toptag$table, x = 'logFC', y = 'FDR', lab = rownames(cad31.AD.cad31.toptag$
```

CAD-31 VS AD.CTRL

EnhancedVolcano

● NS ● Log₂ FC ● p-value ● p – value and log₂ FC



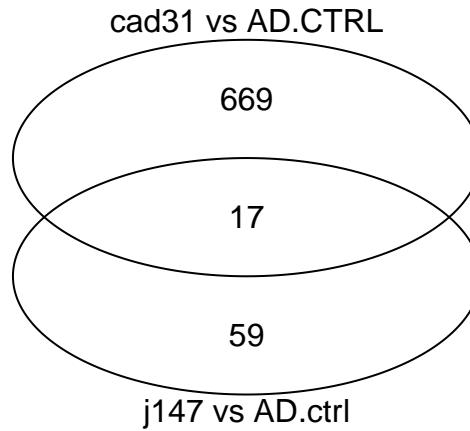
Total = 686 variables

These volcano plots show which deg's are significant and if they're up- or downregulated. Red ones have a

significant p-value and LFC. The first J147 plot shows little significant deg's, all of which are upregulated. The second CAD-31 plot has a lot more variables and really a lot more significant deg's, but the same trend of mostly upregulation is still present.

4.2 Venn diagram

```
# Create a Venn-diagram given just the list of gene-names for both sets
venn(list("j147 vs AD.ctrl" = rownames(j147.toptag),
          "cad31 vs AD.CTRL" = rownames(cad31.AD.cad31.toptag)), )
```



Creating a venn diagram of all the deg's (not only the significant ones) results in seventeen shared deg's

4.3 Signaling Pathway Impact Analysis

4.3.1 Collecting deg's

First we write all the deg's to some files, we then change the gene annotation to entrez id's using an online tool on the DAVID website.

```
degs <- c(rownames(j147.toptag), rownames(cad31.AD.cad31.toptag))
j147.degs <- rownames(j147.toptag)
cad31.degs <- rownames(cad31.AD.cad31.toptag)

# Write gene names to a file
write.table(degs, file = "edger-deg-names.txt",
            row.names = FALSE, quote = FALSE, col.names = FALSE)

write.table(j147.degs, file = "edger-j147deg-names.txt",
            row.names = FALSE, quote = FALSE, col.names = FALSE)

write.table(cad31.degs, file = "edger-cad31deg-names.txt",
            row.names = FALSE, quote = FALSE, col.names = FALSE)
```

4.3.2 Data preparation

In this part we make a table with two columns: entrez ID and the corresponding LFC value.

```
#read table of file downloaded from the david webside
j147.entrezid.gene <- read.table("j147_entrezid_gene.txt", header = T, sep = "\t")
#prep data in markdown for merge
j147.entrezid.gene.conversion <- as.data.frame(rownames(j147.entrezid.gene))
j147.entrezid.gene.conversion$entrez <- j147.entrezid.gene$From
```

```

colnames(j147.entrezid.gene.conversion)[1] <- "gene_name"

#Make a temporary table containing the gene names
temp <- as.data.frame(rownames(j147.toptag$table))
temp$logFc <- j147.toptag$table$logFC
colnames(temp)[1] <- "gene_name"

#merge the data
j147.entrezid.logFC <- merge(j147.entrezid.gene.conversion, temp, by="gene_name")

j147.de <- c(j147.entrezid.logFC$logFc)
names(j147.de) <- c(j147.entrezid.logFC$entrez)

#Do the same for the CAD-31 data
cad31.entrezid.gene <- read.table("cad31_entrezid_gene.txt", header = T, sep = "\t", quote = "")
cad31.entrezid.gene.conversion <- as.data.frame(cad31.entrezid.gene$From)
cad31.entrezid.gene.conversion$entrez <- cad31.entrezid.gene$To
colnames(cad31.entrezid.gene.conversion)[1] <- "gene_name"

temp <- as.data.frame(rownames(cad31.AD.cad31.toptag$table))
temp$logFc <- cad31.AD.cad31.toptag$table$logFC
colnames(temp)[1] <- "gene_name"

cad31.entrezid.logFC <- merge(cad31.entrezid.gene.conversion, temp, by="gene_name")

cad31.de <- c(cad31.entrezid.logFC$logFc)
names(cad31.de) <- c(cad31.entrezid.logFC$entrez)

```

4.3.3 Pathway analysis

```

# Process all signaling pathways to see if they are inhibited or activated
j147.spia.result <- spia(de=j147.de, all=j147.entrezid.logFC$entrez, organism="mmu", plots=TRUE)

##
## Done pathway 1 : RNA transport..
##
## Done pathway 2 : RNA degradation..
##
## Done pathway 3 : PPAR signaling pathway..
## Done pathway 4 : Fanconi anemia pathway..
##
## Done pathway 5 : MAPK signaling pathway..
##
## Done pathway 6 : ErbB signaling pathway..
##
## Done pathway 7 : Calcium signaling pathway..
## Done pathway 8 : Cytokine-cytokine receptor int..
## Done pathway 9 : Chemokine signaling pathway..
## Done pathway 10 : NF-kappa B signaling pathway..

##

```

```

## Done pathway 11 : Phosphatidylinositol signaling..
##
## Done pathway 12 : Neuroactive ligand-receptor in..
## Done pathway 13 : Cell cycle..
## Done pathway 14 : Oocyte meiosis..

##
## Done pathway 15 : p53 signaling pathway..
## Done pathway 16 : Sulfur relay system..
## Done pathway 17 : SNARE interactions in vesicula..
## Done pathway 18 : Regulation of autophagy..

##
## Done pathway 19 : Protein processing in endoplas..
## Done pathway 20 : Lysosome..
## Done pathway 21 : mTOR signaling pathway..
## Done pathway 22 : Apoptosis..

##
## Done pathway 23 : Vascular smooth muscle contrac..
## Done pathway 24 : Wnt signaling pathway..
## Done pathway 25 : Dorso-ventral axis formation..
## Done pathway 26 : Notch signaling pathway..
## Done pathway 27 : Hedgehog signaling pathway..
## Done pathway 28 : TGF-beta signaling pathway..
## Done pathway 29 : Axon guidance..
## Done pathway 30 : VEGF signaling pathway..
## Done pathway 31 : Osteoclast differentiation..

##
## Done pathway 32 : Focal adhesion..

##
## Done pathway 33 : ECM-receptor interaction..
## Done pathway 34 : Cell adhesion molecules (CAMs)..
## Done pathway 35 : Adherens junction..
## Done pathway 36 : Tight junction..
## Done pathway 37 : Gap junction..
## Done pathway 38 : Complement and coagulation cas..
## Done pathway 39 : Antigen processing and present..
## Done pathway 40 : Toll-like receptor signaling p..
## Done pathway 41 : NOD-like receptor signaling pa..
## Done pathway 42 : RIG-I-like receptor signaling ..
## Done pathway 43 : Cytosolic DNA-sensing pathway..

##
## Done pathway 44 : Jak-STAT signaling pathway..
## Done pathway 45 : Natural killer cell mediated c..

##
## Done pathway 46 : T cell receptor signaling path..
## Done pathway 47 : B cell receptor signaling path..
## Done pathway 48 : Fc epsilon RI signaling pathwa..
## Done pathway 49 : Fc gamma R-mediated phagocytos..
## Done pathway 50 : Leukocyte transendothelial mig..
## Done pathway 51 : Intestinal immune network for ..
## Done pathway 52 : Circadian rhythm - mammal..

```

```

##
## Done pathway 53 : Long-term potentiation..

##
## Done pathway 54 : Neurotrophin signaling pathway..
## Done pathway 55 : Retrograde endocannabinoid sig..

##
## Done pathway 56 : Glutamatergic synapse..

##
## Done pathway 57 : Cholinergic synapse..
## Done pathway 58 : Serotonergic synapse..
## Done pathway 59 : GABAergic synapse..

##
## Done pathway 60 : Dopaminergic synapse..
## Done pathway 61 : Long-term depression..
## Done pathway 62 : Olfactory transduction..
## Done pathway 63 : Taste transduction..
## Done pathway 64 : Phototransduction..

##
## Done pathway 65 : Regulation of actin cytoskelet..

##
## Done pathway 66 : Insulin signaling pathway..
## Done pathway 67 : GnRH signaling pathway..
## Done pathway 68 : Progesterone-mediated oocyte m..

##
## Done pathway 69 : Melanogenesis..
## Done pathway 70 : Adipocytokine signaling pathwa..

##
## Done pathway 71 : Type II diabetes mellitus..
## Done pathway 72 : Type I diabetes mellitus..
## Done pathway 73 : Maturity onset diabetes of the..
## Done pathway 74 : Aldosterone-regulated sodium r..
## Done pathway 75 : Endocrine and other factor-reg..

##
## Done pathway 76 : Vasopressin-regulated water re..

##
## Done pathway 77 : Salivary secretion..
## Done pathway 78 : Gastric acid secretion..

##
## Done pathway 79 : Pancreatic secretion..
## Done pathway 80 : Carbohydrate digestion and abs..
## Done pathway 81 : Bile secretion..
## Done pathway 82 : Mineral absorption..

##
## Done pathway 83 : Alzheimer's disease..
## Done pathway 84 : Parkinson's disease..

##
## Done pathway 85 : Amyotrophic lateral sclerosis ..

```

```

##
## Done pathway 86 : Huntington's disease..
## Done pathway 87 : Prion diseases..

##
## Done pathway 88 : Cocaine addiction..

##
## Done pathway 89 : Amphetamine addiction..
## Done pathway 90 : Morphine addiction..

##
## Done pathway 91 : Alcoholism..

##
## Done pathway 92 : Bacterial invasion of epitheli..
## Done pathway 93 : Salmonella infection..
## Done pathway 94 : Pertussis..
## Done pathway 95 : Legionellosis..
## Done pathway 96 : Leishmaniasis..
## Done pathway 97 : Chagas disease (American trypa..
## Done pathway 98 : African trypanosomiasis..
## Done pathway 99 : Malaria..
## Done pathway 100 : Toxoplasmosis..
## Done pathway 101 : Amoebiasis..
## Done pathway 102 : Staphylococcus aureus infectio..
## Done pathway 103 : Tuberculosis..
## Done pathway 104 : Hepatitis C..
## Done pathway 105 : Measles..

##
## Done pathway 106 : Influenza A..

##
## Done pathway 107 : HTLV-I infection..
## Done pathway 108 : Herpes simplex infection..
## Done pathway 109 : Epstein-Barr virus infection..
## Done pathway 110 : Pathways in cancer..
## Done pathway 111 : Transcriptional misregulation ..

##
## Done pathway 112 : Viral carcinogenesis..
## Done pathway 113 : Colorectal cancer..
## Done pathway 114 : Renal cell carcinoma..
## Done pathway 115 : Pancreatic cancer..
## Done pathway 116 : Endometrial cancer..
## Done pathway 117 : Glioma..
## Done pathway 118 : Prostate cancer..
## Done pathway 119 : Thyroid cancer..
## Done pathway 120 : Basal cell carcinoma..
## Done pathway 121 : Melanoma..
## Done pathway 122 : Bladder cancer..

##
## Done pathway 123 : Chronic myeloid leukemia..
## Done pathway 124 : Acute myeloid leukemia..
## Done pathway 125 : Small cell lung cancer..
## Done pathway 126 : Non-small cell lung cancer..

```

```

## Done pathway 127 : Asthma..
## Done pathway 128 : Autoimmune thyroid disease..

##
## Done pathway 129 : Systemic lupus erythematosus..
## Done pathway 130 : Rheumatoid arthritis..
## Done pathway 131 : Allograft rejection..
## Done pathway 132 : Graft-versus-host disease..
## Done pathway 133 : Arrhythmogenic right ventricul..
## Done pathway 134 : Dilated cardiomyopathy..
## Done pathway 135 : Viral myocarditis..

cad31.spia.result <- spia(de=cad31.de, all=cad31.entrezid.logFC$entrez, organism="mmu", plots=TRUE)

##
## Done pathway 1 : RNA transport..

##
## Done pathway 2 : RNA degradation..

##
## Done pathway 3 : PPAR signaling pathway..

##
## Done pathway 4 : Fanconi anemia pathway..

##
## Done pathway 5 : MAPK signaling pathway..

##
## Done pathway 6 : ErbB signaling pathway..

##
## Done pathway 7 : Calcium signaling pathway..

##
## Done pathway 8 : Cytokine-cytokine receptor int..

##
## Done pathway 9 : Chemokine signaling pathway..

##
## Done pathway 10 : NF-kappa B signaling pathway..

##
## Done pathway 11 : Phosphatidylinositol signaling..

##
## Done pathway 12 : Neuroactive ligand-receptor in..

##
## Done pathway 13 : Cell cycle..

##
## Done pathway 14 : Oocyte meiosis..

##
## Done pathway 15 : p53 signaling pathway..
## Done pathway 16 : Sulfur relay system..

##
## Done pathway 17 : SNARE interactions in vesicula..
## Done pathway 18 : Regulation of autophagy..

```

```
##  
## Done pathway 19 : Protein processing in endoplas..  
##  
## Done pathway 20 : Lysosome..  
##  
## Done pathway 21 : mTOR signaling pathway..  
##  
## Done pathway 22 : Apoptosis..  
##  
## Done pathway 23 : Vascular smooth muscle contrac..  
##  
## Done pathway 24 : Wnt signaling pathway..  
##  
## Done pathway 25 : Dorso-ventral axis formation..  
##  
## Done pathway 26 : Notch signaling pathway..  
##  
## Done pathway 27 : Hedgehog signaling pathway..  
##  
## Done pathway 28 : TGF-beta signaling pathway..  
##  
## Done pathway 29 : Axon guidance..  
##  
## Done pathway 30 : VEGF signaling pathway..  
##  
## Done pathway 31 : Osteoclast differentiation..  
##  
## Done pathway 32 : Focal adhesion..  
##  
## Done pathway 33 : ECM-receptor interaction..  
## Done pathway 34 : Cell adhesion molecules (CAMs)..  
## Done pathway 35 : Adherens junction..  
##  
## Done pathway 36 : Tight junction..  
##  
## Done pathway 37 : Gap junction..  
##  
## Done pathway 38 : Complement and coagulation cas..  
##  
## Done pathway 39 : Antigen processing and present..  
##  
## Done pathway 40 : Toll-like receptor signaling p..  
##  
## Done pathway 41 : NOD-like receptor signaling pa..
```

```
##  
## Done pathway 42 : RIG-I-like receptor signaling ..  
##  
## Done pathway 43 : Cytosolic DNA-sensing pathway..  
##  
## Done pathway 44 : Jak-STAT signaling pathway..  
##  
## Done pathway 45 : Natural killer cell mediated c..  
##  
## Done pathway 46 : T cell receptor signaling path..  
## Done pathway 47 : B cell receptor signaling path..  
##  
## Done pathway 48 : Fc epsilon RI signaling pathwa..  
##  
## Done pathway 49 : Fc gamma R-mediated phagocytos..  
##  
## Done pathway 50 : Leukocyte transendothelial mig..  
##  
## Done pathway 51 : Intestinal immune network for ..  
##  
## Done pathway 52 : Circadian rhythm - mammal..  
##  
## Done pathway 53 : Long-term potentiation..  
##  
## Done pathway 54 : Neurotrophin signaling pathway..  
##  
## Done pathway 55 : Retrograde endocannabinoid sig..  
##  
## Done pathway 56 : Glutamatergic synapse..  
##  
## Done pathway 57 : Cholinergic synapse..  
##  
## Done pathway 58 : Serotonergic synapse..  
##  
## Done pathway 59 : GABAergic synapse..  
##  
## Done pathway 60 : Dopaminergic synapse..  
##  
## Done pathway 61 : Long-term depression..  
##  
## Done pathway 62 : Olfactory transduction..  
##  
## Done pathway 63 : Taste transduction..
```

```
##  
## Done pathway 64 : Phototransduction..  
##  
## Done pathway 65 : Regulation of actin cytoskelet..  
##  
## Done pathway 66 : Insulin signaling pathway..  
##  
## Done pathway 67 : GnRH signaling pathway..  
##  
## Done pathway 68 : Progesterone-mediated oocyte m..  
##  
## Done pathway 69 : Melanogenesis..  
##  
## Done pathway 70 : Adipocytokine signaling pathwa..  
##  
## Done pathway 71 : Type II diabetes mellitus..  
## Done pathway 72 : Type I diabetes mellitus..  
## Done pathway 73 : Maturity onset diabetes of the..  
##  
## Done pathway 74 : Aldosterone-regulated sodium r..  
##  
## Done pathway 75 : Endocrine and other factor-reg..  
##  
## Done pathway 76 : Vasopressin-regulated water re..  
##  
## Done pathway 77 : Salivary secretion..  
##  
## Done pathway 78 : Gastric acid secretion..  
##  
## Done pathway 79 : Pancreatic secretion..  
## Done pathway 80 : Carbohydrate digestion and abs..  
##  
## Done pathway 81 : Bile secretion..  
##  
## Done pathway 82 : Mineral absorption..  
##  
## Done pathway 83 : Alzheimer's disease..  
##  
## Done pathway 84 : Parkinson's disease..  
##  
## Done pathway 85 : Amyotrophic lateral sclerosis ..  
##  
## Done pathway 86 : Huntington's disease..
```

```

##
## Done pathway 87 : Prion diseases..

##
## Done pathway 88 : Cocaine addiction..

##
## Done pathway 89 : Amphetamine addiction..

##
## Done pathway 90 : Morphine addiction..

##
## Done pathway 91 : Alcoholism..

##
## Done pathway 92 : Bacterial invasion of epitheli..

##
## Done pathway 93 : Salmonella infection..

##
## Done pathway 94 : Pertussis..

##
## Done pathway 95 : Legionellosis..

##
## Done pathway 96 : Leishmaniasis..
## Done pathway 97 : Chagas disease (American trypa..

##
## Done pathway 98 : African trypanosomiasis..
## Done pathway 99 : Malaria..

##
## Done pathway 100 : Toxoplasmosis..

##
## Done pathway 101 : Amoebiasis..
## Done pathway 102 : Staphylococcus aureus infectio..

##
## Done pathway 103 : Tuberculosis..

##
## Done pathway 104 : Hepatitis C..
## Done pathway 105 : Measles..

##
## Done pathway 106 : Influenza A..

##
## Done pathway 107 : HTLV-I infection..

##
## Done pathway 108 : Herpes simplex infection..

##
## Done pathway 109 : Epstein-Barr virus infection..
## Done pathway 110 : Pathways in cancer..

```

```

## 
## Done pathway 111 : Transcriptional misregulation ..

## 
## Done pathway 112 : Viral carcinogenesis..

## 
## Done pathway 113 : Colorectal cancer..

## 
## Done pathway 114 : Renal cell carcinoma..

## 
## Done pathway 115 : Pancreatic cancer..

## 
## Done pathway 116 : Endometrial cancer..

## 
## Done pathway 117 : Glioma..
## Done pathway 118 : Prostate cancer..
## Done pathway 119 : Thyroid cancer..
## Done pathway 120 : Basal cell carcinoma..

## 
## Done pathway 121 : Melanoma..
## Done pathway 122 : Bladder cancer..

## 
## Done pathway 123 : Chronic myeloid leukemia..

## 
## Done pathway 124 : Acute myeloid leukemia..

## 
## Done pathway 125 : Small cell lung cancer..

## 
## Done pathway 126 : Non-small cell lung cancer..
## Done pathway 127 : Asthma..
## Done pathway 128 : Autoimmune thyroid disease..

## 
## Done pathway 129 : Systemic lupus erythematosus..

## 
## Done pathway 130 : Rheumatoid arthritis..
## Done pathway 131 : Allograft rejection..
## Done pathway 132 : Graft-versus-host disease..

## 
## Done pathway 133 : Arrhythmogenic right ventricul..

## 
## Done pathway 134 : Dilated cardiomyopathy..

## 
## Done pathway 135 : Viral myocarditis..

```

```
j147.spia.result[1:5,]
```

4.3.3.1 J147 pathway results

```

##                               Name      ID pSize NDE      tA pPERT
## 1       Calcium signaling pathway 04020      2   2     1 -0.6386118 0.129
## 2       ErbB signaling pathway 04012      2   2     1 -1.5028731 0.133
## 3       Alzheimer's disease 05010      2   2     1  0.7346170 0.228
## 4       Jak-STAT signaling pathway 04630      1   1     1  0.3109694 0.334
## 5 T cell receptor signaling pathway 04660      1   1     1  0.5692345 0.345
##          pG pGFdr pGFWER      Status
## 1 0.3931846      1     1 Inhibited
## 2 0.4013150      1     1 Inhibited
## 3 0.5650774      1     1 Activated
## 4 0.7002692      1     1 Activated
## 5 0.7121527      1     1 Activated
##
##                                KEGGLINK
## 1 http://www.genome.jp/dbget-bin/show_pathway?mmu04020+13869+12290
## 2 http://www.genome.jp/dbget-bin/show_pathway?mmu04012+13869+12402
## 3 http://www.genome.jp/dbget-bin/show_pathway?mmu05010+14812+16956
## 4      http://www.genome.jp/dbget-bin/show_pathway?mmu04630+12402
## 5      http://www.genome.jp/dbget-bin/show_pathway?mmu04660+12402

```

Here we see the top five most impacted pathway's, very nice to see Alzheimer's disease on third place. Confirming the effect of J147 on the disease.

```
cad31.spia.result[1:5,]
```

4.3.3.2 CAD-31 pathway results

```

##                               Name      ID pSize NDE
## 1       Cholinergic synapse 04725      11   11
## 2       RIG-I-like receptor signaling pathway 04622      3     3
## 3 Endocrine and other factor-regulated calcium reabsorption 04961      3     3
## 4       Toll-like receptor signaling pathway 04620      3     3
## 5       Acute myeloid leukemia 05221      2     2
##          pNDE      tA pPERT      pG pGFdr pGFWER      Status
## 1      1 7.970573477 0.054 0.2116136      1     1 Activated
## 2      1 -4.986830151 0.058 0.2231441      1     1 Inhibited
## 3      1 3.768058304 0.088 0.3018768      1     1 Activated
## 4      1 0.008109592 0.131 0.3972651      1     1 Activated
## 5      1 -1.119277395 0.146 0.4269257      1     1 Inhibited
##
## 1 http://www.genome.jp/dbget-bin/show_pathway?mmu04725+108015+11513+432530+14682+18798+16438+12286+1
## 2
## 3
## 4
## 5

```

In the top five pathways's in of the cad-31 study Alzheimer's disease is not present, but it is further down the lis (39th) meaning significant alteration. We do see Cholinergic synapse on number one, this pathway also plays a role in AD, thus also in this pathway analysis a lot of evidence for the drugs involvement in AD.

5 Quick overview

Its hard to draw conclusions from the study, but the CAD-31 drug seems to affect a lot more genes than its counterpart. Further experiments are needed to determine if these drugs are viable candidates to treat

alzheimer disease. But so far they show promising results.