**R-script for TMM Normilization and edgeR analysis (with FDR cutoff 0.05)**

Reference: edgeR: differential expression analysis of digital gene expression data, User’s Guide, Yunshun Chen, Davis McCarthy, Matthew Ritchie, Mark Robinson, Gordon K. Smyth.

library(edgeR) #loding edgeR

# Importing data

a <- read.csv("/Path/to/the/file/Raw\_counts.csv", header = T, sep="\t", check.names=FALSE)

head(a)

rownames(a) <- a$Probe\_Description

# Removing columns of other time points

a[1] <- NULL

head(a)

a[5:16] <- NULL

head(a)

# Grouping of male and female || duplicates (replace 78 with other time points as necessary)

gro <- c(rep("78F",2), rep("78M",2))

# Creating DEGList object

a1 <- DGEList(counts = a, group =gro)

a1

# Filtering reads with CPM less than 4 for 2 samples

keep <- rowSums(cpm(a1)>4) >= 2

table(keep)

a2 <- a1[keep, , keep.lib.sizes=FALSE]

# Normaliztion

a3 <- calcNormFactors(a2, method="TMM")

a3

a3$samples

#write out Normalization factors and filtered data

write.csv(a3$counts, file="78H\_filtered\_dataset.xls", quote = FALSE)

write.csv(a3$samples, file="78H\_Norm\_factor\_filtered.xls", quote= FALSE)

# Data Exploration

plotMDS(a3)

b1 <- estimateCommonDisp(a3)

b2 <- estimateTagwiseDisp(b1)

#exact test

et <- exactTest(b2)

etp <- topTags(et, n=300000)

summary(etp$table)

etp$table$logFC = -etp$table$logFC

pdf("edgeR-MA\_78\_Normalized-plot.pdf")

plot(

 etp$table$logCPM,

 etp$table$logFC, main="78hrs MA Plot",

 xlab="logCPM", ylab="logFC",

 xlim=c(-3, 20), ylim=c(-12, 12), pch=20, cex=.3,

 col = ifelse( etp$table$FDR < .05, "red", "black" ) )

abline(h=c(-1,1), col="blue")

dev.off()

# output CSV

write.csv(etp$table, "edgeR-78M\_vs\_78F\_Normalized.csv", quote=FALSE)

**R-script for generating Quantile analysis for Unpaired data Figure 3**

# Remove all variables and get input.csv data

rm(list=ls(all=TRUE))

source("/home/keerthana/Desktop/Quartile\_analysis/Gopinath/FPKM\_QUARTILE/NEW/Qbox.R")

dat <- read.csv("FPKM\_Unpaired\_changed.csv",header=TRUE,sep = ",")

#head(dat)

# Create output directory

outdir <- "/home/keerthana/Desktop/Quartile\_analysis/Gopinath/FPKM\_QUARTILE/NEW/"

# Setup plotting parameters

png(filename = paste(outdir, "unpaired\_all\_in\_one.png"), units = "in", w= 15, h = 6, res = 300)

par(mfrow = c(1,4))

yLim = c(-4,8)

Qbx(male = log(dat$X78Male),female = log(dat$X78Female), ylim = yLim, crit = "pmax", sub = "Quartile\_max", main = "Quartiles based max of \n 78h male or female data"  )

Qbx(male = log(dat$X96Male), female = log(dat$X96Female), ylim = yLim, crit = "pmax", sub = "Quartile\_max", main = "Quartiles based max of \n 96h male or female data"  )

Qbx(male = log(dat$X120Male), female = log(dat$X120Female), ylim = yLim, crit = "pmax", sub = "Quartile\_max", main = "Quartiles based max of \n 120h male or female data"  )

Qbx(male = log(dat$Head\_Male), female = log(dat$Head\_Female), ylim = yLim, crit = "pmax", sub = "Quartile\_max", main = "Quartiles based max of \n head male or female data"  )

dev.off()

**R-script for generating Quantile analysis for Paired data Figure 6**

# Remove all variables and get input.csv data

rm(list=ls(all=TRUE))

source("/home/keerthana/Desktop/Quartile\_analysis/Gopinath/FPKM\_QUARTILE/NEW/Qbox.R")

dat <- read.csv("FPKM\_newpaired\_changed.csv",header=TRUE,sep = ",")

#head(dat)

# Create output directory

outdir <- "/home/keerthana/Desktop/Quartile\_analysis/Gopinath/FPKM\_QUARTILE/NEW/"

# Setup plotting parameters

png(filename = paste(outdir, "paired\_all\_in\_one.png"), units = "in", w= 15, h = 6, res = 300)

par(mfrow = c(1,4))

yLim = c(-4,8)

Qbx(male = log(dat$X78Male),female = log(dat$X78Female), ylim = yLim, crit = "pmax", sub = "Quartile\_max", main = "Quartiles based max of \n 78h male or female data"  )

Qbx(male = log(dat$X96Male), female = log(dat$X96Female), ylim = yLim, crit = "pmax", sub = "Quartile\_max", main = "Quartiles based max of \n 96h male or female data"  )

Qbx(male = log(dat$X120Male), female = log(dat$X120Female), ylim = yLim, crit = "pmax", sub = "Quartile\_max", main = "Quartiles based max of \n 120h male or female data"  )

Qbx(male = log(dat$Head\_Male), female = log(dat$Head\_Female), ylim = yLim, crit = "pmax", sub = "Quartile\_max", main = "Quartiles based max of \n head male or female data"  )

dev.off()