

Class17

Nancy Leon-Rivera PID A12925462

#Downstream analysis

```
folders <- list.files(pattern = "_quant")  
files <- paste0(folders, "/abundance.h5")
```

```
file.exists(files)
```

```
[1] FALSE
```

```
names(files) <- sub("_quant", "", folders)  
files
```

```
      <NA>  
"/abundance.h5"
```

Load up the tximport library

```
library(tximport)  
#Files not working  
#txi.kallisto <- tximport(files, type = "kallisto", txOut = T)
```

#Remove zero count genes

Before subsequent analysis, we might want to filter out those annotated transcripts with no reads:

```
#to.keep <- rowSums(txi.kallisto$counts) > 0  
#kset.nonzero <- txi.kallisto$counts[to.keep,]
```

```
#nrow(kset.nonzero)
```

And those with no change over the samples:

```
#keep2 <- apply(kset.nonzero,1,sd)>0  
#x <- kset.nonzero[keep2,]
```

#Try a PCA

```
#pca <- prcomp(t(x), scale=T)  
#summary(pca)
```

```
library(ggplot2)  
library(ggrepel)  
#mycols <- c("blue","blue", "red", "red")  
#ggplot(pca$x)+  
#   aes(PC1, PC2)+  
#   geom_point(col=mycols)+  
#   geom_label_repel(label=rownames(pca$x), col=mycols)
```

#DESeq

```
library(DESeq2)
```

Loading required package: S4Vectors

Loading required package: stats4

Loading required package: BiocGenerics

Attaching package: 'BiocGenerics'

The following objects are masked from 'package:stats':

IQR, mad, sd, var, xtabs

The following objects are masked from 'package:base':

```
anyDuplicated, aperm, 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, saveRDS, setdiff,  
table, tapply, union, unique, unsplit, which.max, which.min
```

Attaching package: 'S4Vectors'

The following object is masked from 'package:utils':

```
findMatches
```

The following objects are masked from 'package:base':

```
expand.grid, I, unname
```

Loading required package: IRanges

Attaching package: 'IRanges'

The following object is masked from 'package:grDevices':

```
windows
```

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, colAvgPerRowSet, 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, rowAvgPerColSet,
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

```
#sampleTable <- data.frame(condition = factor(rep(c("control", "treatment"), each = 2)))  
#rownames(sampleTable) <- colnames(txi.kallisto$counts)
```

```
#sampleTable
```

```
#dds <- DESeqDataSetFromTximport(txi.kallisto,  
#                               sampleTable,  
#                               ~condition)
```

```
#dds <- DESeq(dds)  
#res <- results(dds)  
#res
```