

R workflow

Gajender Aleti

10/09/2020

Import libraries

```
shhh <- suppressPackageStartupMessages # It's a library, so shhh!  
shhh(library(psych))  
shhh(library(decontam))
```

```
## Warning in register(): Can't find generic 'scale_type' in package ggplot2 to  
## register S3 method.
```

```
shhh(library(phyloseq))  
shhh(library(pairwiseAdonis))  
shhh(library(tidyverse))  
shhh(library(qiime2R))  
shhh(library(vegan))  
shhh(library(RVAideMemoire))  
shhh(library(afex))  
shhh(library(dplyr))  
shhh(library(ggplot2))  
shhh(library(readxl))  
shhh(library(tinytex))
```

Pearson correlations - associations among continuous and categorical metadata variables

```
mapfile4<-read.csv("/home/gajender/Documents/R_files/map_file4.csv", header=T, sep="," , row.names=1)  
mapfile <- data.frame(mapfile4[,])  
mat_map2 <- mapfile[which(!duplicated(mapfile$anonymized_name)), c("Age", "BMI", "Total_fat", "Trunk_fa  
mat_map2 <- as.data.frame(apply(mat_map2, 2, FUN = function(x){as.numeric(as.character(x))}))
```

```
## Warning in FUN(newX[, i], ...): NAs introduced by coercion
```

```
## Warning in FUN(newX[, i], ...): NAs introduced by coercion
```

```
## Warning in FUN(newX[, i], ...): NAs introduced by coercion
```

```
## Warning in FUN(newX[, i], ...): NAs introduced by coercion
```

```

mat_map2$BDI_T <- log(mat_map2$BDI_t + 1)
mat_map2$BDI_S <- log(mat_map2$BDI_s + 1)
mat_map2$BDI_C <- log(mat_map2$BDI_c + 1)
mat_map2$BARIC <- (mat_map2$baric)
corr.test(as.matrix(mat_map2), method = "spearman", adjust = "fdr")

## Call:corr.test(x = as.matrix(mat_map2), method = "spearman", adjust = "fdr")
## Correlation matrix
##           Age   BMI Total_fat Trunk_fat BDI_t BDI_s BDI_c baric BDI_T BDI_S
## Age       1.00  0.10      0.10      0.13  0.09  0.18 -0.01 -0.06  0.09  0.18
## BMI       0.10  1.00      0.55      0.61  0.29  0.22  0.27 -0.38  0.29  0.22
## Total_fat 0.10  0.55      1.00      0.96  0.39  0.38  0.29 -0.24  0.39  0.38
## Trunk_fat 0.13  0.61      0.96      1.00  0.33  0.34  0.23 -0.32  0.33  0.34
## BDI_t      0.09  0.29      0.39      0.33  1.00  0.93  0.89 -0.03  1.00  0.93
## BDI_s      0.18  0.22      0.38      0.34  0.93  1.00  0.69 -0.09  0.93  1.00
## BDI_c     -0.01  0.27      0.29      0.23  0.89  0.69  1.00  0.10  0.89  0.69
## baric     -0.06 -0.38     -0.24     -0.32 -0.03 -0.09  0.10  1.00 -0.03 -0.09
## BDI_T      0.09  0.29      0.39      0.33  1.00  0.93  0.89 -0.03  1.00  0.93
## BDI_S      0.18  0.22      0.38      0.34  0.93  1.00  0.69 -0.09  0.93  1.00
## BDI_C     -0.01  0.27      0.29      0.23  0.89  0.69  1.00  0.10  0.89  0.69
## BARIC     -0.06 -0.38     -0.24     -0.32 -0.03 -0.09  0.10  1.00 -0.03 -0.09
##           BDI_C BARIC
## Age       -0.01 -0.06
## BMI        0.27 -0.38
## Total_fat  0.29 -0.24
## Trunk_fat  0.23 -0.32
## BDI_t      0.89 -0.03
## BDI_s      0.69 -0.09
## BDI_c      1.00  0.10
## baric      0.10  1.00
## BDI_T      0.89 -0.03
## BDI_S      0.69 -0.09
## BDI_C      1.00  0.10
## BARIC      0.10  1.00
## Sample Size
##           Age BMI Total_fat Trunk_fat BDI_t BDI_s BDI_c baric BDI_T BDI_S BDI_C
## Age       65  65      63      63      65      65      65      61      65      65      65
## BMI       65  66      64      64      66      66      66      61      66      66      66
## Total_fat 63  64      64      64      64      64      64      59      64      64      64
## Trunk_fat 63  64      64      64      64      64      64      59      64      64      64
## BDI_t     65  66      64      64      66      66      66      61      66      66      66
## BDI_s     65  66      64      64      66      66      66      61      66      66      66
## BDI_c     65  66      64      64      66      66      66      61      66      66      66
## baric     61  61      59      59      61      61      61      61      61      61      61
## BDI_T     65  66      64      64      66      66      66      61      66      66      66
## BDI_S     65  66      64      64      66      66      66      61      66      66      66
## BDI_C     65  66      64      64      66      66      66      61      66      66      66
## BARIC     61  61      59      59      61      61      61      61      61      61      61
##           BARIC
## Age          61
## BMI          61
## Total_fat    59
## Trunk_fat    59

```

```

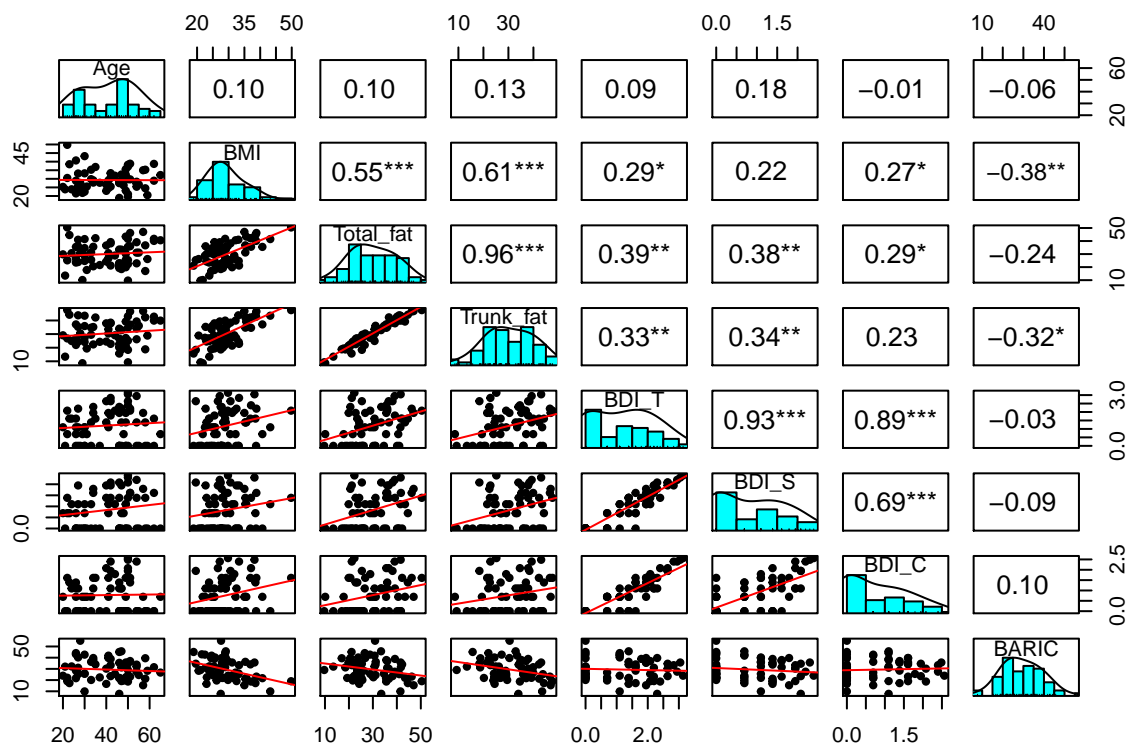
## BDI_t      61
## BDI_s      61
## BDI_c      61
## baric      61
## BDI_T      61
## BDI_S      61
## BDI_C      61
## BARIC      61
## Probability values (Entries above the diagonal are adjusted for multiple tests.)
##           Age  BMI Total_fat Trunk_fat BDI_t BDI_s BDI_c baric BDI_T BDI_S
## Age      0.00 0.56      0.56      0.45 0.56 0.24 0.92 0.72 0.56 0.24
## BMI      0.45 0.00      0.00      0.00 0.04 0.12 0.05 0.01 0.04 0.12
## Total_fat 0.42 0.00      0.00      0.00 0.00 0.01 0.04 0.10 0.00 0.01
## Trunk_fat 0.31 0.00      0.00      0.00 0.02 0.02 0.12 0.03 0.02 0.02
## BDI_t     0.47 0.02      0.00      0.01 0.00 0.00 0.00 0.84 0.00 0.00
## BDI_s     0.16 0.08      0.00      0.01 0.00 0.00 0.00 0.56 0.00 0.00
## BDI_c     0.92 0.03      0.02      0.07 0.00 0.00 0.00 0.56 0.00 0.00
## baric     0.66 0.00      0.06      0.01 0.82 0.50 0.46 0.00 0.84 0.56
## BDI_T     0.47 0.02      0.00      0.01 0.00 0.00 0.00 0.82 0.00 0.00
## BDI_S     0.16 0.08      0.00      0.01 0.00 0.00 0.00 0.50 0.00 0.00
## BDI_C     0.92 0.03      0.02      0.07 0.00 0.00 0.00 0.46 0.00 0.00
## BARIC     0.66 0.00      0.06      0.01 0.82 0.50 0.46 0.00 0.82 0.50
##           BDI_C BARIC
## Age      0.92 0.72
## BMI      0.05 0.01
## Total_fat 0.04 0.10
## Trunk_fat 0.12 0.03
## BDI_t     0.00 0.84
## BDI_s     0.00 0.56
## BDI_c     0.00 0.56
## baric     0.56 0.00
## BDI_T     0.00 0.84
## BDI_S     0.00 0.56
## BDI_C     0.00 0.56
## BARIC     0.46 0.00
##
## To see confidence intervals of the correlations, print with the short=FALSE option

```

```

pairs.panels(mat_map2[, -c(5, 6, 7, 8)], method="spearman", lm=TRUE, density=TRUE, ellipses=F, stars=T, c

```



Decontamination workflow

decontam: https://benjjneb.github.io/decontam/vignettes/decontam_intro.html Identify contaminants in marker-gene and metagenomics data based on prevalence (contaminating DNA can come from several sources, such as the reagents). Prevalence (presence/absence across samples) of each sequence feature in true positive samples is compared to the prevalence in negative controls to identify contaminants Import ASV table, taxonomy and metadata files

```
tax<-read.table("/home/gajender/Documents/projects/git/MicrobiomeDecontamination/taxonomy.tsv", sep= "\t")
otu<-read.table("/home/gajender/Documents/projects/git/MicrobiomeDecontamination/feature-table.txt", sep= "\t")
sample<-read.table("/home/gajender/Documents/projects/git/MicrobiomeDecontamination/metadata.txt", sep= "\t")
OTU = otu_table(otu, taxa_are_rows = TRUE)
TAX = tax_table(as.matrix(tax))
sampledata= sample_data(as.data.frame(sample))
```

Generate a phyloseq object

```
ps=phyloseq(OTU,sampledata,TAX)
ps
```

```
## phyloseq-class experiment-level object
## otu_table() OTU Table: [ 2248 taxa and 369 samples ]
## sample_data() Sample Data: [ 369 samples by 167 sample variables ]
## tax_table() Taxonomy Table: [ 2248 taxa by 2 taxonomic ranks ]
```

```
sample_data(ps)$is.neg <- sample_data(ps)$Sample_or_Control == "Control Sample"
```

Identify contaminants prevalence in blank and true samples with default setting

```
contamdf.prev <- isContaminant(ps, method="prevalence", neg="is.neg")
table(contamdf.prev$contaminant)
```

```
##
## FALSE TRUE
## 2224 24
```

```
head(which(contamdf.prev$contaminant))
```

```
## [1] 212 368 409 414 471 545
```

Stringent filtering (threshold=0.5) to identify contaminant sequences that are more prevalent in negative controls than in positive samples

```
contamdf.prev05 <- isContaminant(ps, method="prevalence", neg="is.neg", threshold=0.5)
table(contamdf.prev05$contaminant)
```

```
##
## FALSE TRUE
## 2114 134
```

```
head(which(contamdf.prev05$contaminant))
```

```
## [1] 20 83 210 212 214 216
```

```
write.table(contamdf.prev05, file = "/home/gajender/Documents/projects/git/MicrobiomeDecontamination/contamdf.prev05.txt")
```

Beta diversity analysis

```
physeq<-qza_to_phyloseq("/home/gajender/Documents/R_files/re_analysis/sample-decontam-filtered-table.qza")
```

PERMANOVA on weighted UniFrac distances

```
wUF.dist <- distance(physeq, method = "wunifrac")
```

```
## Warning in matrix(tree$edge[order(tree$edge[, 1]), ][, 2], byrow = TRUE, : data
## length [2975] is not a sub-multiple or multiple of the number of rows [1488]
```

```
adonis_wUF<-adonis2(wUF.dist~as.numeric(age) + as.factor(sex) + as.factor(race) + as.factor(sample_coll)
adonis_wUF
```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = wUF.dist ~ as.numeric(age) + as.factor(sex) + as.factor(race) + as.factor(sample_collection_time_hongcode))
##
```

| | Df | SumOfSqs | R2 | F | Pr(>F) |
|---|-----|----------|---------|--------|--------|
| ## as.numeric(age) | 1 | 0.004642 | 0.02068 | 5.7647 | 0.001 |
| ## as.factor(sex) | 1 | 0.002785 | 0.01241 | 3.4586 | 0.001 |
| ## as.factor(race) | 3 | 0.007433 | 0.03312 | 3.0771 | 0.001 |
| ## as.factor(sample_collection_time_hongcode) | 4 | 0.003721 | 0.01658 | 1.1552 | 0.172 |
| ## as.numeric(del8) | 1 | 0.001702 | 0.00759 | 2.1141 | 0.012 |
| ## as.factor(bmicategory) | 1 | 0.001529 | 0.00681 | 1.8986 | 0.025 |
| ## as.factor(bditotalcategory) | 1 | 0.002384 | 0.01062 | 2.9603 | 0.003 |
| ## as.factor(bmicat_bditotalcat) | 1 | 0.002140 | 0.00953 | 2.6574 | 0.003 |
| ## Residual | 246 | 0.198085 | 0.88265 | | |
| ## Total | 259 | 0.224420 | 1.00000 | | |

```
##
## as.numeric(age) ***
## as.factor(sex) ***
## as.factor(race) ***
## as.factor(sample_collection_time_hongcode)
## as.numeric(del8) *
## as.factor(bmicategory) *
## as.factor(bditotalcategory) **
## as.factor(bmicat_bditotalcat) **
## Residual
## Total
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

PERMANOVA on unweighted UniFrac distances

```
uUF.dist <- distance(physeq, method = "uunifrac")
```

```
## Warning in matrix(tree$edge[order(tree$edge[, 1]), ][, 2], byrow = TRUE, : data
## length [2975] is not a sub-multiple or multiple of the number of rows [1488]
```

```
adonis_uUF<-adonis2(uUF.dist~as.numeric(age) + as.factor(sex) + as.factor(race) + as.factor(sample_collection_time_hongcode))
adonis_uUF
```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = uUF.dist ~ as.numeric(age) + as.factor(sex) + as.factor(race) + as.factor(sample_collection_time_hongcode))
##
```

| | Df | SumOfSqs | R2 | F | Pr(>F) |
|---|----|----------|---------|--------|--------|
| ## as.numeric(age) | 1 | 0.895 | 0.01404 | 3.8835 | 0.001 |
| ## as.factor(sex) | 1 | 0.585 | 0.00918 | 2.5376 | 0.001 |
| ## as.factor(race) | 3 | 2.017 | 0.03166 | 2.9181 | 0.001 |
| ## as.factor(sample_collection_time_hongcode) | 4 | 0.737 | 0.01157 | 0.7995 | 0.986 |

```
## as.numeric(del8)          1      1.085 0.01703 4.7090 0.001
## as.factor(bmicategory)    1      0.442 0.00694 1.9185 0.006
## as.factor(bditotalcategory) 1      0.650 0.01020 2.8213 0.001
## as.factor(bmicat_bditotalcat) 1      0.622 0.00977 2.7017 0.001
## Residual                246    56.673 0.88962
## Total                   259    63.705 1.00000
##
## as.numeric(age)          ***
## as.factor(sex)           ***
## as.factor(race)          ***
## as.factor(sample_collection_time_hongcode)
## as.numeric(del8)          ***
## as.factor(bmicategory)    **
## as.factor(bditotalcategory) ***
## as.factor(bmicat_bditotalcat) ***
## Residual
## Total
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

pairwiseAdonis on weighted UniFrac distances

```
pairwie_adonis_wUF<-pairwise.adonis2(wUF.dist ~ bmicat_bditotalcat, data = as(sample_data(physeq), "data"))
pairwie_adonis_wUF
```

```
## $parent_call
## [1] "wUF.dist ~ bmicat_bditotalcat , strata = Null"
##
## $OL_vs_OH
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
##              Df SumsOfSqs   MeanSqs F.Model      R2 Pr(>F)
## bmicat_bditotalcat  1    0.00183 0.00182986  2.0799 0.01066 0.006 **
## Residuals          193    0.16980 0.00087978      0.98934
## Total              194    0.17163      1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $OL_vs_HL
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
##              Df SumsOfSqs   MeanSqs F.Model      R2 Pr(>F)
## bmicat_bditotalcat  1    0.001571 0.00157065  1.9415 0.0166 0.033 *
## Residuals          115    0.093033 0.00080898      0.9834
## Total              116    0.094603      1.0000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```

##
## $OL_vs_HH
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
##           Df SumsOfSqs   MeanSqs F.Model      R2 Pr(>F)
## bmicat_bditotalcat  1  0.001133 0.00113283  1.4466 0.01516  0.089 .
## Residuals          94  0.073609 0.00078307           0.98484
## Total              95  0.074742           1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $OH_vs_HL
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
##           Df SumsOfSqs   MeanSqs F.Model      R2 Pr(>F)
## bmicat_bditotalcat  1  0.00196 0.00196034  2.1717 0.01323  0.008 **
## Residuals          162  0.14623 0.00090267           0.98677
## Total              163  0.14819           1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $OH_vs_HH
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
##           Df SumsOfSqs   MeanSqs F.Model      R2 Pr(>F)
## bmicat_bditotalcat  1  0.000952 0.00095210  1.0587 0.00745  0.373
## Residuals          141  0.126809 0.00089936           0.99255
## Total              142  0.127761           1.00000
##
## $HL_vs_HH
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
##           Df SumsOfSqs   MeanSqs F.Model      R2 Pr(>F)
## bmicat_bditotalcat  1  0.001283 0.00128320  1.6154 0.025  0.086 .
## Residuals          63  0.050044 0.00079435           0.975
## Total              64  0.051327           1.000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## attr("class")
## [1] "pwadstrata" "list"

```


pairwiseAdonis on unweighted UniFrac distances

```
pairwise_adonis_uUF<-pairwise.adonis2(uUF.dist ~ bmicat_bditotalcat, data = as(sample_data(physeq), "data"))
pairwise_adonis_uUF
```

```
## $parent_call
## [1] "uUF.dist ~ bmicat_bditotalcat , strata = Null"
##
## $OL_vs_OH
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
##              Df SumsOfSqs MeanSqs F.Model      R2 Pr(>F)
## bmicat_bditotalcat  1      0.530 0.52999  2.1926 0.01123  0.003 **
## Residuals          193      46.653 0.24172           0.98877
## Total              194      47.183           1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $OL_vs_HL
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
##              Df SumsOfSqs MeanSqs F.Model      R2 Pr(>F)
## bmicat_bditotalcat  1      0.4766 0.47663   1.91 0.01634  0.004 **
## Residuals          115      28.6970 0.24954           0.98366
## Total              116      29.1736           1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $OL_vs_HH
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
##              Df SumsOfSqs MeanSqs F.Model      R2 Pr(>F)
## bmicat_bditotalcat  1      0.4966 0.49655  2.0517 0.02136  0.003 **
## Residuals           94      22.7500 0.24202           0.97864
## Total              95      23.2466           1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $OH_vs_HL
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
##              Df SumsOfSqs MeanSqs F.Model      R2 Pr(>F)
```

```
## bmicat_bditotalcat 1 0.623 0.62324 2.5664 0.01559 0.001 ***
## Residuals 162 39.342 0.24285 0.98441
## Total 163 39.965 1.00000
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $OH_vs_HH
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
## Df SumsOfSqs MeanSqs F.Model R2 Pr(>F)
## bmicat_bditotalcat 1 0.510 0.50958 2.1516 0.01503 0.003 **
## Residuals 141 33.395 0.23684 0.98497
## Total 142 33.904 1.00000
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $HL_vs_HH
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
## Df SumsOfSqs MeanSqs F.Model R2 Pr(>F)
## bmicat_bditotalcat 1 0.5789 0.57892 2.3624 0.03614 0.001 ***
## Residuals 63 15.4389 0.24506 0.96386
## Total 64 16.0179 1.00000
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## attr("class")
## [1] "pwadstrata" "list"
```

Betadispersion

```
permutest(betadisper(wUF.dist, cbind(as(sample_data(physeq), "data.frame")$ bmicat_bditotalcat), bias.a
```

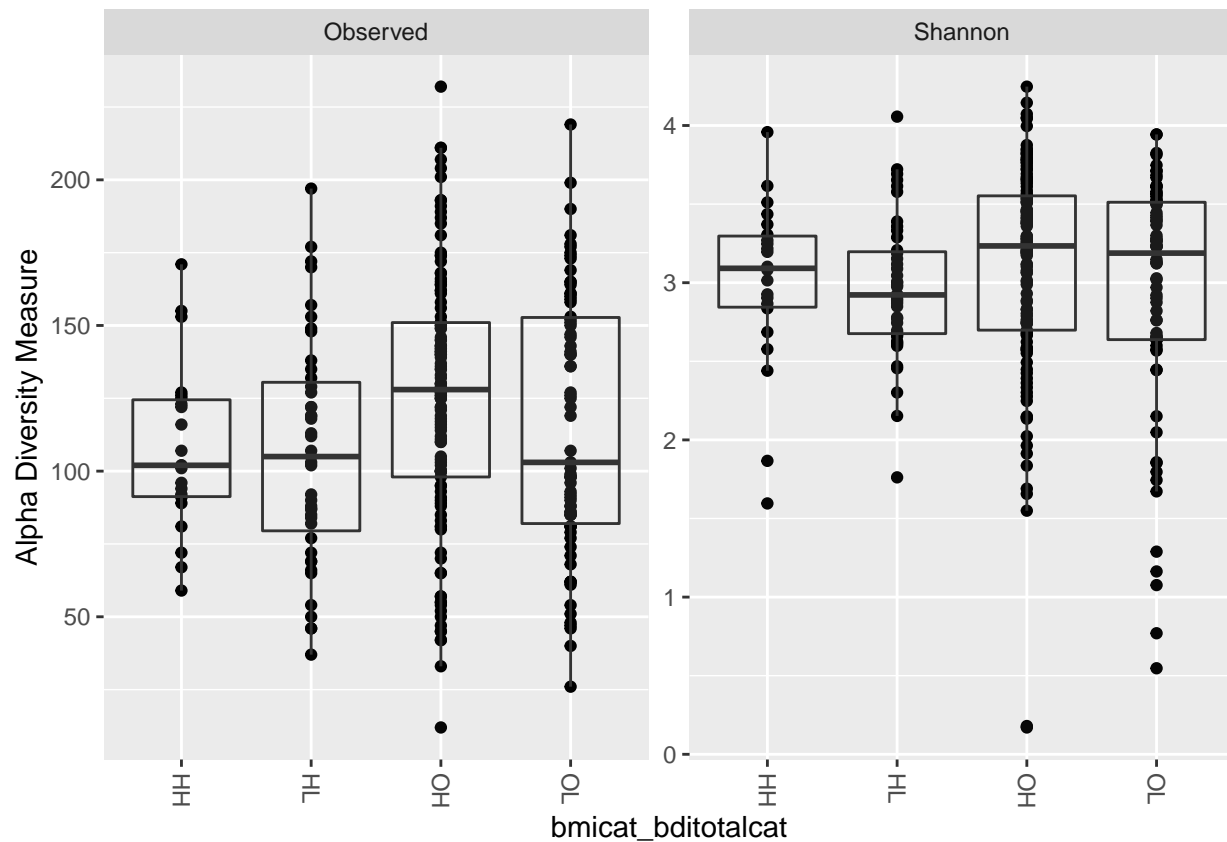
```
##
## Permutation test for homogeneity of multivariate dispersions
## Permutation: free
## Number of permutations: 999
##
## Response: Distances
## Df Sum Sq Mean Sq F N.Perm Pr(>F)
## Groups 3 0.0003644 1.2148e-04 1.4111 999 0.257
## Residuals 256 0.0220393 8.6091e-05
##
## Pairwise comparisons:
## (Observed p-value below diagonal, permuted p-value above diagonal)
## 1 2 3 4
## 1 0.281000 0.096000 0.398
```

```
## 2 0.278579          0.484000 0.692
## 3 0.094905 0.454481          0.180
## 4 0.367189 0.689793 0.169275
```

Alpha diversity analysis

Alpha diversity plots

```
alpha_meas = c("Observed", "Shannon")
p <- plot_richness(physeq, "bmicat_bditotalcat", measures=alpha_meas)
p + geom_boxplot(data=p$physeq, aes(x=bmicat_bditotalcat, y=value, color=NULL), alpha=0.1)
```



Alpha diversity stats - linear mixed effects model to account for repeated measurements

```
div.table<-read.table("~/Documents/projects/Salivary-Microbiome-Metabolome/Gajender-Analyses/Hong-Salivary-Metabolome/DivTable.txt")
summary(lmer(faith_pd ~ sample_collection_time_hongcode+del8+bmi+log(bdi_total3+1) + bmicat_bditotalcat
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula:
## faith_pd ~ sample_collection_time_hongcode + del8 + bmi + log(bdi_total3 +
## 1) + bmicat_bditotalcat + (1 | anonymized_name)
## Data: div.table
##
```

```

## REML criterion at convergence: 1139.3
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.5384 -0.4121  0.0412  0.4508  3.9383
##
## Random effects:
##      Groups             Name             Variance Std.Dev.
## anonymized_name (Intercept) 2.801       1.674
## Residual                   3.271       1.809
## Number of obs: 261, groups: anonymized_name, 60
##
## Fixed effects:
##
##              Estimate Std. Error      df
## (Intercept)      9.993506   2.189309  55.765680
## sample_collection_time_hongcodeCort_1200 -0.103273   0.362545  200.467252
## sample_collection_time_hongcodeCort_1600 -0.302982   0.362358  200.772711
## sample_collection_time_hongcodeCort_1900 -0.533750   0.357643  200.354331
## sample_collection_time_hongcodeCort_Waking -0.066640   0.359098  202.203751
## del18            -0.066751   0.028873   52.010349
## bmi              0.008148   0.057751   56.842042
## log(bdi_total3 + 1) 1.003402   0.519012   56.280209
## bmicat_bditotalcatHL 2.081912   1.359243   53.015794
## bmicat_bditotalcatOH 0.403658   1.091737   54.109264
## bmicat_bditotalcatOL 1.384329   1.299267   53.590415
##
##              t value Pr(>|t|)
## (Intercept)      4.565  2.8e-05 ***
## sample_collection_time_hongcodeCort_1200 -0.285   0.7760
## sample_collection_time_hongcodeCort_1600 -0.836   0.4041
## sample_collection_time_hongcodeCort_1900 -1.492   0.1372
## sample_collection_time_hongcodeCort_Waking -0.186   0.8530
## del18            -2.312   0.0248 *
## bmi              0.141   0.8883
## log(bdi_total3 + 1) 1.933   0.0582 .
## bmicat_bditotalcatHL 1.532   0.1315
## bmicat_bditotalcatOH 0.370   0.7130
## bmicat_bditotalcatOL 1.065   0.2914
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr) s___C_12 s___C_16 s___C_19 s___C_W del18  bmi  l(_3+1
## sm___C_1200 -0.061
## sm___C_1600 -0.071  0.499
## sm___C_1900 -0.061  0.505  0.506
## smpl_c___C_W -0.074  0.508  0.505  0.517
## del18        -0.585  0.000 -0.005 -0.015 -0.008
## bmi          -0.728 -0.012 -0.009 -0.018  0.006  0.366
## lg(bd_t3+1) -0.421 -0.034 -0.010 -0.029 -0.039 -0.014 -0.016
## bmct_bdtHL -0.432 -0.022 -0.010 -0.007 -0.020 -0.145 -0.052  0.629
## bmct_bdtOH  0.128  0.009  0.005  0.023  0.015 -0.198 -0.519 -0.107
## bmct_bdtOL -0.279 -0.015  0.000  0.000 -0.016 -0.097 -0.328  0.590
##
##      bmc_HL bmc_OH
## sm___C_1200

```

```
## sm___C_1600
## sm___C_1900
## smpl_c___C_W
## del8
## bmi
## lg(bd_t3+1)
## bmct_bdtHL
## bmct_bdtOH 0.443
## bmct_bdtOL 0.790 0.613
```

Differential abundance analysis

linear mixed effects model on log-ratios of bacterial features (gram negative/gram positive) and BARIC inflammatory scores - generated from differential analysis

```
sample_plot_data_new <- read_excel("~/Documents/projects/Salivary-Microbiome-Metabolome/Gajender-Analysis")
model<-lmer(gram_Current_Natural_Log_Ratio ~ del8+bmicat_bditotalcat + (1|anonymized_name), data = sample_plot_data_new)
summary(model)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: gram_Current_Natural_Log_Ratio ~ del8 + bmicat_bditotalcat +
##          (1 | anonymized_name)
##      Data: sample_plot_data_new
##
## REML criterion at convergence: 821
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.5104 -0.5309 -0.0504  0.5373  3.1303
##
## Random effects:
##   Groups                Name         Variance Std.Dev.
##   anonymized_name (Intercept) 2.530     1.59
##   Residual                  3.027     1.74
## Number of obs: 190, groups:  anonymized_name, 59
##
## Fixed effects:
##              Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)    1.732615   1.109289  50.930412   1.562   0.1245
## del8            0.005872   0.026846  51.085127   0.219   0.8277
## bmicat_bditotalcat2_OH 1.668612   0.726347  52.425858   2.297   0.0256 *
## bmicat_bditotalcat3_HH 0.745654   1.052546  53.373511   0.708   0.4818
## bmicat_bditotalcat4_OL 1.150424   0.773609  51.380126   1.487   0.1431
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) del8    b_2_OH b_3_HH
## del8          -0.845
## bmct_bd2_OH -0.638  0.240
## bmct_bd3_HH -0.443  0.169  0.500
```

```
## bmct_bd4_OL -0.652  0.287  0.693  0.480
```

boxplot - log-ratios of bacterial features (gram negative/gram positive)

```
box <- ggplot(sample_plot_data_new, aes(x=bmicat_bditotalcat, y=gram_Current_Natural_Log_Ratio, fill=bmicat_bditotalcat)) +  
  geom_boxplot(alpha = 0.4) + geom_jitter(width = 0.1, alpha = 0.4) + scale_fill_manual(values=c("#008c38", "#0000ff", "#d98d38", "#d9534f"))
```

```
## Warning: Removed 70 rows containing non-finite values (stat_boxplot).
```

```
## Warning: Removed 70 rows containing missing values (geom_point).
```

