Generating\_Heatmap\_for\_Microbiomepaper.R

Pedro\_Torres

#Dataset used for Torres et al. 2017 Developmental stage influences metabolic and gut microbiome profile in PCOS mouse model. After doing group\_significance.py in qiime the bacterial genera that had a p>0.05 (after FDR correcting) in the adult model and seperately in the pubertal model were copied from /4week\_8week\_merged/taxa\_plots/merged\_sorted\_L6.txt and pasted into a new file called baccounts.csv and then this was used to create a heatmap for both develpomental stages.  
#install.packages("pheatmap")  
library(pheatmap)  
library(dplyr)

##   
## Attaching package: 'dplyr'

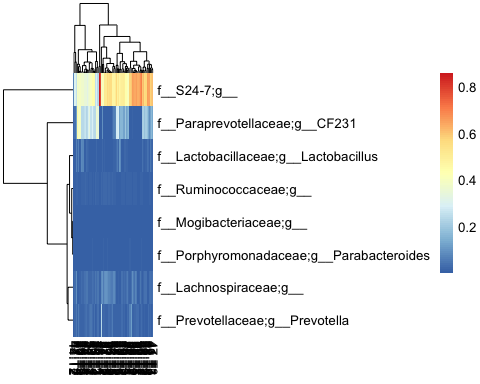
## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

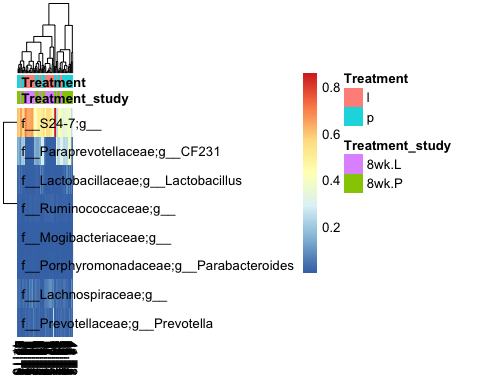
getwd()

## [1] "/Users/Pedro\_Torres/Desktop/PCOS\_Analysis/PCOS.16S.PhD/4weekand8week/heatmap"

setwd("/Users/Pedro\_Torres/Desktop/PCOS\_Analysis/PCOS.16S.PhD/4weekand8week/heatmap/heatmap8weekonly/")  
  
#Looking at week 8 (adult) mouse model only------------------------  
Bac.counts <- read.csv(file = "baccounts8week.csv", header = TRUE, row.names=1,check.names=FALSE)  
Bac.factors <- read.csv(file = "bacfactor8weeks.csv", header = TRUE, row.names=1,check.names=FALSE)  
  
#heatmap  
pheatmap(Bac.counts)



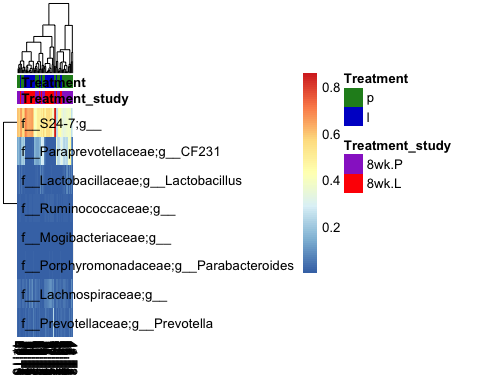
#colour sample groups  
Bac.factorsDS <- select(Bac.factors, Treatment\_study, Treatment)  
pheatmap(Bac.counts, annotation\_col = Bac.factorsDS)



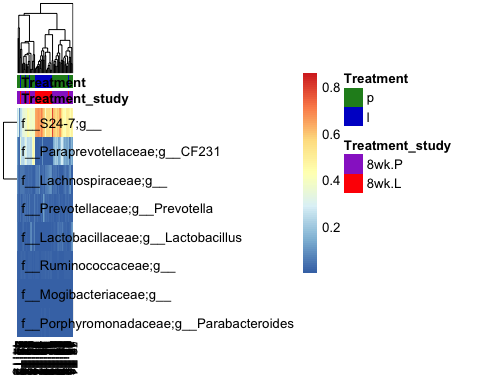
# Reorder Density levels to Sparse, Dense, Other  
Bac.factorsDS$Treatment\_study = factor(Bac.factorsDS$Treatment\_study, levels = c("8wk.P", "8wk.L"))  
DensityCol <- c("darkorchid", "red")  
names(DensityCol) <- levels(Bac.factorsDS$Treatment\_study)  
  
# Reorder TREATMENT to placebo and letrozole  
Bac.factorsDS$Treatment <- factor(Bac.factorsDS$Treatment, levels = c("p", "l"))  
  
SpeciesCol <- c("forestgreen", "blue3")  
names(SpeciesCol) <- levels(Bac.factorsDS$Treatment)  
  
# Add to a list, where names match those in factors dataframe  
AnnColour = list(Treatment\_study = DensityCol,Treatment = SpeciesCol)  
  
# Check the output  
AnnColour

## $Treatment\_study  
## 8wk.P 8wk.L   
## "darkorchid" "red"   
##   
## $Treatment  
## p l   
## "forestgreen" "blue3"

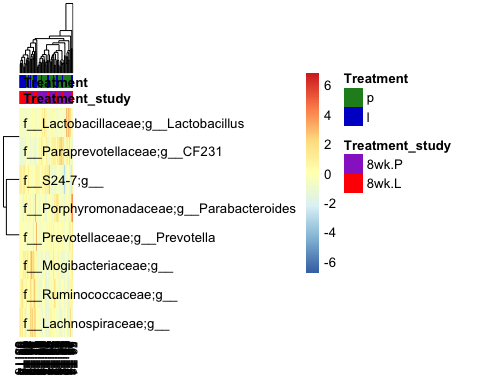
#redraw heatmap  
pheatmap(Bac.counts, annotation\_col = Bac.factorsDS, annotation\_colors = AnnColour)



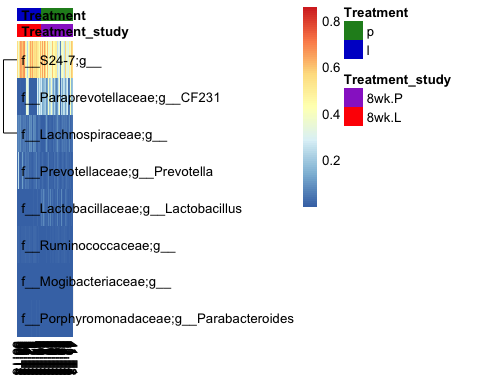
pheatmap(Bac.counts, clustering\_distance\_rows = "manhattan",  
 clustering\_distance\_cols = "manhattan", clustering\_method = 'average',  
 annotation\_colors = AnnColour, annotation\_col = Bac.factorsDS)



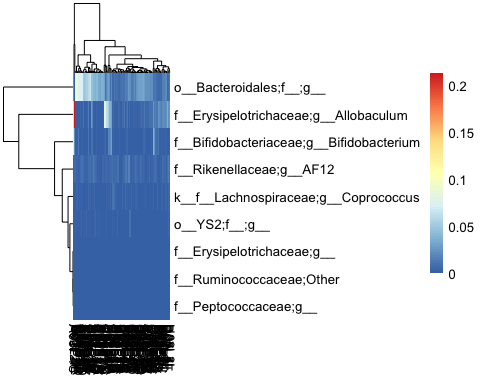
#scaling variables-We can now see how many standard deviations the Log10 abundance of a single OTU is away from the mean for that OTU in a sample compared only with other samples for that OTU  
pheatmap(Bac.counts, scale = "row", clustering\_distance\_rows = "manhattan",  
 clustering\_method = 'average',  
 annotation\_colors = AnnColour, annotation\_col = Bac.factorsDS)



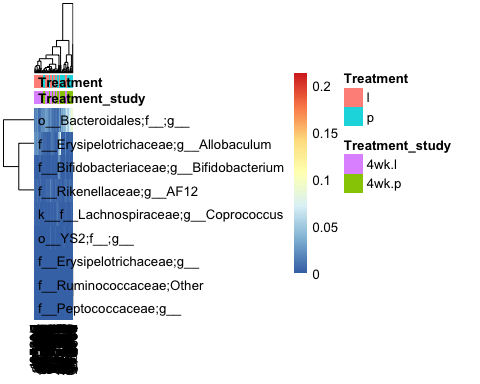
#sort by groups  
SampleOrder = order(Bac.factorsDS$Treatment, Bac.factorsDS$Treatment\_study)  
  
pheatmap(Bac.counts[ , SampleOrder], cluster\_cols = FALSE,  
 clustering\_method = 'average', annotation\_colors = AnnColour, annotation\_col = Bac.factorsDS)



#looking at 4 week (pubertal) mouse model only-----------------  
  
setwd("/Users/Pedro\_Torres/Desktop/PCOS\_Analysis/PCOS.16S.PhD/4weekand8week/heatmap/heatmap4weekonly/")  
  
Bac.counts <- read.csv(file = "bacterialcounts.csv", header = TRUE, row.names=1,check.names=FALSE)  
Bac.factors <- read.csv(file = "bacterialfactors.csv", header = TRUE, row.names=1,check.names=FALSE)  
  
#heatmap  
pheatmap(Bac.counts)



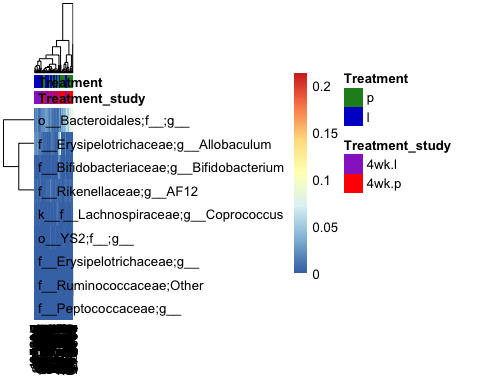
#colour sample groups  
Bac.factorsDS <- select(Bac.factors, Treatment\_study, Treatment)  
pheatmap(Bac.counts, annotation\_col = Bac.factorsDS)



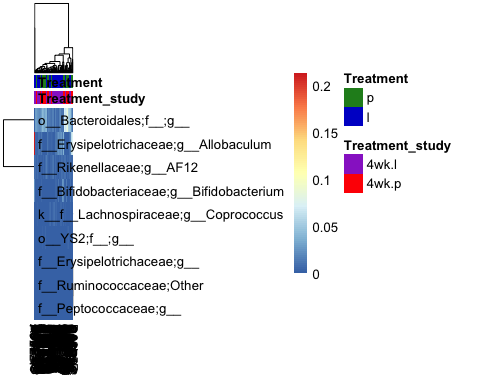
# Reorder Density levels to Sparse, Dense, Other  
Bac.factorsDS$Treatment\_study = factor(Bac.factorsDS$Treatment\_study, levels = c( "4wk.l","4wk.p"))  
DensityCol <- c("darkorchid", "red")  
names(DensityCol) <- levels(Bac.factorsDS$Treatment\_study)  
  
# Reorder TREATMENT to placebo and letrozole  
Bac.factorsDS$Treatment <- factor(Bac.factorsDS$Treatment, levels = c("p", "l"))  
  
SpeciesCol <- c("forestgreen", "blue3")  
names(SpeciesCol) <- levels(Bac.factorsDS$Treatment)  
  
# Add to a list, where names match those in factors dataframe  
AnnColour = list(Treatment\_study = DensityCol,Treatment = SpeciesCol)  
  
# Check the output  
AnnColour

## $Treatment\_study  
## 4wk.l 4wk.p   
## "darkorchid" "red"   
##   
## $Treatment  
## p l   
## "forestgreen" "blue3"

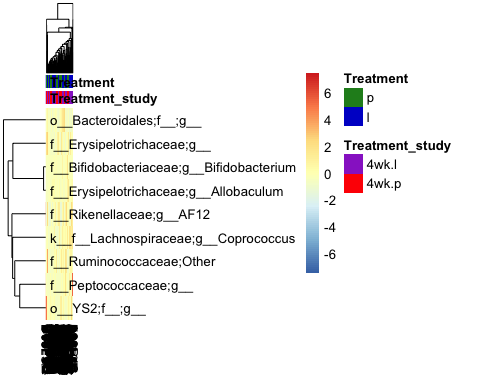
#redraw heatmap  
pheatmap(Bac.counts, annotation\_col = Bac.factorsDS, annotation\_colors = AnnColour)



pheatmap(Bac.counts, clustering\_distance\_rows = "manhattan",  
 clustering\_distance\_cols = "manhattan", clustering\_method = 'average',  
 annotation\_colors = AnnColour, annotation\_col = Bac.factorsDS)



#scaling variables-We can now see how many standard deviations the Log10 abundance of a single OTU is away from the mean for that OTU in a sample compared only with other samples for that OTU  
pheatmap(Bac.counts, scale = "row", clustering\_distance\_rows = "manhattan",  
 clustering\_method = 'average',  
 annotation\_colors = AnnColour, annotation\_col = Bac.factorsDS)



#sort by groups  
SampleOrder = order(Bac.factorsDS$Treatment, Bac.factorsDS$Treatment\_study)  
  
pheatmap(Bac.counts[ , SampleOrder], cluster\_cols = FALSE,  
 clustering\_method = 'average', annotation\_colors = AnnColour, annotation\_col = Bac.factorsDS)

