Redone\_for\_manuscript\_Microbiome.R

Pedro\_Torres

# will do random forest on samples for manuscript.   
  
#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  
# Random Forest was done on L6\_text output with the removal of an unassigned bacterial organism and a TM7 that was extremely low (less than 5 otus in most samples. Then the table was transposed and the columns ‘Treatment\_study; and ‘treatment’ were added from the mapping file to this new transformed ‘merged\_sorted\_L6.txt’ file renamed ‘transposed.csv’ in the Random\_forest folder.  
  
library(randomForest)

## randomForest 4.6-12

## Type rfNews() to see new features/changes/bug fixes.

transposed=read.csv("/Users/Pedro\_Torres/Desktop/transposed.csv", header=T, as.is = T)  
transposed[,3]=factor(transposed[,3])  
  
#transposed[,3]  
#head (transposed)  
#ncol(transposed)  
#transposedoriginal=transposed  
  
#Full oraganism name is too long, will shorted the names to only show order and beyond (e.g, family,genu)  
transposednewname <- sub("^.\*o\_\_", "", colnames(transposed[ ,2:44]), perl=TRUE)  
#transposednewname  
  
names(transposed)[2:44]=transposednewname  
#colnames(transposed[ ,2:44])  
  
#\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
#only week 8  
#transposed  
set.seed(20)  
transposed$Treatment\_study

## [1] "8wk.P" "8wk.P" "8wk.P" "8wk.P" "8wk.P" "8wk.P" "8wk.P" "8wk.P"  
## [9] "8wk.P" "8wk.P" "8wk.P" "8wk.P" "8wk.P" "8wk.P" "8wk.P" "8wk.P"  
## [17] "8wk.P" "8wk.P" "8wk.P" "8wk.P" "8wk.P" "8wk.P" "8wk.P" "8wk.P"  
## [25] "8wk.P" "8wk.P" "8wk.P" "8wk.P" "8wk.P" "8wk.P" "8wk.P" "8wk.P"  
## [33] "8wk.P" "8wk.P" "8wk.P" "8wk.P" "8wk.P" "8wk.P" "8wk.P" "8wk.P"  
## [41] "8wk.L" "8wk.L" "8wk.L" "8wk.L" "8wk.L" "8wk.L" "8wk.L" "8wk.L"  
## [49] "8wk.L" "8wk.L" "8wk.L" "8wk.L" "8wk.L" "8wk.L" "8wk.L" "8wk.L"  
## [57] "8wk.L" "8wk.L" "8wk.L" "8wk.L" "8wk.L" "8wk.L" "8wk.L" "8wk.L"  
## [65] "8wk.L" "8wk.L" "8wk.L" "8wk.L" "8wk.L" "8wk.L" "4wk.p" "4wk.p"  
## [73] "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.l"  
## [81] "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l"  
## [89] "4wk.l" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p"  
## [97] "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p"  
## [105] "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.l" "4wk.l" "4wk.l"  
## [113] "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l"  
## [121] "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l"  
## [129] "4wk.l" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p"  
## [137] "4wk.p" "4wk.p" "4wk.p" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l"  
## [145] "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.p" "4wk.p" "4wk.p"  
## [153] "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.l"  
## [161] "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l"  
## [169] "4wk.l"

transposed[,3]

## [1] p p p p p p p p p p p p p p p p p p p p p p p p p p p p p p p p p p p  
## [36] p p p p p l l l l l l l l l l l l l l l l l l l l l l l l l l l l l l  
## [71] p p p p p p p p p l l l l l l l l l l p p p p p p p p p p p p p p p p  
## [106] p p p p l l l l l l l l l l l l l l l l l l l l p p p p p p p p p p l  
## [141] l l l l l l l l l p p p p p p p p p p l l l l l l l l l l  
## Levels: l p

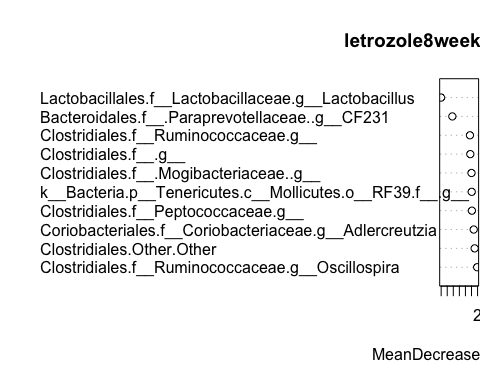
letrozole=subset(transposed,Treatment\_study=="8wk.P"|Treatment\_study=="8wk.L" )  
#training datase  
letrozole$Treatment\_study=factor(letrozole$Treatment\_study)#when i subset i have to turn y variable into a factor or else rf doesnt work  
letrozole$Treatment\_study

## [1] 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P  
## [12] 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P  
## [23] 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P  
## [34] 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P 8wk.P 8wk.L 8wk.L 8wk.L 8wk.L  
## [45] 8wk.L 8wk.L 8wk.L 8wk.L 8wk.L 8wk.L 8wk.L 8wk.L 8wk.L 8wk.L 8wk.L  
## [56] 8wk.L 8wk.L 8wk.L 8wk.L 8wk.L 8wk.L 8wk.L 8wk.L 8wk.L 8wk.L 8wk.L  
## [67] 8wk.L 8wk.L 8wk.L 8wk.L  
## Levels: 8wk.L 8wk.P

xlearn=letrozole[,4:46]  
i=sample(nrow(letrozole),20, rep=F)  
ylearn=letrozole[,3]  
xtest=xlearn[i,]  
xtrain=xlearn[-i,]  
ytest=ylearn[i]  
ytrain=ylearn[-i]  
  
letrozole8weekvs4weekrf=randomForest(xtrain, ytrain, xtest, ytest, importance = T, ntree = 500)  
letrozole8weekvs4weekrf

##   
## Call:  
## randomForest(x = xtrain, y = ytrain, xtest = xtest, ytest = ytest, ntree = 500, importance = T)   
## Type of random forest: classification  
## Number of trees: 500  
## No. of variables tried at each split: 6  
##   
## OOB estimate of error rate: 18%  
## Confusion matrix:  
## l p class.error  
## l 16 5 0.2380952  
## p 4 25 0.1379310  
## Test set error rate: 25%  
## Confusion matrix:  
## l p class.error  
## l 6 3 0.3333333  
## p 2 9 0.1818182

varImpPlot(letrozole8weekvs4weekrf,type = 1,n.var=10)



importance(letrozole8weekvs4weekrf)

## l  
## Bifidobacteriales.f\_\_Bifidobacteriaceae.g\_\_Bifidobacterium 0.09123873  
## Coriobacteriales.f\_\_Coriobacteriaceae.g\_\_ -0.46829897  
## Coriobacteriales.f\_\_Coriobacteriaceae.g\_\_Adlercreutzia 2.40379536  
## Bacteroidales.f\_\_.g\_\_ -0.24680821  
## Bacteroidales.f\_\_Bacteroidaceae.g\_\_Bacteroides -0.25638766  
## Bacteroidales.f\_\_Porphyromonadaceae.g\_\_Parabacteroides -0.73626818  
## Bacteroidales.f\_\_Prevotellaceae.g\_\_Prevotella 1.30825290  
## Bacteroidales.f\_\_Rikenellaceae.g\_\_ 1.02486070  
## Bacteroidales.f\_\_Rikenellaceae.g\_\_AF12 2.29734078  
## Bacteroidales.f\_\_S24.7.g\_\_ 1.95269927  
## Bacteroidales.f\_\_.Paraprevotellaceae..g\_\_CF231 7.68557677  
## YS2.f\_\_.g\_\_ -0.29107968  
## Deferribacterales.f\_\_Deferribacteraceae.g\_\_Mucispirillum 0.52409786  
## Lactobacillales.f\_\_Enterococcaceae.g\_\_Enterococcus -0.65947980  
## Lactobacillales.f\_\_Lactobacillaceae.g\_\_Lactobacillus 12.61538310  
## Clostridiales.Other.Other 2.29928426  
## Clostridiales.f\_\_.g\_\_ 3.85535949  
## Clostridiales.f\_\_Clostridiaceae.g\_\_ -2.45386952  
## Clostridiales.f\_\_Dehalobacteriaceae.g\_\_Dehalobacterium -1.31485559  
## Clostridiales.f\_\_Lachnospiraceae.Other 0.02190857  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_ 1.12424773  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Anaerostipes -0.19625702  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Blautia 0.62472547  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Coprococcus -0.96175776  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Dorea 2.50195320  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Roseburia 0.09410243  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_.Ruminococcus. 1.38327997  
## Clostridiales.f\_\_Peptococcaceae.g\_\_ 2.49987528  
## Clostridiales.f\_\_Peptococcaceae.g\_\_rc4.4 -1.47211368  
## Clostridiales.f\_\_Ruminococcaceae.Other -0.41462523  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_ 2.35540534  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_Oscillospira 2.45665114  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_Ruminococcus -0.49044271  
## Clostridiales.f\_\_.Mogibacteriaceae..g\_\_ 2.38850416  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_ -0.99226270  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_Allobaculum -0.41481412  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_Coprobacillus 0.86911404  
## RF32.f\_\_.g\_\_ -1.58859702  
## Burkholderiales.f\_\_Alcaligenaceae.g\_\_Sutterella 0.12600955  
## Desulfovibrionales.f\_\_Desulfovibrionaceae.g\_\_Bilophila -1.17331212  
## Desulfovibrionales.f\_\_Desulfovibrionaceae.g\_\_Desulfovibrio 1.01959692  
## k\_\_Bacteria.p\_\_Tenericutes.c\_\_Mollicutes.o\_\_RF39.f\_\_.g\_\_ 3.96080245  
## k\_\_Bacteria.p\_\_Verrucomicrobia.c\_\_Verrucomicrobiae.o\_\_Verrucomicrobiales.f\_\_Verrucomicrobiaceae.g\_\_Akkermansia 1.07651845  
## p  
## Bifidobacteriales.f\_\_Bifidobacteriaceae.g\_\_Bifidobacterium -0.734548104  
## Coriobacteriales.f\_\_Coriobacteriaceae.g\_\_ 2.777359547  
## Coriobacteriales.f\_\_Coriobacteriaceae.g\_\_Adlercreutzia 2.726601984  
## Bacteroidales.f\_\_.g\_\_ 1.728506748  
## Bacteroidales.f\_\_Bacteroidaceae.g\_\_Bacteroides -0.439857536  
## Bacteroidales.f\_\_Porphyromonadaceae.g\_\_Parabacteroides 0.981615564  
## Bacteroidales.f\_\_Prevotellaceae.g\_\_Prevotella 0.815539138  
## Bacteroidales.f\_\_Rikenellaceae.g\_\_ -0.509923669  
## Bacteroidales.f\_\_Rikenellaceae.g\_\_AF12 0.735479842  
## Bacteroidales.f\_\_S24.7.g\_\_ 0.127520791  
## Bacteroidales.f\_\_.Paraprevotellaceae..g\_\_CF231 10.071089051  
## YS2.f\_\_.g\_\_ -0.399865714  
## Deferribacterales.f\_\_Deferribacteraceae.g\_\_Mucispirillum 0.025806937  
## Lactobacillales.f\_\_Enterococcaceae.g\_\_Enterococcus -0.242173616  
## Lactobacillales.f\_\_Lactobacillaceae.g\_\_Lactobacillus 12.082399373  
## Clostridiales.Other.Other 1.687234416  
## Clostridiales.f\_\_.g\_\_ 2.586521542  
## Clostridiales.f\_\_Clostridiaceae.g\_\_ -0.344116479  
## Clostridiales.f\_\_Dehalobacteriaceae.g\_\_Dehalobacterium -1.055958057  
## Clostridiales.f\_\_Lachnospiraceae.Other -0.009421689  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_ 1.823364505  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Anaerostipes 0.370693405  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Blautia 2.419674856  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Coprococcus -0.184186019  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Dorea 1.098983404  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Roseburia 1.877379406  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_.Ruminococcus. -0.163287203  
## Clostridiales.f\_\_Peptococcaceae.g\_\_ 3.655527635  
## Clostridiales.f\_\_Peptococcaceae.g\_\_rc4.4 -0.444477318  
## Clostridiales.f\_\_Ruminococcaceae.Other 0.463810103  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_ 4.654981420  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_Oscillospira 0.561382246  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_Ruminococcus 2.071757037  
## Clostridiales.f\_\_.Mogibacteriaceae..g\_\_ 3.665405545  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_ 0.278915639  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_Allobaculum 1.355836230  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_Coprobacillus -0.985892100  
## RF32.f\_\_.g\_\_ 1.135099596  
## Burkholderiales.f\_\_Alcaligenaceae.g\_\_Sutterella 0.307188294  
## Desulfovibrionales.f\_\_Desulfovibrionaceae.g\_\_Bilophila -1.537319879  
## Desulfovibrionales.f\_\_Desulfovibrionaceae.g\_\_Desulfovibrio -0.941964678  
## k\_\_Bacteria.p\_\_Tenericutes.c\_\_Mollicutes.o\_\_RF39.f\_\_.g\_\_ 2.419155197  
## k\_\_Bacteria.p\_\_Verrucomicrobia.c\_\_Verrucomicrobiae.o\_\_Verrucomicrobiales.f\_\_Verrucomicrobiaceae.g\_\_Akkermansia -0.472596807  
## MeanDecreaseAccuracy  
## Bifidobacteriales.f\_\_Bifidobacteriaceae.g\_\_Bifidobacterium -0.40053699  
## Coriobacteriales.f\_\_Coriobacteriaceae.g\_\_ 1.85687972  
## Coriobacteriales.f\_\_Coriobacteriaceae.g\_\_Adlercreutzia 3.27651029  
## Bacteroidales.f\_\_.g\_\_ 0.90657477  
## Bacteroidales.f\_\_Bacteroidaceae.g\_\_Bacteroides -0.49345249  
## Bacteroidales.f\_\_Porphyromonadaceae.g\_\_Parabacteroides 0.24758150  
## Bacteroidales.f\_\_Prevotellaceae.g\_\_Prevotella 1.44616277  
## Bacteroidales.f\_\_Rikenellaceae.g\_\_ 0.68841248  
## Bacteroidales.f\_\_Rikenellaceae.g\_\_AF12 1.75686361  
## Bacteroidales.f\_\_S24.7.g\_\_ 1.47404361  
## Bacteroidales.f\_\_.Paraprevotellaceae..g\_\_CF231 10.30790894  
## YS2.f\_\_.g\_\_ -0.11232542  
## Deferribacterales.f\_\_Deferribacteraceae.g\_\_Mucispirillum 0.13154492  
## Lactobacillales.f\_\_Enterococcaceae.g\_\_Enterococcus -0.47103131  
## Lactobacillales.f\_\_Lactobacillaceae.g\_\_Lactobacillus 14.00869027  
## Clostridiales.Other.Other 2.96081136  
## Clostridiales.f\_\_.g\_\_ 4.39516460  
## Clostridiales.f\_\_Clostridiaceae.g\_\_ -0.91818017  
## Clostridiales.f\_\_Dehalobacteriaceae.g\_\_Dehalobacterium -1.47629111  
## Clostridiales.f\_\_Lachnospiraceae.Other -0.01802041  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_ 1.86191583  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Anaerostipes 0.24857493  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Blautia 1.94790082  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Coprococcus -0.96494874  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Dorea 1.97614199  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Roseburia 1.61947123  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_.Ruminococcus. 0.97535234  
## Clostridiales.f\_\_Peptococcaceae.g\_\_ 3.87053411  
## Clostridiales.f\_\_Peptococcaceae.g\_\_rc4.4 -1.27839456  
## Clostridiales.f\_\_Ruminococcaceae.Other 0.20593680  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_ 4.51310051  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_Oscillospira 2.13238489  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_Ruminococcus 0.83363905  
## Clostridiales.f\_\_.Mogibacteriaceae..g\_\_ 3.99805522  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_ -0.93629482  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_Allobaculum 0.86477505  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_Coprobacillus -0.11787457  
## RF32.f\_\_.g\_\_ -0.10971878  
## Burkholderiales.f\_\_Alcaligenaceae.g\_\_Sutterella 0.12120726  
## Desulfovibrionales.f\_\_Desulfovibrionaceae.g\_\_Bilophila -1.52306800  
## Desulfovibrionales.f\_\_Desulfovibrionaceae.g\_\_Desulfovibrio -0.07173338  
## k\_\_Bacteria.p\_\_Tenericutes.c\_\_Mollicutes.o\_\_RF39.f\_\_.g\_\_ 3.92631500  
## k\_\_Bacteria.p\_\_Verrucomicrobia.c\_\_Verrucomicrobiae.o\_\_Verrucomicrobiales.f\_\_Verrucomicrobiaceae.g\_\_Akkermansia 0.13439620  
## MeanDecreaseGini  
## Bifidobacteriales.f\_\_Bifidobacteriaceae.g\_\_Bifidobacterium 0.3263300  
## Coriobacteriales.f\_\_Coriobacteriaceae.g\_\_ 0.5715491  
## Coriobacteriales.f\_\_Coriobacteriaceae.g\_\_Adlercreutzia 0.6260396  
## Bacteroidales.f\_\_.g\_\_ 0.3129137  
## Bacteroidales.f\_\_Bacteroidaceae.g\_\_Bacteroides 0.3090902  
## Bacteroidales.f\_\_Porphyromonadaceae.g\_\_Parabacteroides 0.3096988  
## Bacteroidales.f\_\_Prevotellaceae.g\_\_Prevotella 0.6515764  
## Bacteroidales.f\_\_Rikenellaceae.g\_\_ 0.4085161  
## Bacteroidales.f\_\_Rikenellaceae.g\_\_AF12 0.5445688  
## Bacteroidales.f\_\_S24.7.g\_\_ 0.5855106  
## Bacteroidales.f\_\_.Paraprevotellaceae..g\_\_CF231 2.2251493  
## YS2.f\_\_.g\_\_ 0.2054602  
## Deferribacterales.f\_\_Deferribacteraceae.g\_\_Mucispirillum 0.3364957  
## Lactobacillales.f\_\_Enterococcaceae.g\_\_Enterococcus 0.3053135  
## Lactobacillales.f\_\_Lactobacillaceae.g\_\_Lactobacillus 4.1285695  
## Clostridiales.Other.Other 0.5193913  
## Clostridiales.f\_\_.g\_\_ 0.8333844  
## Clostridiales.f\_\_Clostridiaceae.g\_\_ 0.2442749  
## Clostridiales.f\_\_Dehalobacteriaceae.g\_\_Dehalobacterium 0.2896436  
## Clostridiales.f\_\_Lachnospiraceae.Other 0.2641785  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_ 0.7758547  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Anaerostipes 0.4693866  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Blautia 0.4576047  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Coprococcus 0.3986170  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Dorea 0.3898145  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Roseburia 0.3993754  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_.Ruminococcus. 0.2763106  
## Clostridiales.f\_\_Peptococcaceae.g\_\_ 0.8307807  
## Clostridiales.f\_\_Peptococcaceae.g\_\_rc4.4 0.2457834  
## Clostridiales.f\_\_Ruminococcaceae.Other 0.2433866  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_ 1.0385315  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_Oscillospira 0.6469097  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_Ruminococcus 0.5656097  
## Clostridiales.f\_\_.Mogibacteriaceae..g\_\_ 0.7372475  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_ 0.1286740  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_Allobaculum 0.3697550  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_Coprobacillus 0.2116429  
## RF32.f\_\_.g\_\_ 0.2069657  
## Burkholderiales.f\_\_Alcaligenaceae.g\_\_Sutterella 0.2563752  
## Desulfovibrionales.f\_\_Desulfovibrionaceae.g\_\_Bilophila 0.1056834  
## Desulfovibrionales.f\_\_Desulfovibrionaceae.g\_\_Desulfovibrio 0.2252417  
## k\_\_Bacteria.p\_\_Tenericutes.c\_\_Mollicutes.o\_\_RF39.f\_\_.g\_\_ 0.5842313  
## k\_\_Bacteria.p\_\_Verrucomicrobia.c\_\_Verrucomicrobiae.o\_\_Verrucomicrobiales.f\_\_Verrucomicrobiaceae.g\_\_Akkermansia 0.3024038

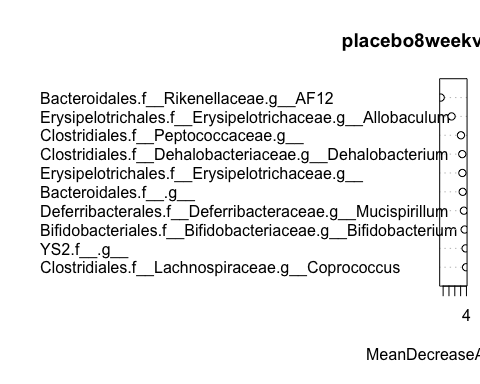
#transposed[,3]  
set.seed(20)  
placebo=subset(transposed, Treatment\_study=="4wk.l"|Treatment\_study=="4wk.p" )  
placebo$Treatment\_study

## [1] "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p"  
## [9] "4wk.p" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l"  
## [17] "4wk.l" "4wk.l" "4wk.l" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p"  
## [25] "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p"  
## [33] "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.l"  
## [41] "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l"  
## [49] "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l"  
## [57] "4wk.l" "4wk.l" "4wk.l" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p"  
## [65] "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.l" "4wk.l" "4wk.l"  
## [73] "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.p"  
## [81] "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p" "4wk.p"  
## [89] "4wk.p" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l" "4wk.l"  
## [97] "4wk.l" "4wk.l" "4wk.l"

# must make my categories into factors after a subset  
placebo$Treatment\_study=factor(placebo$Treatment\_study)  
  
#random forest  
xlearn=placebo[ ,4:46]  
i=sample(nrow(placebo),20, rep=F)  
ylearn=placebo[,3]  
xtest=xlearn[i,]  
xtrain=xlearn[-i,]  
ytest=ylearn[i]  
ytrain=ylearn[-i]  
  
placebo8weekvs4weekrf=randomForest(xtrain, ytrain, xtest, ytest, importance = T, ntree = 500)  
placebo8weekvs4weekrf

##   
## Call:  
## randomForest(x = xtrain, y = ytrain, xtest = xtest, ytest = ytest, ntree = 500, importance = T)   
## Type of random forest: classification  
## Number of trees: 500  
## No. of variables tried at each split: 6  
##   
## OOB estimate of error rate: 21.52%  
## Confusion matrix:  
## l p class.error  
## l 36 8 0.1818182  
## p 9 26 0.2571429  
## Test set error rate: 35%  
## Confusion matrix:  
## l p class.error  
## l 5 1 0.1666667  
## p 6 8 0.4285714

varImpPlot(placebo8weekvs4weekrf,type = 1,n.var=10)



importance(placebo8weekvs4weekrf)

## l  
## Bifidobacteriales.f\_\_Bifidobacteriaceae.g\_\_Bifidobacterium 3.045186102  
## Coriobacteriales.f\_\_Coriobacteriaceae.g\_\_ -0.892517715  
## Coriobacteriales.f\_\_Coriobacteriaceae.g\_\_Adlercreutzia -0.615804713  
## Bacteroidales.f\_\_.g\_\_ 3.154658145  
## Bacteroidales.f\_\_Bacteroidaceae.g\_\_Bacteroides -0.278121514  
## Bacteroidales.f\_\_Porphyromonadaceae.g\_\_Parabacteroides 0.971797665  
## Bacteroidales.f\_\_Prevotellaceae.g\_\_Prevotella -3.111163305  
## Bacteroidales.f\_\_Rikenellaceae.g\_\_ 0.496710950  
## Bacteroidales.f\_\_Rikenellaceae.g\_\_AF12 6.047487230  
## Bacteroidales.f\_\_S24.7.g\_\_ 0.005903735  
## Bacteroidales.f\_\_.Paraprevotellaceae..g\_\_CF231 1.324603642  
## YS2.f\_\_.g\_\_ 2.047980120  
## Deferribacterales.f\_\_Deferribacteraceae.g\_\_Mucispirillum 3.128850953  
## Lactobacillales.f\_\_Enterococcaceae.g\_\_Enterococcus -2.796456979  
## Lactobacillales.f\_\_Lactobacillaceae.g\_\_Lactobacillus 0.015129854  
## Clostridiales.Other.Other 1.960461316  
## Clostridiales.f\_\_.g\_\_ 0.490056782  
## Clostridiales.f\_\_Clostridiaceae.g\_\_ -0.127267589  
## Clostridiales.f\_\_Dehalobacteriaceae.g\_\_Dehalobacterium 3.496343536  
## Clostridiales.f\_\_Lachnospiraceae.Other -0.530516052  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_ 1.275783700  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Anaerostipes -1.728891524  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Blautia -2.213247743  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Coprococcus 4.024175797  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Dorea 0.696214353  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Roseburia 0.568323897  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_.Ruminococcus. 1.616613959  
## Clostridiales.f\_\_Peptococcaceae.g\_\_ 4.117454211  
## Clostridiales.f\_\_Peptococcaceae.g\_\_rc4.4 -0.455190836  
## Clostridiales.f\_\_Ruminococcaceae.Other 0.378894125  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_ 1.392740645  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_Oscillospira 1.912804675  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_Ruminococcus 2.095492235  
## Clostridiales.f\_\_.Mogibacteriaceae..g\_\_ -0.221876681  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_ 2.746313413  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_Allobaculum 4.348577996  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_Coprobacillus 1.299484974  
## RF32.f\_\_.g\_\_ -0.524694056  
## Burkholderiales.f\_\_Alcaligenaceae.g\_\_Sutterella 1.357557296  
## Desulfovibrionales.f\_\_Desulfovibrionaceae.g\_\_Bilophila 0.124343962  
## Desulfovibrionales.f\_\_Desulfovibrionaceae.g\_\_Desulfovibrio 2.974701967  
## k\_\_Bacteria.p\_\_Tenericutes.c\_\_Mollicutes.o\_\_RF39.f\_\_.g\_\_ 1.656063940  
## k\_\_Bacteria.p\_\_Verrucomicrobia.c\_\_Verrucomicrobiae.o\_\_Verrucomicrobiales.f\_\_Verrucomicrobiaceae.g\_\_Akkermansia 3.455383329  
## p  
## Bifidobacteriales.f\_\_Bifidobacteriaceae.g\_\_Bifidobacterium 3.993928345  
## Coriobacteriales.f\_\_Coriobacteriaceae.g\_\_ 1.877157794  
## Coriobacteriales.f\_\_Coriobacteriaceae.g\_\_Adlercreutzia -0.584184016  
## Bacteroidales.f\_\_.g\_\_ 3.886677890  
## Bacteroidales.f\_\_Bacteroidaceae.g\_\_Bacteroides -0.219668505  
## Bacteroidales.f\_\_Porphyromonadaceae.g\_\_Parabacteroides -0.003077478  
## Bacteroidales.f\_\_Prevotellaceae.g\_\_Prevotella 1.108062067  
## Bacteroidales.f\_\_Rikenellaceae.g\_\_ 0.446362301  
## Bacteroidales.f\_\_Rikenellaceae.g\_\_AF12 7.719703550  
## Bacteroidales.f\_\_S24.7.g\_\_ 0.490954591  
## Bacteroidales.f\_\_.Paraprevotellaceae..g\_\_CF231 0.269303980  
## YS2.f\_\_.g\_\_ 4.254347982  
## Deferribacterales.f\_\_Deferribacteraceae.g\_\_Mucispirillum 3.255852834  
## Lactobacillales.f\_\_Enterococcaceae.g\_\_Enterococcus -1.154844299  
## Lactobacillales.f\_\_Lactobacillaceae.g\_\_Lactobacillus 0.839857778  
## Clostridiales.Other.Other -0.098494195  
## Clostridiales.f\_\_.g\_\_ -0.672892683  
## Clostridiales.f\_\_Clostridiaceae.g\_\_ 0.681202741  
## Clostridiales.f\_\_Dehalobacteriaceae.g\_\_Dehalobacterium 4.393544339  
## Clostridiales.f\_\_Lachnospiraceae.Other -0.061677814  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_ 1.439695981  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Anaerostipes 0.049608719  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Blautia -1.310149429  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Coprococcus 2.030009526  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Dorea 1.394275781  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Roseburia 1.043643042  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_.Ruminococcus. -0.022575428  
## Clostridiales.f\_\_Peptococcaceae.g\_\_ 3.333308955  
## Clostridiales.f\_\_Peptococcaceae.g\_\_rc4.4 -1.106801445  
## Clostridiales.f\_\_Ruminococcaceae.Other 0.055953747  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_ 0.956644159  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_Oscillospira 0.386538972  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_Ruminococcus 1.585068045  
## Clostridiales.f\_\_.Mogibacteriaceae..g\_\_ -0.291009788  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_ 4.982373647  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_Allobaculum 6.340041037  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_Coprobacillus -0.651843171  
## RF32.f\_\_.g\_\_ -1.587995922  
## Burkholderiales.f\_\_Alcaligenaceae.g\_\_Sutterella 0.755020125  
## Desulfovibrionales.f\_\_Desulfovibrionaceae.g\_\_Bilophila -0.689883551  
## Desulfovibrionales.f\_\_Desulfovibrionaceae.g\_\_Desulfovibrio 1.826145300  
## k\_\_Bacteria.p\_\_Tenericutes.c\_\_Mollicutes.o\_\_RF39.f\_\_.g\_\_ 1.232347660  
## k\_\_Bacteria.p\_\_Verrucomicrobia.c\_\_Verrucomicrobiae.o\_\_Verrucomicrobiales.f\_\_Verrucomicrobiaceae.g\_\_Akkermansia 1.109364613  
## MeanDecreaseAccuracy  
## Bifidobacteriales.f\_\_Bifidobacteriaceae.g\_\_Bifidobacterium 4.3061063  
## Coriobacteriales.f\_\_Coriobacteriaceae.g\_\_ 0.4989531  
## Coriobacteriales.f\_\_Coriobacteriaceae.g\_\_Adlercreutzia -0.6099921  
## Bacteroidales.f\_\_.g\_\_ 4.6426215  
## Bacteroidales.f\_\_Bacteroidaceae.g\_\_Bacteroides -0.3654303  
## Bacteroidales.f\_\_Porphyromonadaceae.g\_\_Parabacteroides 0.5582761  
## Bacteroidales.f\_\_Prevotellaceae.g\_\_Prevotella -1.2528312  
## Bacteroidales.f\_\_Rikenellaceae.g\_\_ 0.5450073  
## Bacteroidales.f\_\_Rikenellaceae.g\_\_AF12 8.4181308  
## Bacteroidales.f\_\_S24.7.g\_\_ 0.1945786  
## Bacteroidales.f\_\_.Paraprevotellaceae..g\_\_CF231 1.2214419  
## YS2.f\_\_.g\_\_ 4.1440792  
## Deferribacterales.f\_\_Deferribacteraceae.g\_\_Mucispirillum 4.4385356  
## Lactobacillales.f\_\_Enterococcaceae.g\_\_Enterococcus -2.4540168  
## Lactobacillales.f\_\_Lactobacillaceae.g\_\_Lactobacillus 0.5213166  
## Clostridiales.Other.Other 1.3537607  
## Clostridiales.f\_\_.g\_\_ -0.2946579  
## Clostridiales.f\_\_Clostridiaceae.g\_\_ 0.3397081  
## Clostridiales.f\_\_Dehalobacteriaceae.g\_\_Dehalobacterium 4.7194258  
## Clostridiales.f\_\_Lachnospiraceae.Other -0.1582172  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_ 2.1119602  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Anaerostipes -1.2423745  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Blautia -2.2281479  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Coprococcus 4.0693315  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Dorea 1.3296397  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Roseburia 1.2670552  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_.Ruminococcus. 1.0593528  
## Clostridiales.f\_\_Peptococcaceae.g\_\_ 4.9423552  
## Clostridiales.f\_\_Peptococcaceae.g\_\_rc4.4 -0.9670727  
## Clostridiales.f\_\_Ruminococcaceae.Other 0.2673158  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_ 1.4428518  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_Oscillospira 1.6789761  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_Ruminococcus 2.6513578  
## Clostridiales.f\_\_.Mogibacteriaceae..g\_\_ -0.4551168  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_ 4.6996707  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_Allobaculum 6.5513504  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_Coprobacillus 0.4667776  
## RF32.f\_\_.g\_\_ -1.6041087  
## Burkholderiales.f\_\_Alcaligenaceae.g\_\_Sutterella 1.3617480  
## Desulfovibrionales.f\_\_Desulfovibrionaceae.g\_\_Bilophila -0.1498291  
## Desulfovibrionales.f\_\_Desulfovibrionaceae.g\_\_Desulfovibrio 3.2158834  
## k\_\_Bacteria.p\_\_Tenericutes.c\_\_Mollicutes.o\_\_RF39.f\_\_.g\_\_ 1.9495899  
## k\_\_Bacteria.p\_\_Verrucomicrobia.c\_\_Verrucomicrobiae.o\_\_Verrucomicrobiales.f\_\_Verrucomicrobiaceae.g\_\_Akkermansia 3.2222356  
## MeanDecreaseGini  
## Bifidobacteriales.f\_\_Bifidobacteriaceae.g\_\_Bifidobacterium 1.8362972  
## Coriobacteriales.f\_\_Coriobacteriaceae.g\_\_ 0.9096274  
## Coriobacteriales.f\_\_Coriobacteriaceae.g\_\_Adlercreutzia 0.6045852  
## Bacteroidales.f\_\_.g\_\_ 1.6727180  
## Bacteroidales.f\_\_Bacteroidaceae.g\_\_Bacteroides 0.6751342  
## Bacteroidales.f\_\_Porphyromonadaceae.g\_\_Parabacteroides 0.5253559  
## Bacteroidales.f\_\_Prevotellaceae.g\_\_Prevotella 0.4142476  
## Bacteroidales.f\_\_Rikenellaceae.g\_\_ 0.6015880  
## Bacteroidales.f\_\_Rikenellaceae.g\_\_AF12 3.0040886  
## Bacteroidales.f\_\_S24.7.g\_\_ 0.5546118  
## Bacteroidales.f\_\_.Paraprevotellaceae..g\_\_CF231 0.7320278  
## YS2.f\_\_.g\_\_ 1.6036455  
## Deferribacterales.f\_\_Deferribacteraceae.g\_\_Mucispirillum 1.4291107  
## Lactobacillales.f\_\_Enterococcaceae.g\_\_Enterococcus 0.4592077  
## Lactobacillales.f\_\_Lactobacillaceae.g\_\_Lactobacillus 0.6978106  
## Clostridiales.Other.Other 1.0068950  
## Clostridiales.f\_\_.g\_\_ 0.4195818  
## Clostridiales.f\_\_Clostridiaceae.g\_\_ 0.5027492  
## Clostridiales.f\_\_Dehalobacteriaceae.g\_\_Dehalobacterium 1.1819718  
## Clostridiales.f\_\_Lachnospiraceae.Other 0.6020869  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_ 0.8737167  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Anaerostipes 0.4986338  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Blautia 0.3471450  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Coprococcus 1.7104587  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Dorea 0.5656468  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_Roseburia 0.6067979  
## Clostridiales.f\_\_Lachnospiraceae.g\_\_.Ruminococcus. 0.3510623  
## Clostridiales.f\_\_Peptococcaceae.g\_\_ 1.0478108  
## Clostridiales.f\_\_Peptococcaceae.g\_\_rc4.4 0.2742674  
## Clostridiales.f\_\_Ruminococcaceae.Other 0.6887211  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_ 0.6809882  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_Oscillospira 0.6192851  
## Clostridiales.f\_\_Ruminococcaceae.g\_\_Ruminococcus 1.3241538  
## Clostridiales.f\_\_.Mogibacteriaceae..g\_\_ 0.5456512  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_ 1.9985328  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_Allobaculum 2.1010432  
## Erysipelotrichales.f\_\_Erysipelotrichaceae.g\_\_Coprobacillus 0.3642061  
## RF32.f\_\_.g\_\_ 0.4406714  
## Burkholderiales.f\_\_Alcaligenaceae.g\_\_Sutterella 0.9132129  
## Desulfovibrionales.f\_\_Desulfovibrionaceae.g\_\_Bilophila 0.4299663  
## Desulfovibrionales.f\_\_Desulfovibrionaceae.g\_\_Desulfovibrio 0.9022531  
## k\_\_Bacteria.p\_\_Tenericutes.c\_\_Mollicutes.o\_\_RF39.f\_\_.g\_\_ 0.8254440  
## k\_\_Bacteria.p\_\_Verrucomicrobia.c\_\_Verrucomicrobiae.o\_\_Verrucomicrobiales.f\_\_Verrucomicrobiaceae.g\_\_Akkermansia 0.9762815