NSBB552 Final Project (2021) - Phoebe Chum

"Cerebrovascular miRNA Profile in Alzheimer's Disease"

Approach

Mice Model

Triple transgenic Alzheimer's disease "3xTg-AD" mouse model with three human genes Amyloid-beta precursor protein (APP), Presenilin 1 (PSEN1), and Microtubule-associated protein tau (MAPT).

Tissue Type

Whole brain cerebral vessels.

Study Design

- Sample groups:
  - Young control (YC; 1-2 mo)
  - Cognitive impairment (CI; 4-5 mo)
  - Amyloid-beta (AB; 6-8 mo)
  - AB+Tau (ABT; greater than 12 mo).
- n = 3 males and 3 females for each of the four groups.
- Total RNA was extracted from the vessels and sent to NanoString for miRNA expression panel.
- miRNA expression was then analyzed with DESeq2.

Hypothesis

Cerebrovascular microRNA (miRNA) expression profiles corresponding to post-transcriptional regulation can provide a diagnostic map of the early development of AD pathology.

Packages

```
library(DESeq2)
```

```
## Loading required package: S4Vectors
## Loading required package: stats4
## Loading required package: BiocGenerics
## Loading required package: parallel
##
## Attaching package: 'BiocGenerics'
```

```
## The following objects are masked from 'package:parallel':
##
##     clusterApply, clusterApplyLB, clusterCall, clusterEvalQ,
##     clusterExport, clusterMap, parApply, parCapply, parLapply,
##     parLapplyLB, parRapply, parSapply, parSapplyLB

## The following objects are masked from 'package:stats':
##
##     IQR, mad, sd, var, xtabs

## The following objects are masked from 'package:base':
##
##     anyDuplicated, 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, setdiff, sort, table, tapply,
##     union, unique, unsplit, which.max, which.min

##
## Attaching package: 'S4Vectors'

## The following objects are masked from 'package:base':
##
##     expand.grid, I, unname

## Loading required package: IRanges

## 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, colAvgsPerRowSet, 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, rowAvgsPerColSet,
##     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
```

```
library(pheatmap)
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
```

```
library(ggplot2)
```

Resources

Tutorial:

1. https://www.youtube.com/watch?v=wPzeea1Do18

Documentation:

1. http://bioconductor.org/packages/devel/bioc/vignettes/DESeq2/inst/doc/DESeq2.html#countmat
2. https://leaherb.com/add-gradient-colors-to-bar-chart-with-ggplot/
3. https://stackoverflow.com/questions/48633302/reordering-bars-in-a-group-with-ggplot-and-adjusting-groups-by-width

Input Data

Load count matrix and coldata

```
cts <-read.csv("Raw Data_Formatted.csv", header = T, row.names = 1)

coldata <-read.csv("MetaData.csv", header = T, row.names = 1)
```

```
head(cts,10)
```

```
##             Mouse1 Mouse2 Mouse3 Mouse4 Mouse5 Mouse6 Mouse7 Mouse8 Mouse9
## mmu-let-7a     479      1     46     18    356     74      1     89      1
## mmu-let-7b    3377    382    927    363   2547   1141   1591   3194   1420
## mmu-let-7c    3846    193    707    218   2615    774    843   2184    855
## mmu-let-7d    2183    198    639    215   1628    768    837   2414    623
## mmu-let-7e     868    100    230     86    597    197    379    737    289
## mmu-let-7f     177     21     34      9    143     42     92    131     75
## mmu-let-7g    1790     98    371    158   1340    728    637   1630    404
## mmu-let-7i     285     13     58     28    195    146    125    261     69
## mmu-miR-1       48      8     11      3     22     12     20     64     13
## mmu-miR-100    129     16     31     19    123     49     72    168     49
##             Mouse10 Mouse11 Mouse12 Mouse13 Mouse14 Mouse15 Mouse16 Mouse17
## mmu-let-7a      698     234     828     274       1     114     706    1041
## mmu-let-7b     8878    3284    6309    3709    1686    2641    5907    7479
## mmu-let-7c     7024    2882    6182    2852     954    2081    4738    6414
## mmu-let-7d     7041    2693    5331    2327     602    1518    2963    4131
## mmu-let-7e     1751     927    1551     859     280     637    1214    1556
## mmu-let-7f      276     181     350     194      68     148     263     429
## mmu-let-7g     4937    2725    4538    1760     361    1287    2356    4580
## mmu-let-7i     1020     488     864     404      87     233     373     812
## mmu-miR-1        94      56      87      47       9      46      66     156
## mmu-miR-100     335     142     253     181      50     127     193     398
##             Mouse18 Mouse19 Mouse20 Mouse21 Mouse22 Mouse23 Mouse24
## mmu-let-7a      540      62     120       1     178     132     583
## mmu-let-7b     4268    2141    3498    1722    3590    2493    6595
## mmu-let-7c     4413    1532    2075    1158    2766    2024    5025
## mmu-let-7d     2738     960    1629     748    1764    1459    4325
## mmu-let-7e     1239     487     823     344     749     615    1440
## mmu-let-7f      222     111     140      86     140     125     261
## mmu-let-7g     2183     726    1042     567    2012    1075    3377
## mmu-let-7i      304     197     230      96     301     217     617
## mmu-miR-1        45      25      34      20      44      17      70
## mmu-miR-100     198      90     126      75     115      99     222
```

```
coldata
```

```
##         Age Sex
## Mouse1  MCI   F
## Mouse2  MCI   F
## Mouse3  MCI   F
## Mouse4  MCI   M
## Mouse5  MCI   M
## Mouse6  MCI   M
## Mouse7   YC   F
## Mouse8   YC   F
## Mouse9   YC   F
## Mouse10  YC   M
## Mouse11  YC   M
## Mouse12  YC   M
## Mouse13  AB   F
## Mouse14  AB   F
## Mouse15  AB   F
## Mouse16  AB   M
## Mouse17  AB   M
## Mouse18  AB   M
```

```
## Mouse19 ABT   F
## Mouse20 ABT   F
## Mouse21 ABT   F
## Mouse22 ABT   M
## Mouse23 ABT   M
## Mouse24 ABT   M
```

Build an DESeqDataSet from a count matrix and a table of sample information

```
dds <- DESeqDataSetFromMatrix(countData = cts,
                              colData = coldata,
                              design = ~Age + Sex)
```

```
## Warning in DESeqDataSet(se, design = design, ignoreRank): some variables in
## design formula are characters, converting to factors
```

Pre-filtering the dataset/remove rows with zero count

```
dds <- dds[ rowSums(counts(dds)) > 12, ]
```

Differential expression analysis with

```
ddsDE <-DESeq(dds)
```

```
## estimating size factors
```

```
## estimating dispersions
```

```
## gene-wise dispersion estimates
```

```
## mean-dispersion relationship
```

```
## -- note: fitType='parametric', but the dispersion trend was not well captured by the
##    function: y = a/x + b, and a local regression fit was automatically substituted.
##    specify fitType='local' or 'mean' to avoid this message next time.
```

```
## final dispersion estimates
```

```
## fitting model and testing
```

Export normalized read counts

```
normCounts <-counts(ddsDE, normalized = T)
write.csv(normCounts, "normal.all4.csv")
```

Extract result

specify the two groups for comparison

```
res1 <- results(ddsDE, contrast=c("Age", "MCI", "YC"))
res1
```

MCI vs YC

```
## log2 fold change (MLE): Age MCI vs YC
## Wald test p-value: Age MCI vs YC
## DataFrame with 599 rows and 6 columns
##                   baseMean log2FoldChange     lfcSE      stat    pvalue
##                  <numeric>      <numeric> <numeric> <numeric> <numeric>
## mmu-let-7a          220.43      1.5084559  1.049197  1.437725  0.150512
## mmu-let-7b         2737.70      0.2227796  0.302961  0.735340  0.462133
```

```
## mmu-let-7c         2219.55      0.4479122  0.393308  1.138834  0.254773
## mmu-let-7d         1729.28      0.0707983  0.370180  0.191254  0.848327
## mmu-let-7e          624.94      0.2906758  0.316603  0.918109  0.358562
## ...                    ...            ...       ...       ...       ...
## mghv-miR-M1-6      5.78055     -0.1672098  0.514006 -0.325307  0.744949
## mghv-miR-M1-7-3p  25.26609     -0.0840302  0.229111 -0.366767  0.713793
## mghv-miR-M1-7-5p  20.57505      0.2788862  0.244618  1.140091  0.254249
## mghv-miR-M1-8     26.78637     -0.3006724  0.238567 -1.260327  0.207551
## mghv-miR-M1-9      6.73086     -0.7971861  0.523232 -1.523581  0.127613
##                       padj
##                  <numeric>
## mmu-let-7a        0.873083
## mmu-let-7b        0.886469
## mmu-let-7c        0.873083
## mmu-let-7d        0.986410
## mmu-let-7e        0.873083
## ...                    ...
## mghv-miR-M1-6     0.961690
## mghv-miR-M1-7-3p  0.960818
## mghv-miR-M1-7-5p  0.873083
## mghv-miR-M1-8     0.873083
## mghv-miR-M1-9     0.873083
```

```
summary(res1)
```

```
##
## out of 599 with nonzero total read count
## adjusted p-value < 0.1
## LFC > 0 (up)       : 1, 0.17%
## LFC < 0 (down)     : 0, 0%
## outliers [1]       : 0, 0%
## low counts [2]     : 0, 0%
## (mean count < 1)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

```
resultsNames(ddsDE)
```

```
## [1] "Intercept"     "Age_ABT_vs_AB" "Age_MCI_vs_AB" "Age_YC_vs_AB"
## [5] "Sex_M_vs_F"
```

```
res2 <- results(ddsDE, contrast=c("Age", "AB", "YC"))
res2
```

AB vs YC

```
## log2 fold change (MLE): Age AB vs YC
## Wald test p-value: Age AB vs YC
## DataFrame with 599 rows and 6 columns
##                   baseMean log2FoldChange     lfcSE       stat    pvalue
##                  <numeric>      <numeric> <numeric>  <numeric> <numeric>
## mmu-let-7a          220.43      0.7218213  1.048067  0.6887170  0.491001
## mmu-let-7b         2737.70     -0.1184722  0.302529 -0.3916054  0.695350
## mmu-let-7c         2219.55     -0.0146732  0.392940 -0.0373421  0.970212
## mmu-let-7d         1729.28     -0.4964872  0.369668 -1.3430639  0.179251
```

```
## mmu-let-7e          624.94     -0.1327430  0.314917 -0.4215172  0.673377
## ...                    ...            ...       ...        ...       ...
## mghv-miR-M1-6      5.78055     0.09960922  0.365107  0.2728217  0.784990
## mghv-miR-M1-7-3p  25.26609    -0.21844094  0.170360 -1.2822346  0.199760
## mghv-miR-M1-7-5p  20.57505    -0.03119659  0.186220 -0.1675254  0.866957
## mghv-miR-M1-8     26.78637    -0.00609449  0.163632 -0.0372451  0.970290
## mghv-miR-M1-9      6.73086    -0.44557248  0.362786 -1.2281979  0.219373
##                       padj
##                  <numeric>
## mmu-let-7a        0.949709
## mmu-let-7b        0.949709
## mmu-let-7c        0.993418
## mmu-let-7d        0.890748
## mmu-let-7e        0.949709
## ...                    ...
## mghv-miR-M1-6     0.949709
## mghv-miR-M1-7-3p  0.890748
## mghv-miR-M1-7-5p  0.977979
## mghv-miR-M1-8     0.993418
## mghv-miR-M1-9     0.892959
```

```
summary(res2)
```

```
##
## out of 599 with nonzero total read count
## adjusted p-value < 0.1
## LFC > 0 (up)       : 2, 0.33%
## LFC < 0 (down)     : 0, 0%
## outliers [1]       : 0, 0%
## low counts [2]     : 0, 0%
## (mean count < 1)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

```
res3 <- results(ddsDE, contrast=c("Age", "ABT", "YC"))
res3
```

ABT vs YC

```
## log2 fold change (MLE): Age ABT vs YC
## Wald test p-value: Age ABT vs YC
## DataFrame with 599 rows and 6 columns
##                   baseMean log2FoldChange     lfcSE      stat    pvalue
##                  <numeric>      <numeric> <numeric> <numeric> <numeric>
## mmu-let-7a          220.43      -0.457605  1.049107 -0.436185 0.6627025
## mmu-let-7b         2737.70      -0.403170  0.302565 -1.332506 0.1826938
## mmu-let-7c         2219.55      -0.472276  0.393000 -1.201721 0.2294718
## mmu-let-7d         1729.28      -0.881635  0.369744 -2.384449 0.0171047
## mmu-let-7e          624.94      -0.434710  0.315077 -1.379694 0.1676808
## ...                    ...            ...       ...        ...       ...
## mghv-miR-M1-6      5.78055      0.1720641  0.363121  0.473848 0.6356081
## mghv-miR-M1-7-3p  25.26609     -0.3083543  0.172381 -1.788791 0.0736485
## mghv-miR-M1-7-5p  20.57505     -0.0532274  0.187090 -0.284501 0.7760266
## mghv-miR-M1-8     26.78637      0.0297027  0.163289  0.181903 0.8556586
```

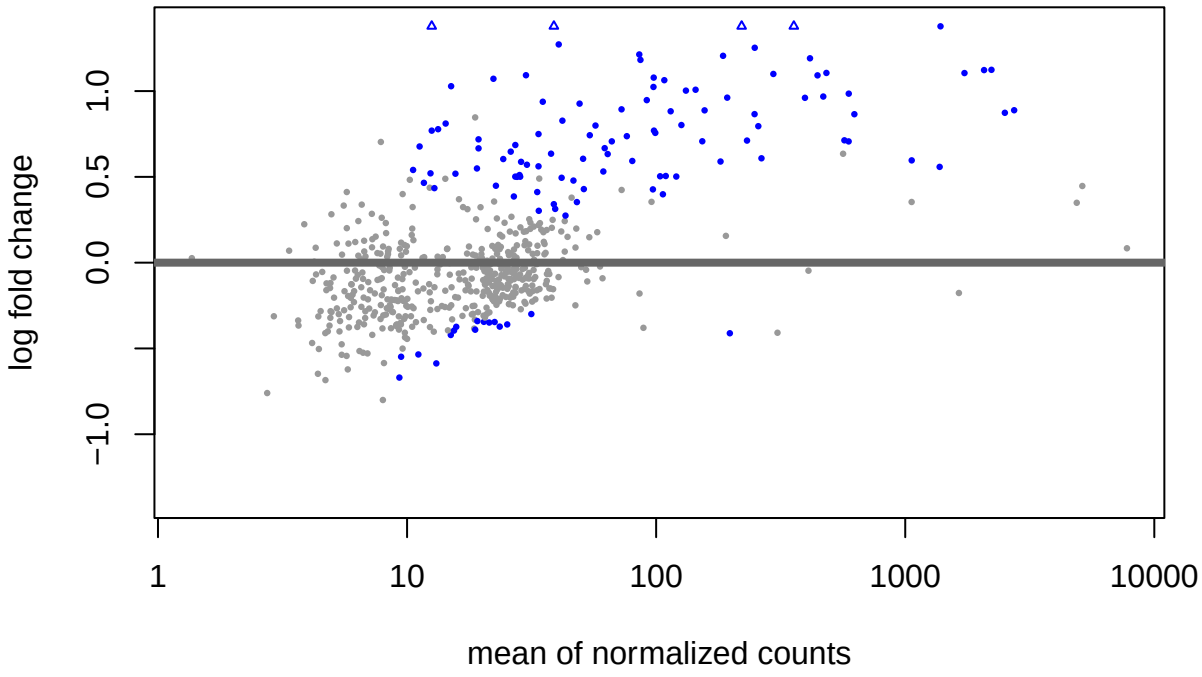
```
## mghv-miR-M1-9      6.73086     -0.6103078  0.369032 -1.653808 0.0981666
##                       padj
##                  <numeric>
## mmu-let-7a        0.914649
## mmu-let-7b        0.644263
## mmu-let-7c        0.683281
## mmu-let-7d        0.368006
## mmu-let-7e        0.612444
## ...                    ...
## mghv-miR-M1-6     0.914649
## mghv-miR-M1-7-3p  0.509852
## mghv-miR-M1-7-5p  0.946749
## mghv-miR-M1-8     0.969636
## mghv-miR-M1-9     0.538780
```

```
summary(res3)
```

```
##
## out of 599 with nonzero total read count
## adjusted p-value < 0.1
## LFC > 0 (up)       : 1, 0.17%
## LFC < 0 (down)     : 0, 0%
## outliers [1]       : 0, 0%
## low counts [2]     : 0, 0%
## (mean count < 1)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

Exploratory Data Analysis (EDA)

```
plotMA(ddsDE)
```



Data Transformation

```
dds <- estimateSizeFactors(dds)
dds <- estimateDispersions(dds)
```

```
## gene-wise dispersion estimates
```

```
## mean-dispersion relationship
```

```
## -- note: fitType='parametric', but the dispersion trend was not well captured by the
##    function: y = a/x + b, and a local regression fit was automatically substituted.
##    specify fitType='local' or 'mean' to avoid this message next time.
```

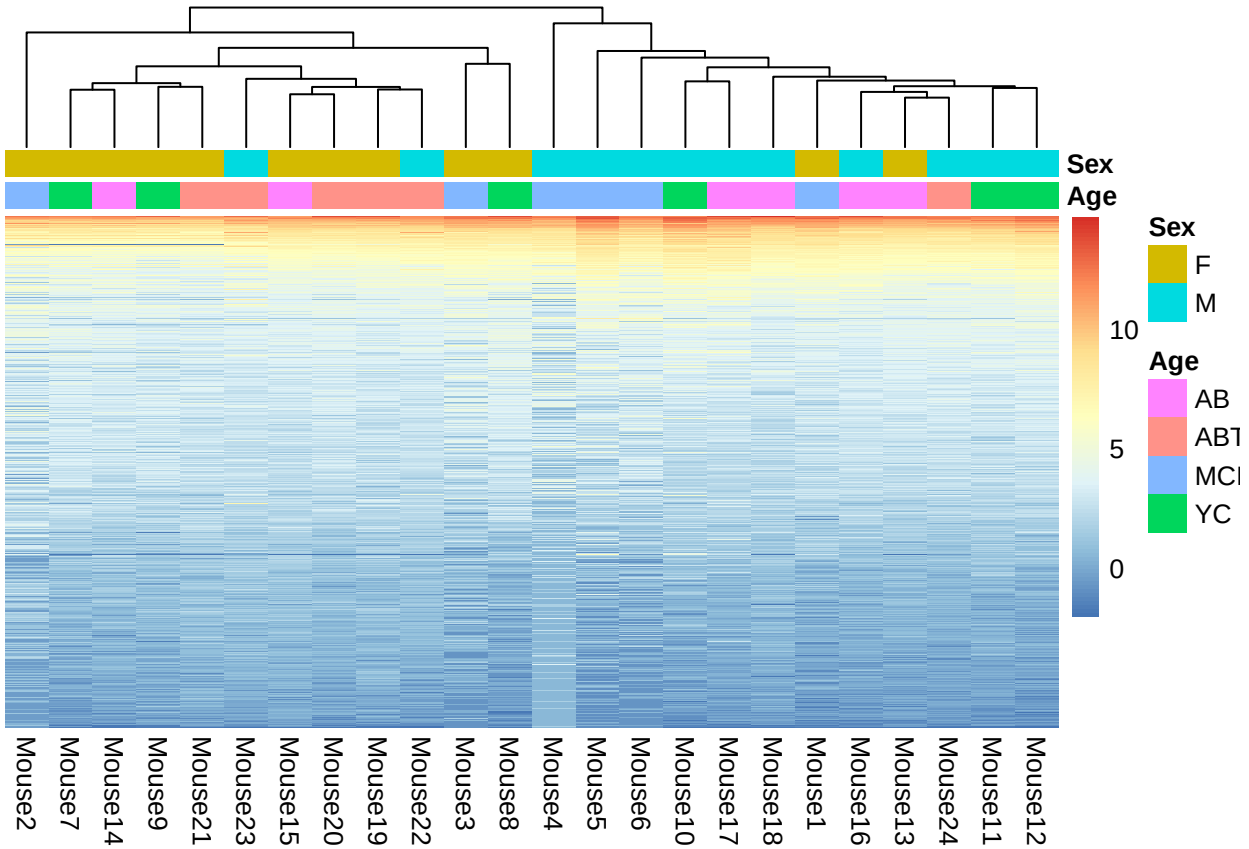
```
## final dispersion estimates
```

```
vsd <- varianceStabilizingTransformation(dds)
```

```
## -- note: fitType='parametric', but the dispersion trend was not well captured by the
##    function: y = a/x + b, and a local regression fit was automatically substituted.
##    specify fitType='local' or 'mean' to avoid this message next time.
```

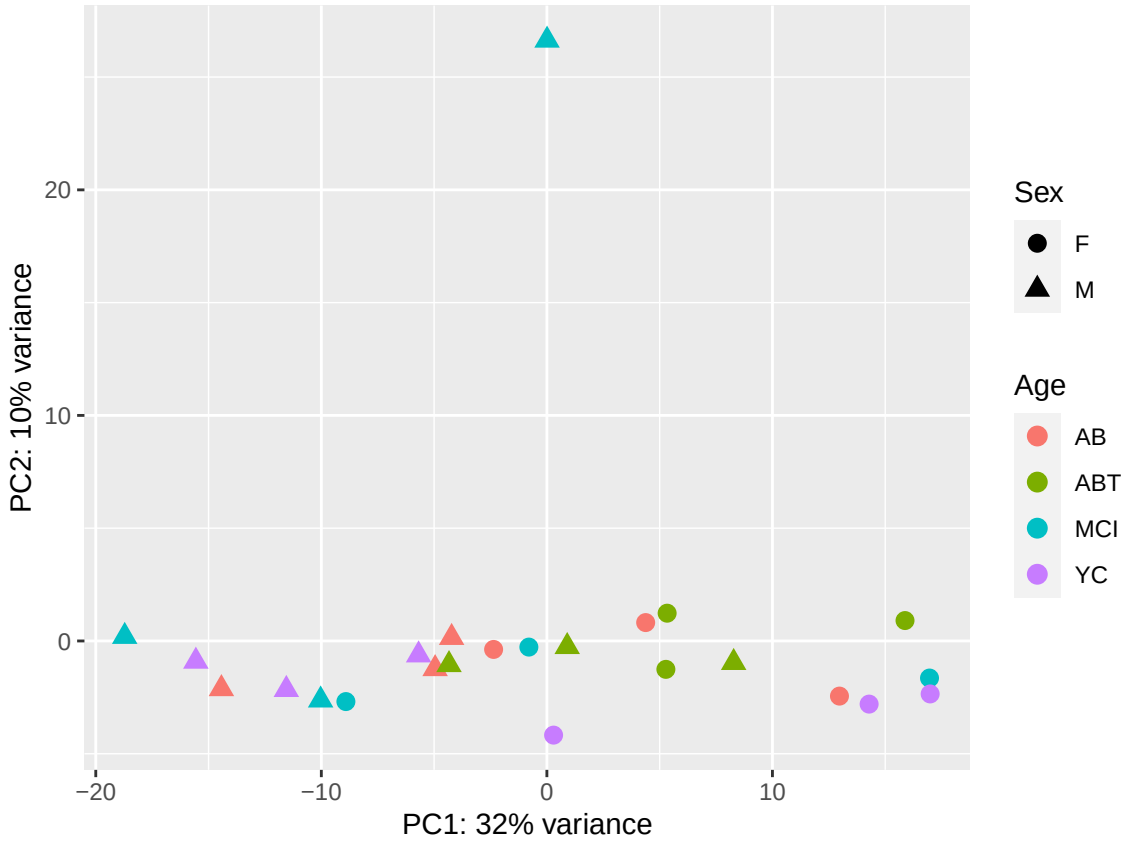
Heatmap

```
library(pheatmap)
select <- order(rowMeans(counts(dds,normalized=TRUE)),
                decreasing=TRUE)
df <- as.data.frame(colData(dds)[,c("Age","Sex")])
pheatmap(assay(vsd)[select,], cluster_rows=FALSE, show_rownames=FALSE,
         cluster_cols=TRUE, annotation_col=df)
```



PCA

```
library(ggplot2)
pcaData <- plotPCA(vsd, intgroup=c("Age", "Sex"), returnData=TRUE)
percentVar <- round(100 * attr(pcaData, "percentVar"))
ggplot(pcaData, aes(PC1, PC2, color=Age, shape=Sex)) +
  geom_point(size=3) +
  xlab(paste0("PC1: ",percentVar[1],"% variance")) +
  ylab(paste0("PC2: ",percentVar[2],"% variance")) +
  coord_fixed()
```



Log2 Fold Change Comparison

MCI vs YC

Convert result into data frame, the filter

```
res1 <- as.data.frame(res1)
res1p <- res1[res1$pvalue<0.05,]
res1padj <- res1[res1$padj<0.05,]
res1logUp <- res1[res1$log2FoldChange>=1,]
res1logDown <- res1[res1$log2FoldChange<=-1,]

res1logUp
```

```
##                baseMean log2FoldChange     lfcSE     stat       pvalue
## mmu-let-7a   220.430307       1.508456 1.0491967 1.437725 1.505122e-01
## mmu-miR-144    7.848560       1.205343 0.4694926 2.567331 1.024848e-02
```

```
## mmu-miR-155    1.367435       1.779606 0.9330175 1.907366 5.647317e-02
## mmu-miR-1947   3.360925       1.290694 0.6737196 1.915773 5.539394e-02
## mmu-miR-2134  35.062603       2.085505 0.5101870 4.087727 4.356212e-05
## mmu-miR-429   38.854111       1.392389 0.6001107 2.320221 2.032893e-02
## mmu-miR-543   12.558773       1.009322 0.9047509 1.115580 2.646020e-01
##                     padj
## mmu-let-7a   0.87308305
## mmu-miR-144  0.78679676
## mmu-miR-155  0.80812046
## mmu-miR-1947 0.80812046
## mmu-miR-2134 0.02609371
## mmu-miR-429  0.79582372
## mmu-miR-543  0.87308305
```

```
res1logDown
```

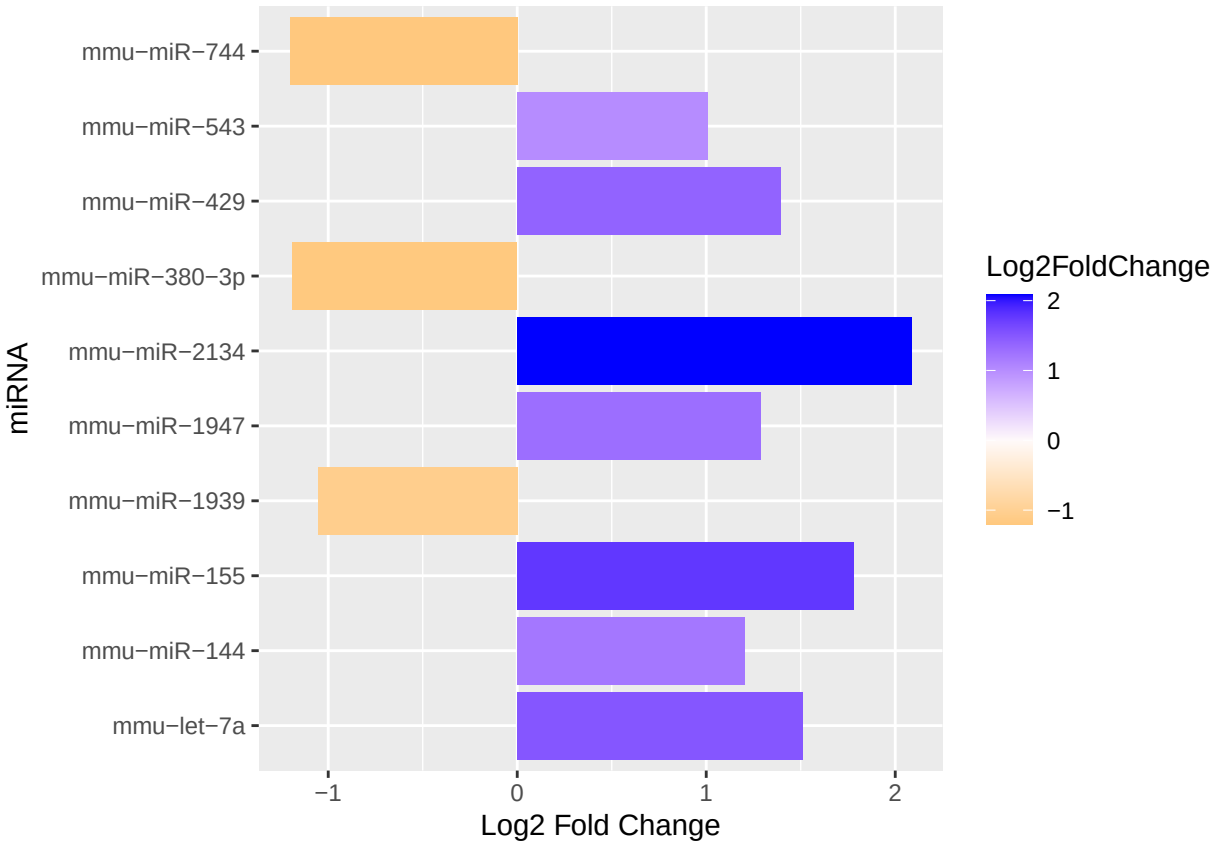
```
##                baseMean log2FoldChange     lfcSE      stat     pvalue      padj
## mmu-miR-1939   4.925136      -1.054161 0.6595054 -1.598411 0.10995149 0.8662072
## mmu-miR-380-3p 8.622660      -1.189481 0.4781933 -2.487449 0.01286629 0.7867968
## mmu-miR-744    4.920099      -1.203220 0.6489005 -1.854244 0.06370430 0.8081205
```

Comments:

log2 Fold Change Comparison	Criteria	Up-regulated	Down-regulated
MCI vs YC (n=6 v 6)	Up-regulated: Fold-change >=1 Down-regulated: Fold-change <=-1	7 miRNAs	3 miRNAs

```
MCI_log2 <-read.csv("MCI vs YC_log2.csv", header = T, row.names = 1)

ggplot(data = MCI_log2,
       aes(x = miRNA, y = Log2FoldChange,
           )) +
  geom_bar(stat = "identity", aes(fill = Log2FoldChange), position=position_dodge()) +
  ylab("Log2 Fold Change") +
  coord_flip() + scale_fill_gradient2(low='orange', mid='snow', high='blue')
```



AB vs YC

Convert result into data frame, the filter

```
res2 <- as.data.frame(res2)
res2p <- res2[res2$pvalue<0.05,]
res2padj <- res2[res2$padj<0.05,]
res2logUp <- res2[res2$log2FoldChange>=1,]
res2logDown <- res2[res2$log2FoldChange<=-1,]

res2logUp
```

```
##                  baseMean log2FoldChange     lfcSE     stat       pvalue
## mmu-miR-144      7.848560       1.027796 0.3943233 2.606482 9.147775e-03
## mmu-miR-291a-5p  5.065948       2.728232 0.5576312 4.892538 9.954385e-07
## mmu-miR-690     72.668587       1.081045 0.2790151 3.874505 1.068416e-04
##                         padj
## mmu-miR-144     0.6062519146
## mmu-miR-291a-5p 0.0005962677
## mmu-miR-690     0.0319990628
```

```
res2logDown
```

```
##             baseMean log2FoldChange     lfcSE      stat    pvalue      padj
## mmu-miR-543 12.55877      -1.936399 0.9228382 -2.098308 0.0358779 0.7410643
```

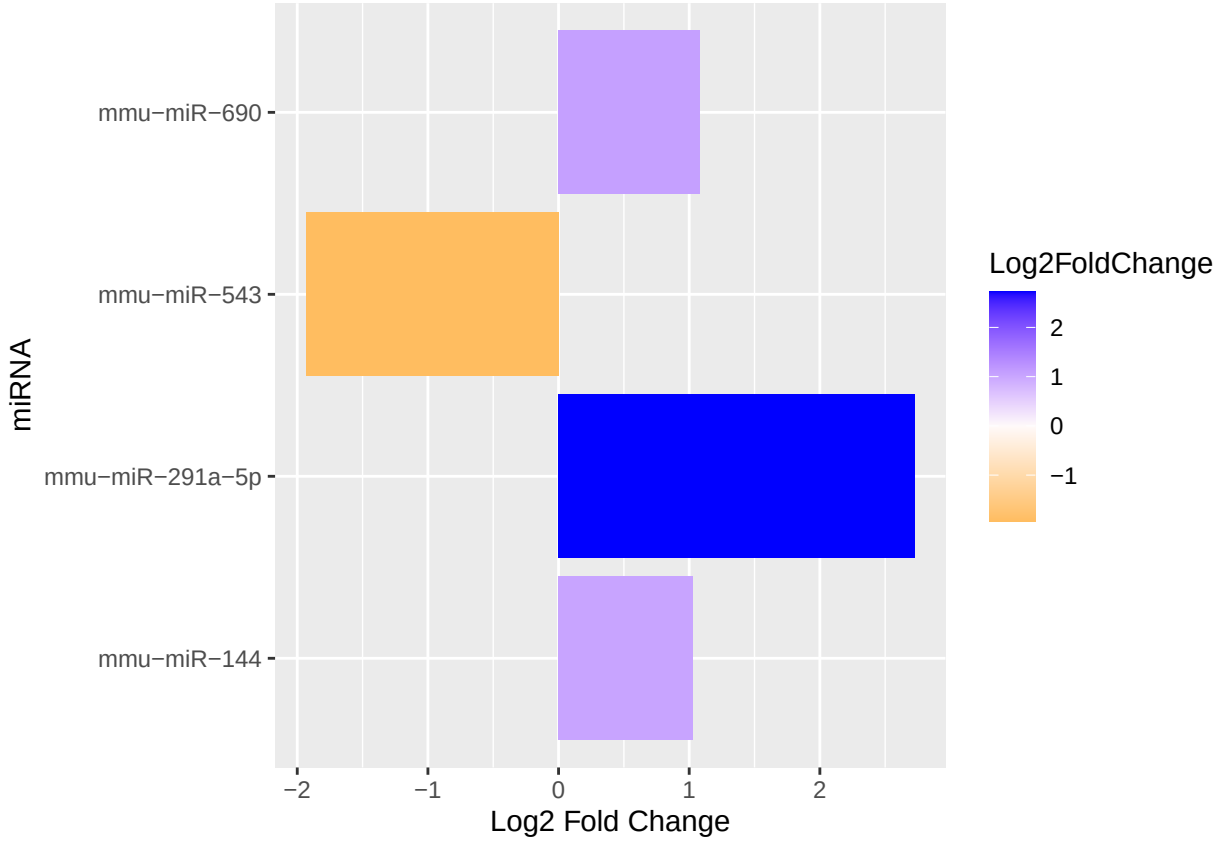
Comments:

log2 Fold Change Comparison	Criteria	Up-regulated	Down-regulated
AB vs YC (n=6 v 6)	Up-regulated: Fold-change >=1 Down-regulated: Fold-change <=-1	3 miRNAs	1 miRNAs

```
AB_log2 <-read.csv("AB vs YC_log2.csv", header = T)
```

```
## Warning in read.table(file = file, header = header, sep = sep, quote = quote, :
## incomplete final line found by readTableHeader on 'AB vs YC_log2.csv'
```

```
ggplot(data = AB_log2,
       aes(x = miRNA, y = Log2FoldChange,
           )) +
  geom_bar(stat = "identity", aes(fill = Log2FoldChange), position=position_dodge()) +
  ylab("Log2 Fold Change") +
  coord_flip() + scale_fill_gradient2(low='orange', mid='snow', high='blue')
```



ABT vs YC

Convert result into data frame, the filter

```
res3 <- as.data.frame(res3)
res3p <- res3[res3$pvalue<0.05,]
res3padj <- res3[res3$padj<0.05,]
res3logUp <- res3[res3$log2FoldChange>=1,]
res3logDown <- res3[res3$log2FoldChange<=-1,]
```

```
res3logUp
```

```
##                  baseMean log2FoldChange     lfcSE     stat       pvalue
## mmu-miR-183     18.766750       1.490320 0.5836500 2.553449 1.066620e-02
## mmu-miR-1932     5.463301       1.045087 0.3984303 2.623010 8.715675e-03
## mmu-miR-291a-5p  5.065948       2.569964 0.5606528 4.583879 4.564294e-06
## mmu-miR-295      4.181448       1.059709 0.4949342 2.141110 3.226514e-02
## mmu-miR-429     38.854111       1.014256 0.5865755 1.729114 8.378877e-02
##                        padj
## mmu-miR-183     0.354319640
## mmu-miR-1932    0.348045973
## mmu-miR-291a-5p 0.002734012
## mmu-miR-295     0.449163975
## mmu-miR-429     0.509852291
```

```
res3logDown
```

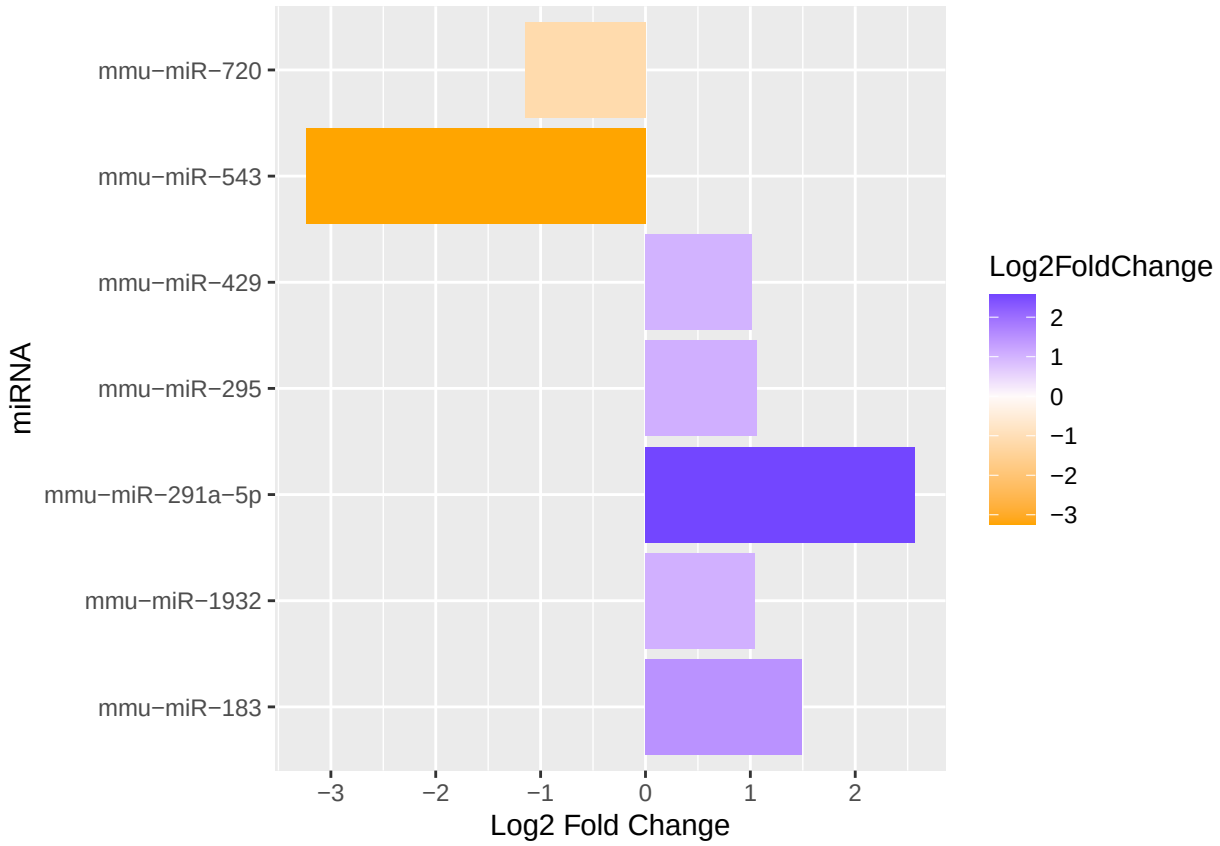
```
##               baseMean log2FoldChange     lfcSE      stat       pvalue      padj
## mmu-miR-543   12.55877      -3.241302 0.9754260 -3.322961 0.0008906742 0.1628592
## mmu-miR-720 563.09853      -1.149735 0.4621617 -2.487733 0.0128560191 0.3560996
```

Comments:

log2 Fold Change Comparison	Criteria	Up-regulated	Down-regulated
ABT vs YC (n=6 v 6)	Up-regulated: Fold-change >=1 Down-regulated: Fold-change <=-1	5 miRNAs	2 miRNAs

```
ABT_log2 <-read.csv("ABT vs YC_log2.csv", header = T)

ggplot(data = ABT_log2,
       aes(x = miRNA, y = Log2FoldChange,
           )) +
  geom_bar(stat = "identity", aes(fill = Log2FoldChange), position=position_dodge()) +
  ylab("Log2 Fold Change") +
  coord_flip() + scale_fill_gradient2(low='orange', mid='snow', high='blue')
```



Summary

Log2 Fold Change: Up-regulated

MCI vs YC	AB vs YC	ABT vs YC
mmu-let-7a		mmu-miR-183
mmu-miR-144	**mmu-miR-144**	mmu-miR-1932
mmu-miR-155	**mmu-miR-291a-5p**	**mmu-miR-291a-5p**
mmu-miR-1947	mmu-miR-690	mmu-miR-295
mmu-miR-2134		
mmu-miR-429		**mmu-miR-429**
mmu-miR-543		

Literature Search

- Elevated Levels of miR-144-3p Induce Cholinergic Degeneration by Impairing the Maturation of NGF in Alzheimer's Disease (doi: 10.3389/fcell.2021.667412)
- Knockdown of miR-429 Attenuates A -Induced Neuronal Damage by Targeting SOX2 and BCL2 in Mouse Cortical Neurons (doi: 10.1007/s11064-018-2643-3)

Log2 Fold Change: Down-regulated

MCI vs YC	AB vs YC	ABT vs YC
mmu-miR-1939		
mmu-miR-380-3p	***mmu-miR-543***	***mmu-miR-543***

MCI vs YC	AB vs YC	ABT vs YC
mmu-miR-744		mmu-miR-720

Literature Search

- miR-543: PubMed search - 211 results

Conclusion

Observation

For this project, I used DESeq2 to analyze NanoString miRNA count data. The data did not show clear clustering with heatmap and PCA plot. PC1 shows 32% variance and PC2 shows 10% variance. Neither of the principle component suggests a strong distinction to separate the sample groups. Log2 fold change suggests some upregulated and downregulated miRNAs. In particular, I was interested in any miRNAs expressions with at least 2-fold differences. According to the summary table shown above, mmu-miR-144 marks the transition between mild cognitive impairment (MCI) and Amyloid-beta (A) stage whereas mmu-miR-429 marks the transition between A and Amyloid-beta + Tau (A T)stage. Interenstingly, mmu-miR-543 is upregulated in MCI, but is downregulated in A and A T.

Future Direction

Find out the pathways these miRNAs are involved in for further investigation of how they may contribute to the AD pathology.

Limitation

Due to smaller input (less than 600 miRNAs) of the dataset, the DESeq2 normalization may not have been the most appropriate tool to assess this dataset. Therefore, I have also explored another package, NanoStringDiff (doi:%2010.1093/nar/gkw677). However, I have yet to figure out the appropriate design input to receive a result from this package (Error in Beta.full %*% contrast : non-conformable arguments).

NanoStringDiff (Work in Progress) ## Data Input

directory <- "/Users/phoebechum/Desktop/NanoString/Data Analysis/R Analysis/Raw Data.csv"

designs = data.frame(group=c("MCI1","MCI2","MCI3","MCI4","MCI5","MCI6","YC1","YC2","YC3","YC4","YC5","YC6

library("NanoStringDiff")

NanoStringData=createNanoStringSetFromCsv(directory,header=TRUE,designs)

NanoStringData

pheno=pData(NanoStringData)

group=pheno$group

design.full=model.matrix(~0+group)

design.full

contrast = c(-1,1)

NanoStringData=estNormalizationFactors(NanoStringData)

positiveFactor(NanoStringData)

negativeFactor(NanoStringData)

housekeepingFactor(NanoStringData)

result=glm.LRT(NanoStringData,design.full,contrast=contrast)

head(result$table)

result

Pairwise Comparisons

endogenous=matrix(rpois(300,50),25,12)

colnames(endogenous)=paste("Sample", 1:12)

colnames(endogenous)=paste("Sample",1:12)

positive=matrix(rpois(72,c(128,32,8,2,0.5,0.125)*80),6,12)

negative=matrix(rpois(96,10),8,12)

housekeeping=matrix(rpois(36,100),3,12)

designs=data.frame(group=c(rep("YC",4),rep("MCI",4),rep("AB",4)),

- gender=rep(c("Male","Male","Female","Female"),3),
- age=c(20,40,39,37,29,47,23,45,34,65,35,64))

NanoStringData2=createNanoStringSet(endogenous,positive,

- negative,housekeeping,designs)

NanoStringData2