Day 10: Statistical Analysis (Global Gut)

Back to

All of the code in this page is meant to be run in `R` unless otherwise specified.

Loading a genus table and the metadata into R

Before loading data into R, this QIIME command must be run on the command line to collapse OTU counts
into genus (L6) and phylum (L2) count tables:

```
# (run on command line)
summarize_taxa.py -i otu_table.biom -L 6

# convert to JSON BIOM format to load into R using R biom package:
biom convert -i otu_table_L6.biom -o otu_table_L6_json.biom --to-json
```

Inside `R`, Install biom package and vegan package if not installed.

```
install.packages(c('biom','vegan'),repo='http://cran.wustl.edu')
```

Load biom package, vegan package; load data

```
library('biom')
library('vegan')

# load biom file
genus.biom <- read_biom('otu_table_L6_json.biom')

# Extract data matrix (genus counts) from biom table
genus <- as.matrix(biom_data(genus.biom))

# transpose so that rows are samples and columns are genera
genus <- t(genus)

# load mapping file
map <- read.table('map.txt', sep='\t', comment='', head=T, row.names=1)
```

It is extremely important to ensure that your genus table and metadata table sample IDs are lined up
correctly.

```
# find the overlapping samples
common.ids <- intersect(rownames(map), rownames(genus))

# get just the overlapping samples
genus <- genus[common.ids,]
map <- map[common.ids,]
```

See dimensions of genus table. Then drop genera present in < 10% of samples.

```r
dim(genus)
```

```
## [1]  66 205
```

```r
genus <- genus[,colMeans(genus > 0) >= .1]
dim(genus)
```

```
## [1] 66 97
```

Show only the first ten genera in genus table

```r
colnames(genus)[1:10]
```

```
##  [1] "k__Archaea;p__Euryarchaeota;c__Methanobacteria;o__Methanobacteriales;f__Methanobacteriaceae;g__
##  [2] "k__Archaea;p__Euryarchaeota;c__Methanobacteria;o__Methanobacteriales;f__Methanobacteriaceae;g__
##  [3] "k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Actinomycetaceae;g__Acti
##  [4] "k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Bifidobacteriales;f__Bifidobacteriaceae;g__
##  [5] "k__Bacteria;p__Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__"
##  [6] "k__Bacteria;p__Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__Ad
##  [7] "k__Bacteria;p__Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__Co
##  [8] "k__Bacteria;p__Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__Sl
##  [9] "k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__;g__"
## [10] "k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides"
```

Strip the kingdom level from the names so that these are easier to display.

```r
colnames(genus) <- sapply(strsplit(colnames(genus),'p__'),'[',2)
```

Show the first 10 rows and first 2 columns of the genus table

```r
genus[1:10,1:2]
```

```
##                  Euryarchaeota;c__Methanobacteria;o__Methanobacteriales;f__Methanobacteriaceae;g__Meth
## h122M.1.418534
## h85M.1.418596
## h95M.1.418831
## h9M.1.418588
## k278A.2.418424
## h146M.1.418838
## h147M.1.418531
## h101M.1.418586
## h165M.1.418394
## h186M.1.418788
##                  Euryarchaeota;c__Methanobacteria;o__Methanobacteriales;f__Methanobacteriaceae;g__Meth
## h122M.1.418534                                                                                     0.0
## h85M.1.418596                                                                                      0.0
## h95M.1.418831                                                                                      0.0
## h9M.1.418588                                                                                       0.0
## k278A.2.418424                                                                                     0.0
## h146M.1.418838                                                                                     0.0
```

```
## h147M.1.418531                                                              0.
## h101M.1.418586                                                              0.
## h165M.1.418394                                                              0.
## h186M.1.418788                                                              0.
```

See available columns in the metadata

```
colnames(map)
```

```
##  [1] "BarcodeSequence"           "LinkerPrimerSequence"
##  [3] "AGE"                       "AGE_GROUP"
##  [5] "BODY_SITE"                 "COUNTRY"
##  [7] "ELEVATION"                 "EXPERIMENT_DESIGN_DESCRIPTION"
##  [9] "FAMILY_RELATIONSHIP"       "HOST_COMMON_NAME"
## [11] "HOST_SUBJECT_ID"           "LATITUDE"
## [13] "LONGITUDE"                 "PCR_PRIMERS"
## [15] "REGION"                    "RUN"
## [17] "RUN_CENTER"                "RUN_DATE"
## [19] "RUN_LANE"                  "SEX"
## [21] "STUDY_ABSTRACT"            "Description"
```

Show how many samples are from each Country

```
table(map$COUNTRY)
```

```
##
##                  GAZ:Malawi GAZ:United States of America
##                          22                            22
##              GAZ:Venezuela
##                          22
```

Basic association testing

Let's run some tests on Prevotella. First extract the Prevotella column and save it to a variable `prevotella` for convenience.

```
# find out what column Prevotella is in
grep('g__Prevotella',colnames(genus))
```
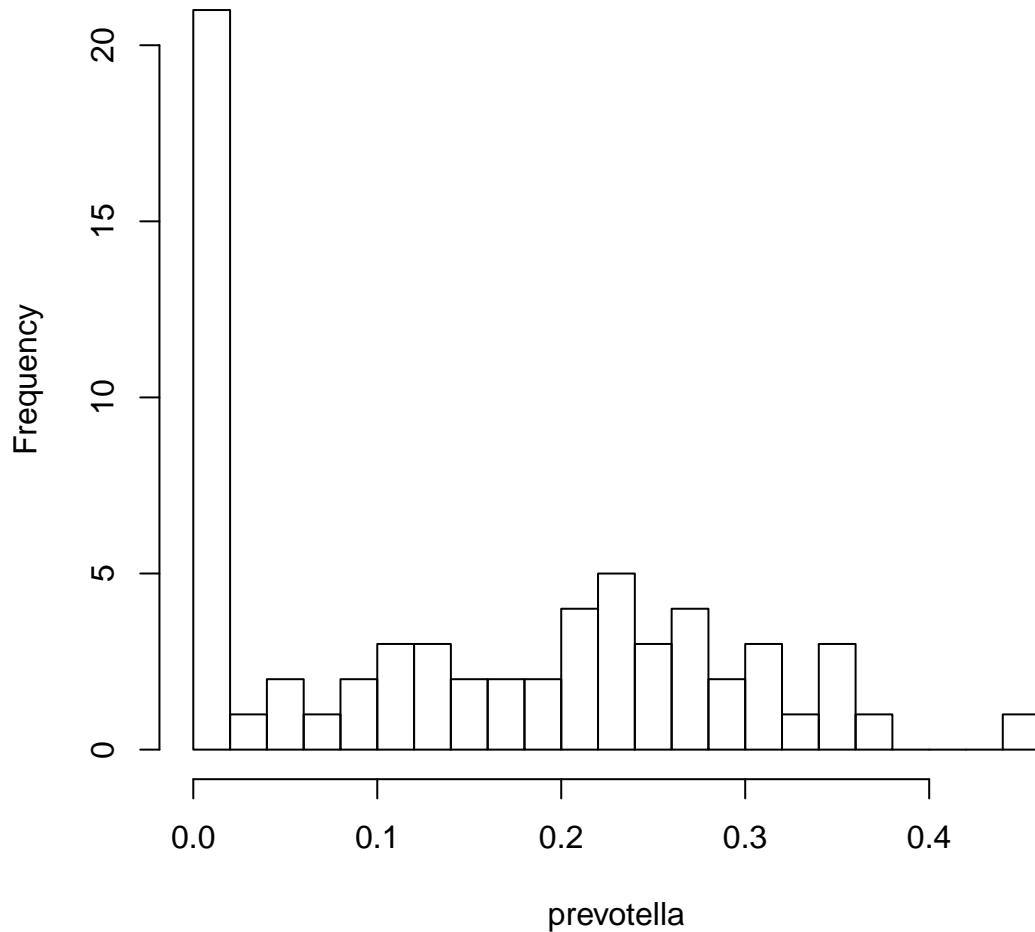
```
## [1] 13
```

```
# save that column in a variable
prevotella <- genus[,grep('g__Prevotella',colnames(genus))]
```

Visualize the distribution of Prevotella

```
# find out what column Prevotella is in
hist(prevotella, br=30)
```

Histogram of prevotella

Run a test of Pearson's correlation of Prevotella and age. Note that the result is not quite significant (p=0.0531).

```
cor.test(prevotella, map$AGE)
```

```
##
##  Pearson's product-moment correlation
##
## data:  prevotella and map$AGE
## t = -1.9704, df = 64, p-value = 0.05312
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.454858210  0.003055594
## sample estimates:
##       cor
## -0.239154
```

Now run a linear regression of prevotella against age. Notice that statistically this is equivalent to running the Pearson's correlation. The p-value in row 2 column 4 of the "Coefficients" table is the same as the p-value from the correlation test.

```
# fit a linear model. The "~" means "as a function of"
fit <- lm(prevotella ~ map$AGE)

# print a summary of the results
summary(fit)
```

```
##
## Call:
## lm(formula = prevotella ~ map$AGE)
##
## Residuals:
##      Min      1Q   Median      3Q      Max
## -0.18946 -0.12186 -0.01805  0.10726  0.30337
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.2009925  0.0319288   6.295 3.15e-08 ***
## map$AGE     -0.0019218  0.0009753  -1.970   0.0531 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.124 on 64 degrees of freedom
## Multiple R-squared:  0.05719,    Adjusted R-squared:  0.04246
## F-statistic: 3.883 on 1 and 64 DF,  p-value: 0.05312
```

```
# A nice way to get the exact p-value for the age regression coefficient using the anova function
pval <- anova(fit)['map$AGE','Pr(>F)']
pval
```

```
## [1] 0.0531206
```

## Testing for normally distributed data

We can test whether the residuals are normally distributed Kolmogorov-Smirnov test. If $p < 0.05$, we can reject the null hypothesis that the data came from a normal distribution, meaning that the linear test is not appropriate.

```
# Make a quantile-quantile plot of the (studentized) residuals vs. a normal distribution
qqnorm(rstudent(fit)); abline(0,1)
```

Normal Q–Q Plot
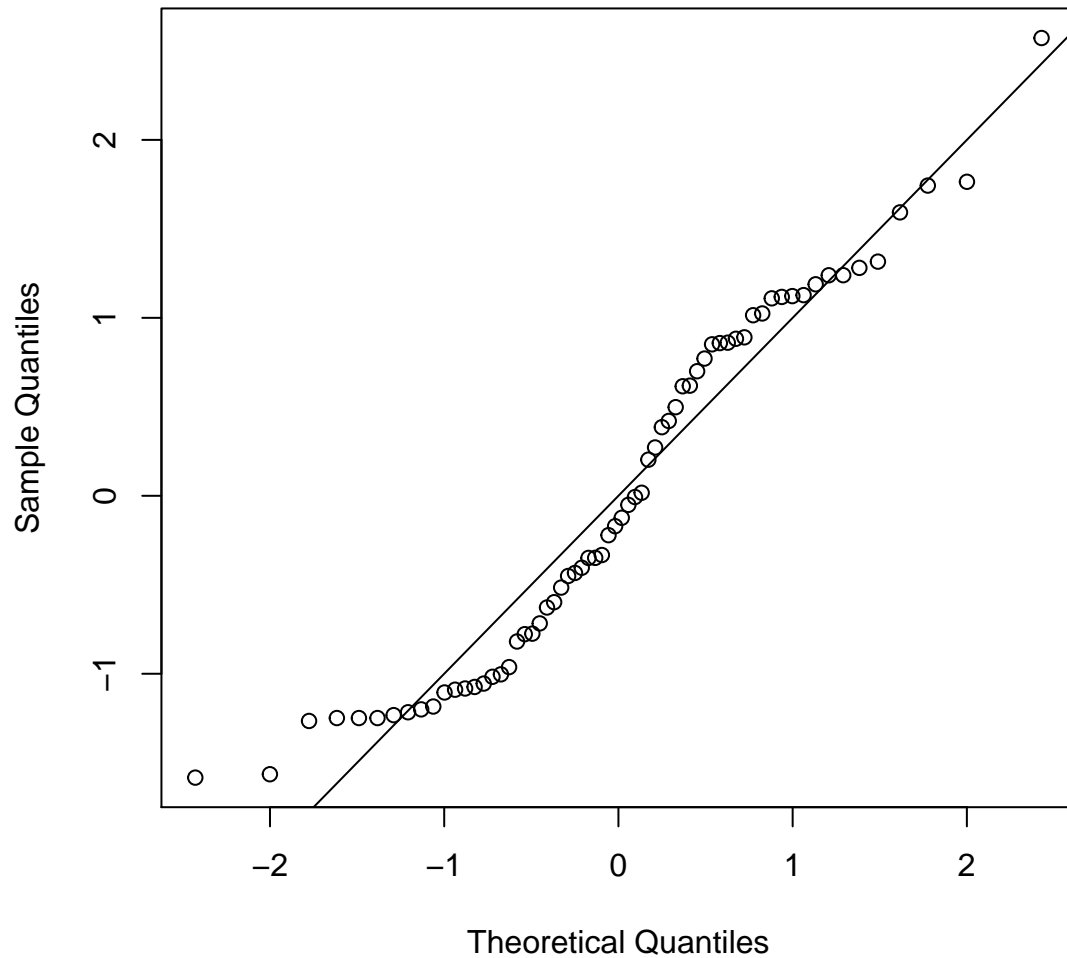

```r
# Kolmogorov-Smirnov test
ks.test(rstudent(fit), pnorm, mean=mean(rstudent(fit)), sd=sd(rstudent(fit)))
```

```
## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit))): ties should not be present for the Kolmogorov-Smirnov
## test
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstudent(fit)
## D = 0.10218, p-value = 0.496
## alternative hypothesis: two-sided
```

Controling for confounders

Perhaps country of origin is a confounder that is obscuring the association of Prevotella and Age. Using lm() we can add confounders to the regression. Now after removing the effects of country, there is a strong association of Prevotella and age.

```
# fit a linear model. The "~" means "as a function of"
fit <- lm(prevotella ~ map$AGE + map$COUNTRY)

# print a summary of the results
summary(fit)
```

```
##
## Call:
## lm(formula = prevotella ~ map$AGE + map$COUNTRY)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.150373 -0.040611 -0.008627  0.033487  0.199252
##
## Coefficients:
##                                     Estimate Std. Error t value
## (Intercept)                        0.2982044  0.0238895  12.483
## map$AGE                           -0.0016456  0.0006193  -2.657
## map$COUNTRYGAZ:United States of America -0.2311571  0.0237310  -9.741
## map$COUNTRYGAZ:Venezuela          -0.0843013  0.0237296  -3.553
##                                     Pr(>|t|)
## (Intercept)                         < 2e-16 ***
## map$AGE                             0.010010 *
## map$COUNTRYGAZ:United States of America 4.08e-14 ***
## map$COUNTRYGAZ:Venezuela            0.000736 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.07861 on 62 degrees of freedom
## Multiple R-squared:  0.6329, Adjusted R-squared:  0.6152
## F-statistic: 35.63 on 3 and 62 DF,  p-value: 1.641e-13
```

Testing multiple hypotheses

We have so far only tested one genus. Let's test them all using a loop.

```
# pvals is a vector initialized with zeroes
# with enough slots for the different genera
pvals <- numeric(ncol(genus))

# "name" the pvalues after the genera
names(pvals) <- colnames(genus)

# Loop through the columns of the genus table, testing each one
for(i in 1:ncol(genus)) {
    fit <- lm(genus[,i] ~ map$AGE + map$COUNTRY)
    pvals[i] <-  anova(fit)['map$AGE','Pr(>F)']
```

```
}

# note, you could put this all on one line with:
# for(i in 1:ncol(genus)) pvals[i] <-  anova(lm(genus[,i] ~ map$AGE + map$COUNTRY))['map$AGE','Pr(>F)']

# print the 10 smallest p-values:
sort(pvals)[1:10]
```

```
##                                                                WPS-2;c__;o__;f__;g__
##                                                                           0.002762885
##                  Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella
##                                                                           0.002841431
##                                    Verrucomicrobia;c__Verruco-5;o__WCHB1-41;f__RFP12;g__
##                                                                           0.004747721
##                        Firmicutes;c__Clostridia;o__Clostridiales;f__Christensenellaceae;g__
##                                                                           0.010042840
##                    Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Lactococcus
##                                                                           0.011344588
##                              Proteobacteria;c__Alphaproteobacteria;o__RF32;f__;g__
##                                                                           0.019484449
## Euryarchaeota;c__Methanobacteria;o__Methanobacteriales;f__Methanobacteriaceae;g__Methanobrevibacter
##                                                                           0.028070760
##                                  Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__
##                                                                           0.030610821
##              Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium
##                                                                           0.035887736
##                                    Proteobacteria;c__Alphaproteobacteria;o__;f__;g__
##                                                                           0.045431657
```

Note, you could put this all on one line with this:

```
# "apply" with genus, 2 means do something to every column of genus
# ("apply" with genus, 1 would mean do something every row)
# the last part defines a new function to do to each column, which
# will be passed in the temporary variable named "xx"
pvals <- apply(genus, 2, function(xx) anova(lm(xx ~ map$AGE + map$COUNTRY))['map$AGE','Pr(>F)'])

# print the 10 smallest p-values:
sort(pvals)[1:10]
```

```
##                                                                WPS-2;c__;o__;f__;g__
##                                                                           0.002762885
##                  Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella
##                                                                           0.002841431
##                                    Verrucomicrobia;c__Verruco-5;o__WCHB1-41;f__RFP12;g__
##                                                                           0.004747721
##                        Firmicutes;c__Clostridia;o__Clostridiales;f__Christensenellaceae;g__
##                                                                           0.010042840
##                    Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Lactococcus
##                                                                           0.011344588
##                              Proteobacteria;c__Alphaproteobacteria;o__RF32;f__;g__
##                                                                           0.019484449
```

```
## Euryarchaeota;c__Methanobacteria;o__Methanobacteriales;f__Methanobacteriaceae;g__Methanobrevibacter
##                                                                                    0.028070760
##                                        Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__
##                                                                                    0.030610821
##               Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium
##                                                                                    0.035887736
##                                        Proteobacteria;c__Alphaproteobacteria;o__;f__;g__
##                                                                                    0.045431657
```

Looks like there are some significant p-values. But did they happen just by chance? There are 97 columns in the genus table, so that means we did 97 tests, and about 5% of them should be $p < 0.05$ just by chance. To correct for this, we can adjust the p-values using the `p.adjust` function. Here we are correcting for multiple hypothesis testing use False Discovery Rate (FDR). The adjusted p-values are often called "q-values."

```r
qvals <- p.adjust(pvals,'fdr')

# print the lowest 10 q-values
sort(qvals)[1:10]
```

```
##                       Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella
##                                                                                    0.1378094
##                                                                        WPS-2;c__;o__;f__;g__
##                                                                                    0.1378094
##                                  Verrucomicrobia;c__Verruco-5;o__WCHB1-41;f__RFP12;g__
##                                                                                    0.1535096
##            Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Lactococcus
##                                                                                    0.2200850
##                      Firmicutes;c__Clostridia;o__Clostridiales;f__Christensenellaceae;g__
##                                                                                    0.2200850
##                                 Proteobacteria;c__Alphaproteobacteria;o__RF32;f__;g__
##                                                                                    0.3149986
## Euryarchaeota;c__Methanobacteria;o__Methanobacteriales;f__Methanobacteriaceae;g__Methanobrevibacter
##                                                                                    0.3711562
##                                        Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__
##                                                                                    0.3711562
##               Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium
##                                                                                    0.3867900
##                                        Proteobacteria;c__Alphaproteobacteria;o__;f__;g__
##                                                                                    0.4406871
```

Testing with generalized linear regression.

Let's test whether the residuals from linear regression were normally distributed for all of the taxa. To follow along, you can put the following code in a separate text file called `stats.r` (or whatever you want to call it, and then call it using `source('stats.r')`.

```r
ks.pvals <- numeric(ncol(genus))

# "name" the pvalues after the genera
names(ks.pvals) <- colnames(genus)

# Loop through the columns of the genus table, testing each one
```

```
for(i in 1:ncol(genus)) {
    fit <- lm(genus[,i] ~ map$AGE + map$COUNTRY)
    ks.pvals[i] <- ks.test(rstudent(fit), pnorm, mean=mean(rstudent(fit)), sd=sd(rstudent(fit)),exact=
}
```

```
## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test
```

```
## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test
```

```
## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test
```

```
## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test
```

```
## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test
```

```
## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test
```

```
## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test
```

```
## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test
```

```
## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test
```

```
## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test
```

```
## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test
```

```
## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test
```

```
## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
```

```
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test
```

```
## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
```

```
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test
```

```
## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

## Warning in ks.test(rstudent(fit), pnorm, mean = mean(rstudent(fit)), sd =
## sd(rstudent(fit)), : ties should not be present for the Kolmogorov-Smirnov
## test

# Now since we ran 97 tests we should correct for multiple hypothesis testing.
ks.qvals <- p.adjust(ks.pvals,'fdr')

# print the lowest 10 q-values
sort(ks.qvals)[1:10]

##                          Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__RF16;g__
##                                                                        1.076916e-14
##                    Firmicutes;c__Bacilli;o__Lactobacillales;f__Leuconostocaceae;g__
##                                                                        5.761502e-13
```

```
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Citrobacter
##                                                                    2.408703e-12
##          Lentisphaerae;c__[Lentisphaeria];o__Victivallales;f__Victivallaceae;g__Victivallis
##                                                                    2.278486e-11
##            Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__
##                                                                    2.695500e-10
##             Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__Alistipes
##                                                                    9.155081e-10
##          Tenericutes;c__Mollicutes;o__Anaeroplasmatales;f__Anaeroplasmataceae;g__Anaeroplasma
##                                                                    2.454626e-09
##              Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus
##                                                                    4.203562e-09
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Escherichia
##                                                                    4.203562e-09
##      Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Coprobacillus
##                                                                    4.632104e-09
```

Wow. There are a lot of these whose residuals are highly non-normal. Let's use a negative binomial distribution instead. We will use the `edgeR` package. If you don't have it, you can install it with:

```
source("https://bioconductor.org/biocLite.R")

# If the previous command doesn't work, try http://

biocLite("edgeR")
```

Now use a convenient wrapper function that I provided in `repo/src/wrap.edgeR.r`. Load this script using `source`:

```
source('../../src/wrap.edgeR.r')
```

Note: the negative binomial uses raw counts of sequences (rarefied or not), not the relative abundances. Therefore we must re-run `summarize_taxa.py` with the `-a` flag to use absolute abundance. **These commands are run on the command line (not in R)**

```
# (run on command line)
summarize_taxa.py -i otu_table.biom -L 6 -a -o taxa-absolute

# convert to JSON BIOM format to load into R using R biom package:
biom convert -i taxa-absolute/otu_table_L6.biom -o taxa-absolute/otu_table_L6_json.biom --to-json
```

Now we need to load the genus table again.

```
genus.biom <- read_biom('taxa-absolute/otu_table_L6_json.biom')
genus.a <- as.matrix(biom_data(genus.biom))
genus.a <- t(genus.a)
genus.a <- genus.a[common.ids,]
genus.a <- genus.a[,colMeans(genus.a > 0) >= .1]
colnames(genus.a) <- sapply(strsplit(colnames(genus.a),'p__'),'[',2)
```

First let us run the regression without covariates. The main function in this script is `glm.edgeR()`. This function will perform multiple hypothesis testing on all columns of a data matrix. It has the following main

parameters: -x: the independent variable. If discrete must be binary; OK to be continuous. -Y: A matrix with samples in rows and dependent variables in columns -covariates: a matrix of additional covariates you want to control for (default NULL)

```
result <- glm.edgeR(x=map$AGE, Y=genus.a)
```

```
## Making DGEList...
## calcNormFactors...
## estimate common dispersion...
## estimate trended dispersion...
## estimate tagwise dispersion...
## fit glm...
## likelihood ratio test...
```

We can print the top "tags" (genera) using the topTags function. Note that two genera are significant even after correction for multiple hypothesis testing.

```
topTags(result)
```

```
## Coefficient:  x
##
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Lactococcus                     0.1(
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus                   -0.0(
## Firmicutes;c__Clostridia;o__Clostridiales;f__Christensenellaceae;g__                             0.0(
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium                -0.0:
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Veillonella                      -0.0(
## Actinobacteria;c__Actinobacteria;o__Bifidobacteriales;f__Bifidobacteriaceae;g__Bifidobacterium   -0.0(
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__              0.0:
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Citrobacter   0.1:
## Proteobacteria;c__Alphaproteobacteria;o__RF32;f__;g__                                            0.0!
## Verrucomicrobia;c__Verruco-5;o__WCHB1-41;f__RFP12;g__                                            0.0:
##                                                                                                  lo(
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Lactococcus                      11.0(
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus                    13.3`
## Firmicutes;c__Clostridia;o__Clostridiales;f__Christensenellaceae;g__                             12.9!
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium                 15.8:
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Veillonella                      11.0(
## Actinobacteria;c__Actinobacteria;o__Bifidobacteriales;f__Bifidobacteriaceae;g__Bifidobacterium   14.1
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__             13.3`
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Citrobacter 10.9:
## Proteobacteria;c__Alphaproteobacteria;o__RF32;f__;g__                                            11.5(
## Verrucomicrobia;c__Verruco-5;o__WCHB1-41;f__RFP12;g__                                            10.4`
##
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Lactococcus                      11.9:
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus                    11.4:
## Firmicutes;c__Clostridia;o__Clostridiales;f__Christensenellaceae;g__                              4.2:
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium                 4.2:
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Veillonella                       3.4:
## Actinobacteria;c__Actinobacteria;o__Bifidobacteriales;f__Bifidobacteriaceae;g__Bifidobacterium    3.3(
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__              3.0(
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Citrobacter   2.9(
## Proteobacteria;c__Alphaproteobacteria;o__RF32;f__;g__                                             2.7:
## Verrucomicrobia;c__Verruco-5;o__WCHB1-41;f__RFP12;g__                                             2.1`
```

```
## 
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Lactococcus                           0.000
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus                          0.000
## Firmicutes;c__Clostridia;o__Clostridiales;f__Christensenellaceae;g__                                    0.039
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium                       0.039
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Veillonella                            0.064
## Actinobacteria;c__Actinobacteria;o__Bifidobacteriales;f__Bifidobacteriaceae;g__Bifidobacterium          0.069
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__                    0.083
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Citrobacter 0.083
## Proteobacteria;c__Alphaproteobacteria;o__RF32;f__;g__                                                   0.098
## Verrucomicrobia;c__Verruco-5;o__WCHB1-41;f__RFP12;g__                                                   0.140
## 
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Lactococcus                           0.035
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus                          0.035
## Firmicutes;c__Clostridia;o__Clostridiales;f__Christensenellaceae;g__                                    0.969
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium                       0.969
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Veillonella                            0.98
## Actinobacteria;c__Actinobacteria;o__Bifidobacteriales;f__Bifidobacteriaceae;g__Bifidobacterium          0.98
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__                    0.98
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Citrobacter 0.98
## Proteobacteria;c__Alphaproteobacteria;o__RF32;f__;g__                                                   0.98
## Verrucomicrobia;c__Verruco-5;o__WCHB1-41;f__RFP12;g__                                                   0.98
```

However, we have not controlled for COUNTRY, which may be a confounder. We can do this with edgeR.

```r
result <- glm.edgeR(x=map$AGE, Y=genus.a, covariates=map$COUNTRY)
```

```
## Making DGEList...
## calcNormFactors...
## estimate common dispersion...
## estimate trended dispersion...
## estimate tagwise dispersion...
## fit glm...
## likelihood ratio test...
```

```r
topTags(result)
```

```
## Coefficient:  x
## 
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus                           -0.05
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__[Prevotella]                    0.048
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Lactococcus                              0.078
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium                        -0.015
## Proteobacteria;c__Alphaproteobacteria;o__RF32;f__;g__                                                     0.055
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella                           -0.030
## Actinobacteria;c__Actinobacteria;o__Bifidobacteriales;f__Bifidobacteriaceae;g__Bifidobacterium          -0.050
## Tenericutes;c__Mollicutes;o__Anaeroplasmatales;f__Anaeroplasmataceae;g__Anaeroplasma                    -0.024
## Firmicutes;c__Clostridia;o__Clostridiales;f__Christensenellaceae;g__                                      0.037
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Bulleidia                  0.028
##                                                                                                          logC
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus                           13.379
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__[Prevotella]                   13.873
```

```
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Lactococcus                              11.048
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium                         15.826
## Proteobacteria;c__Alphaproteobacteria;o__RF32;f__;g__                                                     11.565
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella                            17.375
## Actinobacteria;c__Actinobacteria;o__Bifidobacteriales;f__Bifidobacteriaceae;g__Bifidobacterium 14.141
## Tenericutes;c__Mollicutes;o__Anaeroplasmatales;f__Anaeroplasmataceae;g__Anaeroplasma                     10.471
## Firmicutes;c__Clostridia;o__Clostridiales;f__Christensenellaceae;g__                                      12.958
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Bulleidia                  11.681
##
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus                             6.890
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__[Prevotella]                     6.810
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Lactococcus                               6.494
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium                          5.120
## Proteobacteria;c__Alphaproteobacteria;o__RF32;f__;g__                                                      4.547
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella                             4.424
## Actinobacteria;c__Actinobacteria;o__Bifidobacteriales;f__Bifidobacteriaceae;g__Bifidobacterium 4.076
## Tenericutes;c__Mollicutes;o__Anaeroplasmatales;f__Anaeroplasmataceae;g__Anaeroplasma                      3.816
## Firmicutes;c__Clostridia;o__Clostridiales;f__Christensenellaceae;g__                                       3.665
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Bulleidia                   3.525
##
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus                             0.008
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__[Prevotella]                     0.009
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Lactococcus                               0.010
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium                          0.023
## Proteobacteria;c__Alphaproteobacteria;o__RF32;f__;g__                                                      0.032
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella                             0.035
## Actinobacteria;c__Actinobacteria;o__Bifidobacteriales;f__Bifidobacteriaceae;g__Bifidobacterium 0.043
## Tenericutes;c__Mollicutes;o__Anaeroplasmatales;f__Anaeroplasmataceae;g__Anaeroplasma                      0.050
## Firmicutes;c__Clostridia;o__Clostridiales;f__Christensenellaceae;g__                                       0.055
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Bulleidia                   0.060
##
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus                             0.3498
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__[Prevotella]                     0.3498
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Lactococcus                               0.3498
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium                          0.4871
## Proteobacteria;c__Alphaproteobacteria;o__RF32;f__;g__                                                      0.4871
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella                             0.4871
## Actinobacteria;c__Actinobacteria;o__Bifidobacteriales;f__Bifidobacteriaceae;g__Bifidobacterium 0.4871
## Tenericutes;c__Mollicutes;o__Anaeroplasmatales;f__Anaeroplasmataceae;g__Anaeroplasma                      0.4871
## Firmicutes;c__Clostridia;o__Clostridiales;f__Christensenellaceae;g__                                       0.4871
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Bulleidia                   0.4871
```

Note that no genera are significantly associated with age after controlling for country, and using the negative binomial as the assumed distribution for the residuals.

Let us test whether any genera are significantly associated with being from the USA, while controling for age:

```r
# make a vector that is TRUE if sample is from USA, FALSE otherwise
# the "==" means "test if equal"
is.USA <- map$COUNTRY == "GAZ:United States of America"

# run with is.USA as the independent variable
result <- glm.edgeR(x=is.USA, Y=genus.a, covariates=map$AGE)
```

```
## Making DGEList...
```

```
## calcNormFactors...
## estimate common dispersion...
## estimate trended dispersion...
## estimate tagwise dispersion...
## fit glm...
## likelihood ratio test...
```

topTags(result)

```
## Coefficient:  xTRUE
##                                                                                                          l
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__                                  6.348
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides                      5.31
## Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__Succinivibrionaceae;g__Succinivibrio -9.118
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Bulleidia          -5.810
## Cyanobacteria;c__4C0d-2;o__YS2;f__;g__                                                             -6.150
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae;g__Parabacteroides             3.040
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__;g__                                              -5.439
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__[Prevotella]             -5.104
## Spirochaetes;c__Spirochaetes;o__Spirochaetales;f__Spirochaetaceae;g__Treponema                    -7.063
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Clostridium                        -2.807
##                                                                                                       logC
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__                                 13.335
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides                     16.174
## Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__Succinivibrionaceae;g__Succinivibrio 14.638
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Bulleidia          11.682
## Cyanobacteria;c__4C0d-2;o__YS2;f__;g__                                                             12.512
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae;g__Parabacteroides            13.464
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__;g__                                             12.893
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__[Prevotella]            13.873
## Spirochaetes;c__Spirochaetes;o__Spirochaetales;f__Spirochaetaceae;g__Treponema                   12.576
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Clostridium                       12.711
##
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__                                176.08
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides                    141.20
## Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__Succinivibrionaceae;g__Succinivibrio  85.44
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Bulleidia           72.09
## Cyanobacteria;c__4C0d-2;o__YS2;f__;g__                                                             60.40
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae;g__Parabacteroides            48.9
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__;g__                                             47.72
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__[Prevotella]            42.17
## Spirochaetes;c__Spirochaetes;o__Spirochaetales;f__Spirochaetaceae;g__Treponema                    31.76
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Clostridium                        31.65
##
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__                                3.4640
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides                    1.4095
## Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__Succinivibrionaceae;g__Succinivibrio 2.3880
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Bulleidia           2.0551
## Cyanobacteria;c__4C0d-2;o__YS2;f__;g__                                                             7.5050
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae;g__Parabacteroides            2.6758
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__;g__                                             4.9129
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__[Prevotella]            8.3353
## Spirochaetes;c__Spirochaetes;o__Spirochaetales;f__Spirochaetaceae;g__Treponema                    1.7359
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Clostridium                        1.8430
```

```
## 
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__                           3.3601
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides                6.8361
## Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__Succinivibrionaceae;g__Succinivibrio 7.7212
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Bulleidia        4.9838
## Cyanobacteria;c__4C0d-2;o__YS2;f__;g__                                                         1.4559
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae;g__Parabacteroides          4.3258
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__;g__                                          6.8079
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__[Prevotella]          1.0100
## Spirochaetes;c__Spirochaetes;o__Spirochaetales;f__Spirochaetaceae;g__Treponema                 1.7883
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Clostridium                      1.7883
```

Now we're talking! Many significant results. Print all with:

```r
topTags(result, n=Inf)
```

```
## Coefficient:  xTRUE
## 
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides
## Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__Succinivibrionaceae;g__Succinivibrio       -
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Bulleidia
## Cyanobacteria;c__4C0d-2;o__YS2;f__;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae;g__Parabacteroides
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__[Prevotella]               -
## Spirochaetes;c__Spirochaetes;o__Spirochaetales;f__Spirochaetaceae;g__Treponema
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Clostridium
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__p-75-a5             -
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Barnesiellaceae];g__
## Proteobacteria;c__Alphaproteobacteria;o__;f__;g__                                                -
## Proteobacteria;c__Deltaproteobacteria;o__Desulfovibrionales;f__Desulfovibrionaceae;g__Bilophila
## Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__Succinivibrionaceae;g__Ruminobacter
## Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__              -
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Odoribacteraceae];g__Odoribacter
## Elusimicrobia;c__Elusimicrobia;o__Elusimicrobiales;f__Elusimicrobiaceae;g__
## Euryarchaeota;c__Methanobacteria;o__Methanobacteriales;f__Methanobacteriaceae;g__Methanosphaera
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__CF231
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Veillonella                      -
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Sarcina
## Verrucomicrobia;c__Verruco-5;o__WCHB1-41;f__RFP12;g__
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Holdemania
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Odoribacteraceae];g__Butyricimonas
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__                                        -
## Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Haemophilus
## Proteobacteria;c__Epsilonproteobacteria;o__Campylobacterales;f__Campylobacteraceae;g__Campylobacter  -
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Catenibacterium     -
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Serratia
```

```
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Coprococcus
## Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__Adlercreutzia
## Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Alcaligenaceae;g__Sutterella
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Leuconostocaceae;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Butyrivibrio
## Verrucomicrobia;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Verrucomicrobiaceae;g__Akkermansia
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Anaerovibrio
## Firmicutes;c__Clostridia;o__Clostridiales;f__[Mogibacteriaceae];g__
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Escherichia
## Firmicutes;c__Bacilli;o__Lactobacillales;f__;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__Alistipes
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__[Ruminococcus]
## Tenericutes;c__Mollicutes;o__Anaeroplasmatales;f__Anaeroplasmataceae;g__
## WPS-2;c__;o__;f__;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Roseburia
## Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__Slackia
## Tenericutes;c__Mollicutes;o__Anaeroplasmatales;f__Anaeroplasmataceae;g__Anaeroplasma
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Blautia
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Ruminococcus
## Firmicutes;c__Clostridia;o__;f__;g__
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus
## Tenericutes;c__Mollicutes;o__RF39;f__;g__
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Enterococcaceae;g__Enterococcus
## Firmicutes;c__Clostridia;o__Clostridiales;f__Peptococcaceae;g__Peptococcus
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__RF16;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__02d06
## Lentisphaerae;c__[Lentisphaeria];o__Victivallales;f__Victivallaceae;g__
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Lactococcus
## Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__
## Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Actinomycetaceae;g__Actinomyces
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Citrobacter
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Phascolarctobacterium
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Anaerostipes
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__SMB53
## Firmicutes;c__Clostridia;o__Clostridiales;f__Christensenellaceae;g__
## Actinobacteria;c__Actinobacteria;o__Bifidobacteriales;f__Bifidobacteriaceae;g__Bifidobacterium
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Lachnospira
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__
## Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__Collinsella
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Dialister
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Dorea
## Euryarchaeota;c__Methanobacteria;o__Methanobacteriales;f__Methanobacteriaceae;g__Methanobrevibacter
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Epulopiscium
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Oscillospira
## Proteobacteria;c__Alphaproteobacteria;o__RF32;f__;g__
## Tenericutes;c__RF3;o__ML615J-28;f__;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Megasphaera
## Lentisphaerae;c__[Lentisphaeria];o__Victivallales;f__Victivallaceae;g__Victivallis
## Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae;g__Oxalobacter
## Firmicutes;c__Clostridia;o__Clostridiales;f__Dehalobacteriaceae;g__Dehalobacterium
## Firmicutes;c__Bacilli;o__Turicibacterales;f__Turicibacteraceae;g__Turicibacter
```

```
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__;g__
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__[Eubacterium]
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Lachnobacterium
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Coprobacillus
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Peptostreptococcaceae;g__
## Proteobacteria;c__Deltaproteobacteria;o__Desulfovibrionales;f__Desulfovibrionaceae;g__Desulfovibrio
##
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides
## Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__Succinivibrionaceae;g__Succinivibrio
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Bulleidia
## Cyanobacteria;c__4C0d-2;o__YS2;f__;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae;g__Parabacteroides
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__[Prevotella]
## Spirochaetes;c__Spirochaetes;o__Spirochaetales;f__Spirochaetaceae;g__Treponema
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Clostridium
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__p-75-a5
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Barnesiellaceae];g__
## Proteobacteria;c__Alphaproteobacteria;o__;f__;g__
## Proteobacteria;c__Deltaproteobacteria;o__Desulfovibrionales;f__Desulfovibrionaceae;g__Bilophila
## Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__Succinivibrionaceae;g__Ruminobacter
## Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Odoribacteraceae];g__Odoribacter
## Elusimicrobia;c__Elusimicrobia;o__Elusimicrobiales;f__Elusimicrobiaceae;g__
## Euryarchaeota;c__Methanobacteria;o__Methanobacteriales;f__Methanobacteriaceae;g__Methanosphaera
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__CF231
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Veillonella
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Sarcina
## Verrucomicrobia;c__Verruco-5;o__WCHB1-41;f__RFP12;g__
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Holdemania
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Odoribacteraceae];g__Butyricimonas
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__
## Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Haemophilus
## Proteobacteria;c__Epsilonproteobacteria;o__Campylobacterales;f__Campylobacteraceae;g__Campylobacter
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Catenibacterium
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Serratia
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Coprococcus
## Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__Adlercreutzia
## Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Alcaligenaceae;g__Sutterella
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Leuconostocaceae;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Butyrivibrio
## Verrucomicrobia;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Verrucomicrobiaceae;g__Akkermansia
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Anaerovibrio
## Firmicutes;c__Clostridia;o__Clostridiales;f__[Mogibacteriaceae];g__
```

```
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Escherichia
## Firmicutes;c__Bacilli;o__Lactobacillales;f__;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__Alistipes
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__[Ruminococcus]
## Tenericutes;c__Mollicutes;o__Anaeroplasmatales;f__Anaeroplasmataceae;g__
## WPS-2;c__;o__;f__;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Roseburia
## Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__Slackia
## Tenericutes;c__Mollicutes;o__Anaeroplasmatales;f__Anaeroplasmataceae;g__Anaeroplasma
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Blautia
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Ruminococcus
## Firmicutes;c__Clostridia;o__;f__;g__
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus
## Tenericutes;c__Mollicutes;o__RF39;f__;g__
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Enterococcaceae;g__Enterococcus
## Firmicutes;c__Clostridia;o__Clostridiales;f__Peptococcaceae;g__Peptococcus
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__RF16;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__02d06
## Lentisphaerae;c__[Lentisphaeria];o__Victivallales;f__Victivallaceae;g__
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Lactococcus
## Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__
## Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Actinomycetaceae;g__Actinomyces
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Citrobacter
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Phascolarctobacterium
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Anaerostipes
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__SMB53
## Firmicutes;c__Clostridia;o__Clostridiales;f__Christensenellaceae;g__
## Actinobacteria;c__Actinobacteria;o__Bifidobacteriales;f__Bifidobacteriaceae;g__Bifidobacterium
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Lachnospira
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__
## Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__Collinsella
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Dialister
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Dorea
## Euryarchaeota;c__Methanobacteria;o__Methanobacteriales;f__Methanobacteriaceae;g__Methanobrevibacter
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Epulopiscium
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Oscillospira
## Proteobacteria;c__Alphaproteobacteria;o__RF32;f__;g__
## Tenericutes;c__RF3;o__ML615J-28;f__;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Megasphaera
## Lentisphaerae;c__[Lentisphaeria];o__Victivallales;f__Victivallaceae;g__Victivallis
## Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae;g__Oxalobacter
## Firmicutes;c__Clostridia;o__Clostridiales;f__Dehalobacteriaceae;g__Dehalobacterium
## Firmicutes;c__Bacilli;o__Turicibacterales;f__Turicibacteraceae;g__Turicibacter
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__;g__
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__[Eubacterium]
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Lachnobacterium
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Coprobacillus
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Peptostreptococcaceae;g__
## Proteobacteria;c__Deltaproteobacteria;o__Desulfovibrionales;f__Desulfovibrionaceae;g__Desulfovibrio
##
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__
```

```
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides
## Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__Succinivibrionaceae;g__Succinivibrio
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Bulleidia
## Cyanobacteria;c__4C0d-2;o__YS2;f__;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae;g__Parabacteroides
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__[Prevotella]
## Spirochaetes;c__Spirochaetes;o__Spirochaetales;f__Spirochaetaceae;g__Treponema
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Clostridium
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__p-75-a5
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Barnesiellaceae];g__
## Proteobacteria;c__Alphaproteobacteria;o__;f__;g__
## Proteobacteria;c__Deltaproteobacteria;o__Desulfovibrionales;f__Desulfovibrionaceae;g__Bilophila
## Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__Succinivibrionaceae;g__Ruminobacter
## Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Odoribacteraceae];g__Odoribacter
## Elusimicrobia;c__Elusimicrobia;o__Elusimicrobiales;f__Elusimicrobiaceae;g__
## Euryarchaeota;c__Methanobacteria;o__Methanobacteriales;f__Methanobacteriaceae;g__Methanosphaera
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__CF231
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Veillonella
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Sarcina
## Verrucomicrobia;c__Verruco-5;o__WCHB1-41;f__RFP12;g__
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Holdemania
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Odoribacteraceae];g__Butyricimonas
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__
## Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Haemophilus
## Proteobacteria;c__Epsilonproteobacteria;o__Campylobacterales;f__Campylobacteraceae;g__Campylobacter
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Catenibacterium
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Serratia
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Coprococcus
## Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__Adlercreutzia
## Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Alcaligenaceae;g__Sutterella
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Leuconostocaceae;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Butyrivibrio
## Verrucomicrobia;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Verrucomicrobiaceae;g__Akkermansia
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Anaerovibrio
## Firmicutes;c__Clostridia;o__Clostridiales;f__[Mogibacteriaceae];g__
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Escherichia
## Firmicutes;c__Bacilli;o__Lactobacillales;f__;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__Alistipes
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__[Ruminococcus]
## Tenericutes;c__Mollicutes;o__Anaeroplasmatales;f__Anaeroplasmataceae;g__
## WPS-2;c__;o__;f__;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Roseburia
## Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__Slackia
## Tenericutes;c__Mollicutes;o__Anaeroplasmatales;f__Anaeroplasmataceae;g__Anaeroplasma
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Blautia
```

```
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Ruminococcus
## Firmicutes;c__Clostridia;o__;f__;g__
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus
## Tenericutes;c__Mollicutes;o__RF39;f__;g__
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Enterococcaceae;g__Enterococcus
## Firmicutes;c__Clostridia;o__Clostridiales;f__Peptococcaceae;g__Peptococcus
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__RF16;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__02d06
## Lentisphaerae;c__[Lentisphaeria];o__Victivallales;f__Victivallaceae;g__
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Lactococcus
## Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__
## Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Actinomycetaceae;g__Actinomyces
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Citrobacter
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Phascolarctobacterium
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Anaerostipes
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__SMB53
## Firmicutes;c__Clostridia;o__Clostridiales;f__Christensenellaceae;g__
## Actinobacteria;c__Actinobacteria;o__Bifidobacteriales;f__Bifidobacteriaceae;g__Bifidobacterium
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Lachnospira
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__
## Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__Collinsella
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Dialister
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Dorea
## Euryarchaeota;c__Methanobacteria;o__Methanobacteriales;f__Methanobacteriaceae;g__Methanobrevibacter
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Epulopiscium
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Oscillospira
## Proteobacteria;c__Alphaproteobacteria;o__RF32;f__;g__
## Tenericutes;c__RF3;o__ML615J-28;f__;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Megasphaera
## Lentisphaerae;c__[Lentisphaeria];o__Victivallales;f__Victivallaceae;g__Victivallis
## Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae;g__Oxalobacter
## Firmicutes;c__Clostridia;o__Clostridiales;f__Dehalobacteriaceae;g__Dehalobacterium
## Firmicutes;c__Bacilli;o__Turicibacterales;f__Turicibacteraceae;g__Turicibacter
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__;g__
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__[Eubacterium]
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Lachnobacterium
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Coprobacillus
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Peptostreptococcaceae;g__
## Proteobacteria;c__Deltaproteobacteria;o__Desulfovibrionales;f__Desulfovibrionaceae;g__Desulfovibrio
## 
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides
## Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__Succinivibrionaceae;g__Succinivibrio
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Bulleidia
## Cyanobacteria;c__4C0d-2;o__YS2;f__;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae;g__Parabacteroides
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__[Prevotella]
## Spirochaetes;c__Spirochaetes;o__Spirochaetales;f__Spirochaetaceae;g__Treponema
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Clostridium
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella
```

```
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__p-75-a5
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Barnesiellaceae];g__
## Proteobacteria;c__Alphaproteobacteria;o__;f__;g__
## Proteobacteria;c__Deltaproteobacteria;o__Desulfovibrionales;f__Desulfovibrionaceae;g__Bilophila
## Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__Succinivibrionaceae;g__Ruminobacter
## Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Odoribacteraceae];g__Odoribacter
## Elusimicrobia;c__Elusimicrobia;o__Elusimicrobiales;f__Elusimicrobiaceae;g__
## Euryarchaeota;c__Methanobacteria;o__Methanobacteriales;f__Methanobacteriaceae;g__Methanosphaera
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__CF231
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Veillonella
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Sarcina
## Verrucomicrobia;c__Verruco-5;o__WCHB1-41;f__RFP12;g__
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Holdemania
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Odoribacteraceae];g__Butyricimonas
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__
## Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Haemophilus
## Proteobacteria;c__Epsilonproteobacteria;o__Campylobacterales;f__Campylobacteraceae;g__Campylobacter
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Catenibacterium
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Serratia
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Coprococcus
## Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__Adlercreutzia
## Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Alcaligenaceae;g__Sutterella
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Leuconostocaceae;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Butyrivibrio
## Verrucomicrobia;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Verrucomicrobiaceae;g__Akkermansia
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Anaerovibrio
## Firmicutes;c__Clostridia;o__Clostridiales;f__[Mogibacteriaceae];g__
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Escherichia
## Firmicutes;c__Bacilli;o__Lactobacillales;f__;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__Alistipes
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__[Ruminococcus]
## Tenericutes;c__Mollicutes;o__Anaeroplasmatales;f__Anaeroplasmataceae;g__
## WPS-2;c__;o__;f__;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Roseburia
## Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__Slackia
## Tenericutes;c__Mollicutes;o__Anaeroplasmatales;f__Anaeroplasmataceae;g__Anaeroplasma
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Blautia
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Ruminococcus
## Firmicutes;c__Clostridia;o__;f__;g__
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus
## Tenericutes;c__Mollicutes;o__RF39;f__;g__
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Enterococcaceae;g__Enterococcus
## Firmicutes;c__Clostridia;o__Clostridiales;f__Peptococcaceae;g__Peptococcus
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__RF16;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__02d06
## Lentisphaerae;c__[Lentisphaeria];o__Victivallales;f__Victivallaceae;g__
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Lactococcus
```

```
## Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__
## Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Actinomycetaceae;g__Actinomyces
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Citrobacter
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Phascolarctobacterium
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Anaerostipes
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__SMB53
## Firmicutes;c__Clostridia;o__Clostridiales;f__Christensenellaceae;g__
## Actinobacteria;c__Actinobacteria;o__Bifidobacteriales;f__Bifidobacteriaceae;g__Bifidobacterium
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Lachnospira
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__
## Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__Collinsella
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Dialister
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Dorea
## Euryarchaeota;c__Methanobacteria;o__Methanobacteriales;f__Methanobacteriaceae;g__Methanobrevibacter
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Epulopiscium
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Oscillospira
## Proteobacteria;c__Alphaproteobacteria;o__RF32;f__;g__
## Tenericutes;c__RF3;o__ML615J-28;f__;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Megasphaera
## Lentisphaerae;c__[Lentisphaeria];o__Victivallales;f__Victivallaceae;g__Victivallis
## Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae;g__Oxalobacter
## Firmicutes;c__Clostridia;o__Clostridiales;f__Dehalobacteriaceae;g__Dehalobacterium
## Firmicutes;c__Bacilli;o__Turicibacterales;f__Turicibacteraceae;g__Turicibacter
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__;g__
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__[Eubacterium]
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Lachnobacterium
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Coprobacillus
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Peptostreptococcaceae;g__
## Proteobacteria;c__Deltaproteobacteria;o__Desulfovibrionales;f__Desulfovibrionaceae;g__Desulfovibrio
##
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides
## Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__Succinivibrionaceae;g__Succinivibrio
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Bulleidia
## Cyanobacteria;c__4C0d-2;o__YS2;f__;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae;g__Parabacteroides
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__[Prevotella]
## Spirochaetes;c__Spirochaetes;o__Spirochaetales;f__Spirochaetaceae;g__Treponema
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Clostridium
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__p-75-a5
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Barnesiellaceae];g__
## Proteobacteria;c__Alphaproteobacteria;o__;f__;g__
## Proteobacteria;c__Deltaproteobacteria;o__Desulfovibrionales;f__Desulfovibrionaceae;g__Bilophila
## Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__Succinivibrionaceae;g__Ruminobacter
## Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Odoribacteraceae];g__Odoribacter
```

```
## Elusimicrobia;c__Elusimicrobia;o__Elusimicrobiales;f__Elusimicrobiaceae;g__
## Euryarchaeota;c__Methanobacteria;o__Methanobacteriales;f__Methanobacteriaceae;g__Methanosphaera
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__CF231
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Veillonella
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Sarcina
## Verrucomicrobia;c__Verruco-5;o__WCHB1-41;f__RFP12;g__
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Holdemania
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Odoribacteraceae];g__Butyricimonas
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__
## Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Haemophilus
## Proteobacteria;c__Epsilonproteobacteria;o__Campylobacterales;f__Campylobacteraceae;g__Campylobacter
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Catenibacterium
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Serratia
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Coprococcus
## Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__Adlercreutzia
## Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Alcaligenaceae;g__Sutterella
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Leuconostocaceae;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Butyrivibrio
## Verrucomicrobia;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Verrucomicrobiaceae;g__Akkermansia
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Anaerovibrio
## Firmicutes;c__Clostridia;o__Clostridiales;f__[Mogibacteriaceae];g__
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Escherichia
## Firmicutes;c__Bacilli;o__Lactobacillales;f__;g__
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__Alistipes
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__[Ruminococcus]
## Tenericutes;c__Mollicutes;o__Anaeroplasmatales;f__Anaeroplasmataceae;g__
## WPS-2;c__;o__;f__;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Roseburia
## Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__Slackia
## Tenericutes;c__Mollicutes;o__Anaeroplasmatales;f__Anaeroplasmataceae;g__Anaeroplasma
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Blautia
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Ruminococcus
## Firmicutes;c__Clostridia;o__;f__;g__
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus
## Tenericutes;c__Mollicutes;o__RF39;f__;g__
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Enterococcaceae;g__Enterococcus
## Firmicutes;c__Clostridia;o__Clostridiales;f__Peptococcaceae;g__Peptococcus
## Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__RF16;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__02d06
## Lentisphaerae;c__[Lentisphaeria];o__Victivallales;f__Victivallaceae;g__
## Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Lactococcus
## Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__
## Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Actinomycetaceae;g__Actinomyces
## Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Citrobacter
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Phascolarctobacterium
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Anaerostipes
## Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__SMB53
## Firmicutes;c__Clostridia;o__Clostridiales;f__Christensenellaceae;g__
## Actinobacteria;c__Actinobacteria;o__Bifidobacteriales;f__Bifidobacteriaceae;g__Bifidobacterium
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Lachnospira
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__
```

```
## Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__Collinsella
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Dialister
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Dorea
## Euryarchaeota;c__Methanobacteria;o__Methanobacteriales;f__Methanobacteriaceae;g__Methanobrevibacter
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Epulopiscium
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Oscillospira
## Proteobacteria;c__Alphaproteobacteria;o__RF32;f__;g__
## Tenericutes;c__RF3;o__ML615J-28;f__;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Megasphaera
## Lentisphaerae;c__[Lentisphaeria];o__Victivallales;f__Victivallaceae;g__Victivallis
## Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae;g__Oxalobacter
## Firmicutes;c__Clostridia;o__Clostridiales;f__Dehalobacteriaceae;g__Dehalobacterium
## Firmicutes;c__Bacilli;o__Turicibacterales;f__Turicibacteraceae;g__Turicibacter
## Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__;g__
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__[Eubacterium]
## Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Lachnobacterium
## Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Coprobacillus
## Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__
## Firmicutes;c__Clostridia;o__Clostridiales;f__Peptostreptococcaceae;g__
## Proteobacteria;c__Deltaproteobacteria;o__Desulfovibrionales;f__Desulfovibrionaceae;g__Desulfovibrio
```

Write the results to a tab-delimited file (to open in Excel) with:

```r
write.table(topTags(result, n=Inf)$table, file='edgeR_results.txt',sep='\t',quote=FALSE, col.names=NA)
```

How many genera were significantly associated with the USA?

```r
sum(topTags(result,n=Inf)$table$FDR <= 0.05)
```

```
## [1] 62
```

Non-parametric tests

Whenever we do not need to control for confounding variables, we can use non-parametric tests. These are the safest because they don't rely on any null distribution (e.g. normal). They typically consider only the **ranks** of values (order of values), not the actual values themselves.

For testing differences between two categories, we can use the Mann-Whitney U test, sometimes called the Wilcoxon Signed Rank test or Wilcoxon Rank Sum test. There are slight differences between these tests. Here we will test the difference in **Prevotella** abundance between USA and non-USA. Make sure to use the relative abundances, not the absolute abundances.

```r
wilcox.test(prevotella ~ is.USA, exact=FALSE)
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  prevotella by is.USA
## W = 934, p-value = 7.732e-10
## alternative hypothesis: true location shift is not equal to 0
```

```
# get the exact p-value
wilcox.test(prevotella ~ is.USA, exact=FALSE)$p.value
```

```
## [1] 7.732064e-10
```

We can do a test for differentiation across multiple categories, analogous to ANOVA, using the Kruskal-Wallis test.

```
kruskal.test(prevotella ~ map$COUNTRY)
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  prevotella by map$COUNTRY
## Kruskal-Wallis chi-squared = 42.001, df = 2, p-value = 7.581e-10
```

For continuous variables, we can use Spearman correlation instead of Pearson correlation.

```
cor.test(prevotella, map$AGE, method='spearman', exact=FALSE)
```

```
##
##  Spearman's rank correlation rho
##
## data:  prevotella and map$AGE
## S = 60496, p-value = 0.033
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.2628328
```