#Normalization Microarray data

```
if (!requireNamespace("BiocManager", quietly = TRUE))
    install.packages("BiocManager")
BiocManager::install("limma")
library(limma)
```

```
File->Change dir
library(limma)
targets <- readTargets("SwirlSample.txt")
rg<-read.maimages(targets, source="spot")
rg$genes<-readGAL()
rg$printer<-getLayout(rg$genes)
imageplot(rg$R[,1], rg$printer, low="white", high="red")
imageplot(rg$Rb[,1], rg$printer, low="white", high="red")
red<-(rg$R[,1]-rg$Rb[,1])
green<-(rg$G[,1]-rg$Gb[,1])
plot(green,red,main=" Raw Scatter plot")
hist(green)
hist(log2(green))
# Produce a scatter plot for log2(green) vs log2(red)
plot((green+red)/2,red-green,main="MA Plot")
MA <- normalizeWithinArrays(rg, method="none")
plotDensities(MA)
plotPrintTipLoess(MA)
MA <- normalizeWithinArrays(rg)
plotPrintTipLoess(MA)
plotDensities(MA)
boxplot(MA$M)
MA <- normalizeBetweenArrays(MA,method="scale")
boxplot(MA$M~col(MA$M),names=colnames(MA$M))
```

#Identification of differentially Expressed Genes from Spotted Array Data

```
use both swirl and coral data
Write your interpretation based on the gene functionalities
Commands
library(limma)
targets<- readTargets("SwirlSample.txt")
RG <- read.maimages(targets, source="spot")
RG$genes<- readGAL("fish.gal")
RG$printer<- getLayout(RG$genes)
MA <- normalizeWithinArrays(RG, method="loess")
MA2 <- normalizeBetweenArrays(MA,method="scale")
design <- modelMatrix(targets, ref="wild type")
fit<-lmFit(MA2,design)
```

```
fit2<-eBayes(fit)
volcanoplot(fit2,highlight=20,names=fit2$genes$Name)
topTable(fit2,number=20,adjust='BH')
```

```
#Identification of Differentially Expressed Genes (DEGs) From RNAseq data
if (!requireNamespace("BiocManager", quietly = TRUE))
    install.packages("BiocManager")

BiocManager::install("edgeR")
```

```
library(edgeR)
counts <- read.csv("samples.csv", row.names = 1)
data1 <- DGEList(counts)
data1 <- calcNormFactors(data1)
cutoff <- 5
drop <- which(apply(cpm(data1), 1, max) < cutoff)
d <- data1[-drop,]
dim(d)
snames <- colnames(counts)
group<-interaction(substr(snames, 1, nchar(snames) - 2))
plotMDS(d, col = as.numeric(group))
mm <- model.matrix(~0 + group)
y <- voom(d, mm, plot = T)
fit <- lmFit(y, mm)
contr <- makeContrasts(groupLum.A. - groupTNBC., levels =
colnames(coef(fit)))
tmp <- contrasts.fit(fit, contr)
tmp <- eBayes(tmp)
top.table <- topTable(tmp, sort.by = "P", n = Inf)
head(top.table, 20)
```

#Identification of DEGs from Affymetrix Array data

```
dat = read.csv("bc.csv", header=T)
library(affy)
Dat = ReadAffy(filenames=paste(dat$Samples,".cel",sep=""))
Dat
annotation(Dat)
phenoData(Dat)
sample.names = dat$Tumor
colnames(exprs(Dat)) = sample.names
image(Dat)
boxplot(Dat,col=c(rep("Green",6),rep("Blue",6), rep("red",6)))
MAplot(Dat[,c(1,8)], pair=T)
```

```
eset = rma(Dat)
dim(exprs(eset))
exprs(eset)[1:20,1:3]
boxplot(data.frame(exprs(eset)),col=c(rep("Green",6),rep("Blue",6),rep("re
d",6)))
library(limma)
design <- model.matrix(~ 0+factor(c(rep("apocrine",6), rep("basal",6),
rep("luminal",6))))
colnames(design) <- c("apocrine", "basal", "luminal")
design
fit=lmFit(eset,design)
cont.matrix=makeContrasts(Comp2to1=basal-apocrine,
Comp3to1=luminal-apocrine,Comp3to2=luminal-basal,levels=design)
fit2=contrasts.fit(fit,cont.matrix)
fit2=eBayes(fit2)
options(digits=3)
topTable(fit2,coef=1,adjust="BH")
volcanoplot(fit2,coef=2, highlight=10)
volcanoplot(fit2,coef=1, highlight=10)
results = decideTests(fit2)
vennDiagram(results,include=c("up","down"), counts.col=c("red","green"))
```

Clustering of Gene Expression Data
Library(pvclust)

data(lung)

lung[1:3,1:5]

result <- pvclust(lung, method.dist="cor",
method.hclust="average", nboot=1000)

plot(result)

pvrect(result, alpha=0.95)