

# STAT646 - Homework 3

Daniel Osorio - dcosoriorh@tamu.edu  
Department of Veterinary Integrative Biosciences  
Texas A&M University

1. Read the minfi tutorial:  
<https://www.bioconductor.org/help/course-materials/2015/BioC2015/methylation450k.html>
2. Read the Breacher analysis2.pdf. Take the RGset data, the phenotype data (uploaded at eCampus) and perform the following steps:

- (a) Normalize the data using SWAN normalization.

```
> #load("rgset_breacher.Rdata")
```

- (b) Recreate all the plots from the Breacher\_analysis2.pdf file.

```
> phenoData <- read.csv(file = "Phenotype_Breacher_HM450.csv",  
+                       row.names = 1,  
+                       stringsAsFactors = FALSE)  
> excludeS <- c('879', '4667', '5505', '7725', '3F05_1_2',  
+               '3F05_10_2', 'Leuk_ZC', 'Leuk_FH', '3D03_10_2')  
> phenoData <- phenoData[!rownames(phenoData) %in% excludeS,]  
> bL <- phenoData[phenoData$sample == 1,]  
> colnames(bL)[7] <- "Breaches"  
> bL$Breaches[bL$Breaches == "1/9/2016"] <- "1-9"  
> bL$Breaches[bL$Breaches %in% c("400+", "200-399")] <- "200+"  
> bL$Breaches <- factor(x = bL$Breaches,  
+                       levels = c("0", "1-9", "10-39",  
+                                   "40-99", "100-199", "200+"))  
> par(mar=c(3,3,2,1), mgp = c(1.5,0.5,0), mfrow = c(1,2))  
> plot(x = bL$Breaches,  
+      main = "# Breaches",  
+      xlab = "Categories",  
+      ylab = "Breaches",  
+      las = 1)  
> pie(x = c('65.625%' = 70, '34.375%' = 30),  
+     col = c("red", "skyblue"))  
> legend("topright",  
+       legend = c("<39", ">39"),  
+       fill = c("red", "skyblue"), bty = "n")
```

