

Plasmid Genes Occurrence Frequency Diversity Analysis

Ilhan Cem Duru

2024-08-02

Load and prepare data for phyloseq

Create counts matrix

```
# Load counts
cog_function_occurence <- read.csv("cog_function_pivot.tsv", header = TRUE,
                                   skipNul = TRUE, sep = "\t", as.is = TRUE)

#cog_function_occurence
#str(cog_function_occurence)
#row.names(cog_function_occurence)
#class(cog_function_occurence)
colnames(cog_function_occurence)
```

##	[1]	"COG_Function"	"C1"	"C102"	"C103"	"C104"
##	[6]	"C105"	"C107"	"C111"	"C114"	"C116"
##	[11]	"C118"	"C119"	"C123"	"C124"	"C134"
##	[16]	"C135"	"C136"	"C137"	"C140"	"C142"
##	[21]	"C146"	"C147"	"C148"	"C15"	"C152IIP"
##	[26]	"C18"	"C19"	"C20"	"C21"	"C23"
##	[31]	"C24"	"C26"	"C28"	"C30"	"C32"
##	[36]	"C33"	"C34old"	"C35"	"C40"	"C44"
##	[41]	"C46"	"C47"	"C48"	"C49"	"C5"
##	[46]	"C51"	"C54"	"C59"	"C65"	"C68"
##	[51]	"C69"	"C7"	"C70"	"C72"	"C74"
##	[56]	"C75"	"C76II"	"C80"	"C82"	"C85"
##	[61]	"C86"	"C87"	"C88"	"C89"	"C9"
##	[66]	"C90"	"C95"	"C96"	"C98"	"P100"
##	[71]	"P103"	"P104IIP"	"P105"	"P107"	"P10old"
##	[76]	"P114"	"P115"	"P116"	"P118"	"P119"
##	[81]	"P11old"	"P12"	"P120"	"P14"	"P15"
##	[86]	"P16"	"P17"	"P18"	"P19"	"P20"
##	[91]	"P24"	"P26"	"P28"	"P31"	"P34"
##	[96]	"P37"	"P38"	"P4"	"P42"	"P43"
##	[101]	"P45"	"P46"	"P47"	"P48"	"P5"
##	[106]	"P50"	"P51"	"P52"	"P53"	"P56"
##	[111]	"P57"	"P58"	"P59"	"P60"	"P61"
##	[116]	"P62P"	"P63"	"P66"	"P67"	"P68"
##	[121]	"P69"	"P70"	"P71"	"P72"	"P73"
##	[126]	"P74"	"P77"	"P79"	"P8"	"P83"
##	[131]	"P85"	"P87"	"P88"	"P9"	"P94"
##	[136]	"P95II"	"P99"			

```

#head(colnames(cog_function_occurence))
#dim(cog_function_occurence)
#cog_function_occurence$phage

## Put this in `matrix` class for `phyloseq`'s `otu table`, and create it:
counts <- as.matrix(sapply(cog_function_occurence, as.numeric))
#head(counts)
#head(colnames(counts))
rownames(counts) <- cog_function_occurence$COG_Function
#head(counts)
#head(colnames(counts))

# delete/remove the newly created "phage" column:
counts <- counts[, -grep("COG_Function", colnames(counts))]
#head(counts)
#head(colnames(counts))
#dim(counts)

otu.table <- otu_table(counts, taxa_are_rows = TRUE)
head(otu.table)

```

```

## OTU Table:          [6 taxa and 136 samples]
##                    taxa are rows
##
##                    C1 C102 C103 C104 C105 C107
## (p)ppGpp synthase/hydrolase, HD superfamily      0    0    0    2    0    2
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit  0    0    0    0    0    0
## 1,4-Dihydroxy-2-naphthoyl-CoA synthase            1    1    2    0    0    0
## 1,4-alpha-glucan branching enzyme                 0    0    0    0    0    1
## 1,6-Anhydro-N-acetylmuramate kinase               0    0    0    1    0    0
## 1-acyl-sn-glycerol-3-phosphate acyltransferase     0    0    0    1    0    0
##
##                    C111 C114 C116 C118 C119 C123
## (p)ppGpp synthase/hydrolase, HD superfamily      0    4    0    0    0    0
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit  0    2    0    0    0    0
## 1,4-Dihydroxy-2-naphthoyl-CoA synthase            0    0    1    0    0    0
## 1,4-alpha-glucan branching enzyme                 0    0    0    0    0    0
## 1,6-Anhydro-N-acetylmuramate kinase               0    0    0    0    0    0
## 1-acyl-sn-glycerol-3-phosphate acyltransferase     0    1    0    1    0    0
##
##                    C124 C134 C135 C136 C137 C140
## (p)ppGpp synthase/hydrolase, HD superfamily      0    0    0    0    0    0
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit  0    0    0    0    0    0
## 1,4-Dihydroxy-2-naphthoyl-CoA synthase            1    1    1    2    0    0
## 1,4-alpha-glucan branching enzyme                 0    0    0    0    0    0
## 1,6-Anhydro-N-acetylmuramate kinase               0    0    0    0    0    0
## 1-acyl-sn-glycerol-3-phosphate acyltransferase     0    0    0    0    0    0
##
##                    C142 C146 C147 C148 C15
## (p)ppGpp synthase/hydrolase, HD superfamily      0    0    0    0    0
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit  0    0    0    0    0
## 1,4-Dihydroxy-2-naphthoyl-CoA synthase            0    0    0    0    0
## 1,4-alpha-glucan branching enzyme                 0    0    0    0    0
## 1,6-Anhydro-N-acetylmuramate kinase               0    0    0    0    0
## 1-acyl-sn-glycerol-3-phosphate acyltransferase     1    0    0    0    0
##
##                    C152IIP C18 C19 C20 C21 C23
## (p)ppGpp synthase/hydrolase, HD superfamily      2    0    2    0    0    0
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit  0    0    0    0    0    0

```

## 1,4-Dihydroxy-2-naphthoyl-CoA synthase	2	2	2	2	0	0
## 1,4-alpha-glucan branching enzyme	0	0	0	0	0	0
## 1,6-Anhydro-N-acetylmuramate kinase	0	0	0	0	0	0
## 1-acyl-sn-glycerol-3-phosphate acyltransferase	0	0	0	0	0	1
##	C24	C26	C28	C30	C32	C33
## (p)ppGpp synthase/hydrolase, HD superfamily	0	0	0	0	0	0
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit	0	0	0	0	0	0
## 1,4-Dihydroxy-2-naphthoyl-CoA synthase	0	0	0	0	2	0
## 1,4-alpha-glucan branching enzyme	0	0	0	0	0	0
## 1,6-Anhydro-N-acetylmuramate kinase	0	0	0	0	0	0
## 1-acyl-sn-glycerol-3-phosphate acyltransferase	0	0	0	0	0	0
##	C34old	C35	C40	C44	C46	C47
## (p)ppGpp synthase/hydrolase, HD superfamily	0	0	0	0	0	0
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit	0	0	0	0	0	0
## 1,4-Dihydroxy-2-naphthoyl-CoA synthase	1	3	0	0	1	0
## 1,4-alpha-glucan branching enzyme	0	0	0	0	0	0
## 1,6-Anhydro-N-acetylmuramate kinase	0	0	0	0	0	0
## 1-acyl-sn-glycerol-3-phosphate acyltransferase	0	0	0	0	0	1
##	C48	C49	C5	C51	C54	C59
## (p)ppGpp synthase/hydrolase, HD superfamily	0	0	0	2	2	0
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit	0	0	0	0	0	0
## 1,4-Dihydroxy-2-naphthoyl-CoA synthase	1	0	1	0	0	1
## 1,4-alpha-glucan branching enzyme	0	0	0	0	0	0
## 1,6-Anhydro-N-acetylmuramate kinase	0	0	0	0	0	0
## 1-acyl-sn-glycerol-3-phosphate acyltransferase	0	0	0	0	0	0
##	C68	C69	C7	C70	C72	C74
## (p)ppGpp synthase/hydrolase, HD superfamily	0	0	4	0	0	0
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit	0	0	0	0	0	0
## 1,4-Dihydroxy-2-naphthoyl-CoA synthase	1	0	1	2	0	0
## 1,4-alpha-glucan branching enzyme	0	0	0	0	0	0
## 1,6-Anhydro-N-acetylmuramate kinase	0	0	0	0	0	0
## 1-acyl-sn-glycerol-3-phosphate acyltransferase	1	1	0	0	0	0
##	C76II	C80	C82	C85	C86	C87
## (p)ppGpp synthase/hydrolase, HD superfamily	2	0	0	0	2	0
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit	0	0	0	0	0	0
## 1,4-Dihydroxy-2-naphthoyl-CoA synthase	0	0	0	0	0	0
## 1,4-alpha-glucan branching enzyme	0	0	0	0	0	0
## 1,6-Anhydro-N-acetylmuramate kinase	0	0	0	0	0	0
## 1-acyl-sn-glycerol-3-phosphate acyltransferase	0	1	0	0	0	0
##	C89	C9	C90	C95	C96	C98
## (p)ppGpp synthase/hydrolase, HD superfamily	0	0	0	0	0	0
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit	0	0	0	0	0	0
## 1,4-Dihydroxy-2-naphthoyl-CoA synthase	1	0	0	0	0	1
## 1,4-alpha-glucan branching enzyme	0	0	0	0	0	0
## 1,6-Anhydro-N-acetylmuramate kinase	0	0	0	0	0	0
## 1-acyl-sn-glycerol-3-phosphate acyltransferase	0	0	0	0	0	1
##	P103	P104IIP	P105	P107	P10old	
## (p)ppGpp synthase/hydrolase, HD superfamily	0	0	0	0	0	0
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit	0	0	0	0	0	0
## 1,4-Dihydroxy-2-naphthoyl-CoA synthase	0	1	0	1	1	1
## 1,4-alpha-glucan branching enzyme	0	0	0	0	0	0
## 1,6-Anhydro-N-acetylmuramate kinase	0	0	0	0	0	0
## 1-acyl-sn-glycerol-3-phosphate acyltransferase	0	0	0	0	0	0
##	P114	P115	P116	P118	P119	

## (p)ppGpp synthase/hydrolase, HD superfamily	0	2	2	0	0	
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit	0	0	0	0	0	
## 1,4-Dihydroxy-2-naphthoyl-CoA synthase	0	0	0	0	0	
## 1,4-alpha-glucan branching enzyme	0	0	0	0	0	
## 1,6-Anhydro-N-acetylmuramate kinase	0	0	0	0	0	
## 1-acyl-sn-glycerol-3-phosphate acyltransferase	0	0	0	0	0	
##	P11old	P12	P120	P14	P15	P16
## (p)ppGpp synthase/hydrolase, HD superfamily	0	0	0	0	0	2
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit	0	0	0	0	0	0
## 1,4-Dihydroxy-2-naphthoyl-CoA synthase	0	0	1	0	0	0
## 1,4-alpha-glucan branching enzyme	0	0	0	0	0	0
## 1,6-Anhydro-N-acetylmuramate kinase	0	0	0	0	0	0
## 1-acyl-sn-glycerol-3-phosphate acyltransferase	0	0	0	0	0	0
##	P17	P18	P19	P20	P24	P26
## (p)ppGpp synthase/hydrolase, HD superfamily	0	0	0	0	0	0
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit	0	0	0	0	0	0
## 1,4-Dihydroxy-2-naphthoyl-CoA synthase	0	1	0	0	0	1
## 1,4-alpha-glucan branching enzyme	0	0	0	0	0	0
## 1,6-Anhydro-N-acetylmuramate kinase	0	0	0	0	0	0
## 1-acyl-sn-glycerol-3-phosphate acyltransferase	0	1	0	0	0	0
##	P31	P34	P37	P38	P4	P42
## (p)ppGpp synthase/hydrolase, HD superfamily	0	0	0	0	0	0
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit	0	0	0	0	0	0
## 1,4-Dihydroxy-2-naphthoyl-CoA synthase	0	0	1	0	0	2
## 1,4-alpha-glucan branching enzyme	0	0	0	0	0	0
## 1,6-Anhydro-N-acetylmuramate kinase	0	0	0	0	0	0
## 1-acyl-sn-glycerol-3-phosphate acyltransferase	0	0	0	0	1	0
##	P45	P46	P47	P48	P5	P50
## (p)ppGpp synthase/hydrolase, HD superfamily	0	0	2	0	0	0
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit	0	0	0	0	0	0
## 1,4-Dihydroxy-2-naphthoyl-CoA synthase	0	0	0	0	0	0
## 1,4-alpha-glucan branching enzyme	0	0	0	0	0	0
## 1,6-Anhydro-N-acetylmuramate kinase	0	0	0	0	0	0
## 1-acyl-sn-glycerol-3-phosphate acyltransferase	0	0	0	0	0	0
##	P52	P53	P56	P57	P58	P59
## (p)ppGpp synthase/hydrolase, HD superfamily	2	0	2	0	0	2
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit	0	0	0	0	0	0
## 1,4-Dihydroxy-2-naphthoyl-CoA synthase	1	0	1	0	0	1
## 1,4-alpha-glucan branching enzyme	0	0	0	0	0	0
## 1,6-Anhydro-N-acetylmuramate kinase	0	0	0	0	0	0
## 1-acyl-sn-glycerol-3-phosphate acyltransferase	0	0	0	1	0	0
##	P61	P62P	P63	P66	P67	P68
## (p)ppGpp synthase/hydrolase, HD superfamily	0	0	0	0	0	0
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit	0	0	0	0	0	0
## 1,4-Dihydroxy-2-naphthoyl-CoA synthase	0	1	1	0	0	1
## 1,4-alpha-glucan branching enzyme	0	0	0	0	0	0
## 1,6-Anhydro-N-acetylmuramate kinase	0	0	0	0	0	0
## 1-acyl-sn-glycerol-3-phosphate acyltransferase	0	0	0	0	0	0
##	P70	P71	P72	P73	P74	P77
## (p)ppGpp synthase/hydrolase, HD superfamily	0	0	0	0	0	2
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit	0	0	0	0	0	1
## 1,4-Dihydroxy-2-naphthoyl-CoA synthase	1	0	0	1	0	1
## 1,4-alpha-glucan branching enzyme	0	0	0	0	0	0
## 1,6-Anhydro-N-acetylmuramate kinase	0	0	0	0	0	0

```
## 1-acyl-sn-glycerol-3-phosphate acyltransferase      0  1  0  1  0  0  0
##                                                    P8 P83 P85 P87 P88 P9 P94
## (p)ppGpp synthase/hydrolase, HD superfamily      0  0  0  0  0  0  0
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit  0  0  0  0  0  0  0
## 1,4-Dihydroxy-2-naphthoyl-CoA synthase            2  1  0  1  2  0  2
## 1,4-alpha-glucan branching enzyme                 0  0  0  0  0  0  0
## 1,6-Anhydro-N-acetylmuramate kinase               0  0  0  0  0  0  0
## 1-acyl-sn-glycerol-3-phosphate acyltransferase     0  0  0  0  0  0  0
##                                                    P95II P99
## (p)ppGpp synthase/hydrolase, HD superfamily      0  0
## 1,2-phenylacetyl-CoA epoxidase, catalytic subunit  0  0
## 1,4-Dihydroxy-2-naphthoyl-CoA synthase            1  0
## 1,4-alpha-glucan branching enzyme                 0  0
## 1,6-Anhydro-N-acetylmuramate kinase               0  0
## 1-acyl-sn-glycerol-3-phosphate acyltransferase     0  0
```

```
head(colnames(otu.table))
```

```
## [1] "C1" "C102" "C103" "C104" "C105" "C107"
```

```
dim(otu.table)
```

```
## [1] 1638 136
```

```
#otu.table
```

Create dummy tax table for TAX

```
taxmat = matrix(sample(letters, 1638, replace = TRUE),
                 nrow = nrow(otu.table), ncol = 7)
rownames(taxmat) <- rownames(otu.table)
colnames(taxmat) <- c("Domain", "Phylum", "Class", "Order",
                     "Family", "Genus", "Species")
#taxmat
TAX = tax_table(taxmat)
```

Import the metadata:

```
sampladata <- as.data.frame(read.csv
                           (file = "pd_meta_with_ffq_and_scfa_only_oursamples_3variables_ordered.csv",
                             header = TRUE, sep = ",", row.names = 1))
#rownames(sampladata)
#colnames(otu.table)
identical(rownames(sampladata), colnames(otu.table))

## [1] TRUE

sampladata = sample_data(sampladata)
```

Create physeq

```
physeqfinal <- phyloseq(otu.table, TAX, sampladata)
#physeqfinal
#sample_data(physeqfinal)
#summary(sample_data(physeqfinal))
```

```

physeqfinal.2 <- subset_taxa(physeqfinal, taxa_sums(physeqfinal) > 0)
#head(sort(taxa_sums(physeqfinal.2), decreasing = FALSE))

physeqfinal.2

## phyloseq-class experiment-level object
## otu_table() OTU Table: [ 1638 taxa and 136 samples ]
## sample_data() Sample Data: [ 136 samples by 4 sample variables ]
## tax_table() Taxonomy Table: [ 1638 taxa by 7 taxonomic ranks ]

physeqfinal

## phyloseq-class experiment-level object
## otu_table() OTU Table: [ 1638 taxa and 136 samples ]
## sample_data() Sample Data: [ 136 samples by 4 sample variables ]
## tax_table() Taxonomy Table: [ 1638 taxa by 7 taxonomic ranks ]

summary(sample_data(physeqfinal.2))

##      Group      gender      age_at_stool_collection      BMI
## Length:136      Length:136      Min. :51.00      Min. :17.51
## Class :character Class :character 1st Qu.:61.00      1st Qu.:24.13
## Mode :character  Mode :character Median :65.00      Median :26.31
##                                     Mean :64.98      Mean :26.69
##                                     3rd Qu.:69.00      3rd Qu.:28.62
##                                     Max. :78.00      Max. :37.87
##                                     NA's :13

```

Alpha Diversity for Group (C vs P)

```

richness.table <- estimate_richness(physeqfinal.2, split = TRUE,
                                   measures = c("Observed", "Chao1",
                                                "ACE", "Shannon", "InvSimpson"))

head(richness.table)

##      Observed   Chao1 se.chao1      ACE   se.ACE  Shannon InvSimpson
## C1          214 360.2708 36.16543 439.1438 14.04301 4.352056 30.16385
## C102         162 258.1842 27.72289 308.5497 11.55420 4.519859 49.59434
## C103         123 215.8846 30.19955 264.2992 11.60073 4.086501 27.34375
## C104         260 464.3061 47.14729 511.9811 14.34224 4.598662 37.43706
## C105         127 209.4516 26.03126 247.1548  9.51239 4.015784 23.79298
## C107         186 351.0000 43.25324 399.4231 12.92949 4.383748 32.12203

richness.table$Group <- sample_data(physeqfinal.2)$Group

```

Plot Alpha Diversity

```

theme_set(theme_bw())

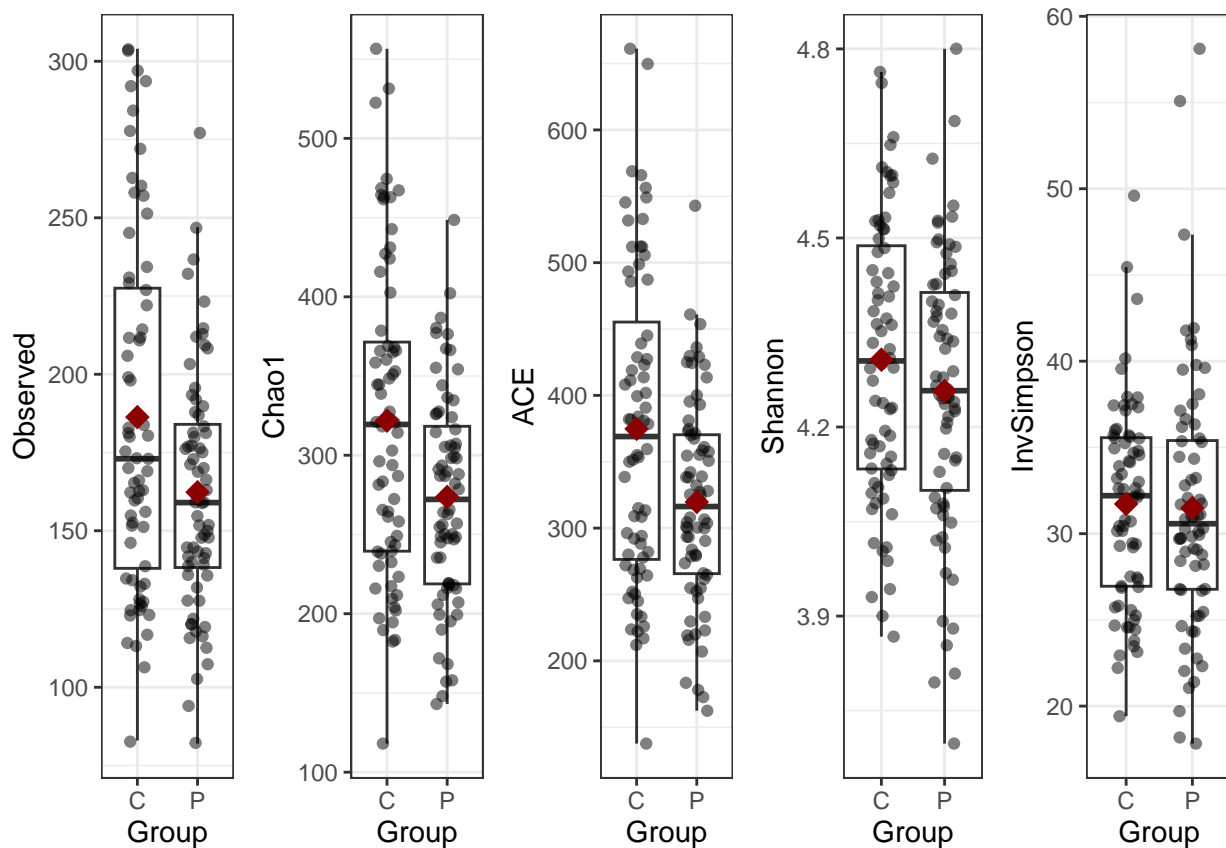
grid.arrange(
  ggplot(richness.table, aes(x = Group, y = Observed)) +
    geom_boxplot(outlier.shape = NA) + geom_jitter(width = 0.2, alpha = .5) +
    stat_summary(fun = mean, color = "darkred", geom = "point", shape = 18, size = 4),

```

```

ggplot(richness.table, aes(x = Group, y = Chao1)) +
  geom_boxplot(outlier.shape = NA) + geom_jitter(width = 0.2, alpha = .5) +
  stat_summary(fun = mean, color = "darkred", geom = "point", shape = 18, size = 4),
ggplot(richness.table, aes(x = Group, y = ACE)) +
  geom_boxplot(outlier.shape = NA) + geom_jitter(width = 0.2, alpha = .5) +
  stat_summary(fun = mean, color = "darkred", geom = "point", shape = 18, size = 4),
ggplot(richness.table, aes(x = Group, y = Shannon)) +
  geom_boxplot(outlier.shape = NA) + geom_jitter(width = 0.2, alpha = .5) +
  stat_summary(fun = mean, color = "darkred", geom = "point", shape = 18, size = 4),
ggplot(richness.table, aes(x = Group, y = InvSimpson)) +
  geom_boxplot(outlier.shape = NA) + geom_jitter(width = 0.2, alpha = .5) +
  stat_summary(fun = mean, color = "darkred", geom = "point", shape = 18, size = 4),
nrow = 1)

```



Calculate mean and standard deviation for each alpha diversity metric

```

# Function to calculate mean and SD, ignoring NaN values
calc_mean_sd <- function(x) {
  x_clean <- x[!is.nan(x)]
  if (length(x_clean) > 0) {
    c(mean = mean(x_clean, na.rm = TRUE),
      sd = sd(x_clean, na.rm = TRUE))
  } else {
    c(mean = NA, sd = NA)
  }
}

```

```

# Calculate mean and standard deviation for each index by group, ignoring NaN values
richness_stats <- richness.table %>%
  group_by(Group) %>%
  summarise(
    across(c(Observed, Chao1, ACE, Shannon, InvSimpson),
      list(Mean = ~calc_mean_sd(.)[1],
           SD = ~calc_mean_sd(.)[2]))
  )

```

```

## group_by: one grouping variable (Group)
## summarise: now 2 rows and 11 columns, ungrouped

```

```

# Print the results
print(richness_stats)

```

```

## # A tibble: 2 x 11
##   Group Observed_Mean Observed_SD Chao1_Mean Chao1_SD ACE_Mean ACE_SD
##   <chr>      <dbl>      <dbl>      <dbl>    <dbl>  <dbl>  <dbl>
## 1 C          186.        56.9        322.     99.2   375.   118.
## 2 P          162.        38.4        274.     67.3   320.    77.8
## # i 4 more variables: Shannon_Mean <dbl>, Shannon_SD <dbl>,
## #   InvSimpson_Mean <dbl>, InvSimpson_SD <dbl>
# write the results to a CSV file
write.csv(richness_stats, "group_richness_statistics.csv", row.names = FALSE)

```

Observed Richness Wilcoxon rank sum test (Group)

```

wilcox.test(richness.table$Observed ~ sample_data(physeqfinal.2)$Group,
  conf.level = 0.95, conf.int = TRUE)

```

```

##
## Wilcoxon rank sum test with continuity correction
##
## data: richness.table$Observed by sample_data(physeqfinal.2)$Group
## W = 2814, p-value = 0.02905
## alternative hypothesis: true location shift is not equal to 0
## 95 percent confidence interval:
##  2.999956 35.999918
## sample estimates:
## difference in location
##          18.00002

```

CHAO1 Wilcoxon rank sum test (Group)

```

wilcox.test(richness.table$Chao1 ~ sample_data(physeqfinal.2)$Group,
  conf.level = 0.95, conf.int = TRUE)

```

```

##
## Wilcoxon rank sum test with continuity correction
##
## data: richness.table$Chao1 by sample_data(physeqfinal.2)$Group
## W = 2916, p-value = 0.008623
## alternative hypothesis: true location shift is not equal to 0

```



```
## 95 percent confidence interval:
## 12.27292 72.96276
## sample estimates:
## difference in location
## 42.03116
```

ACE Wilcoxon rank sum test (Group)

```
wilcox.test(richness.table$ACE ~ sample_data(physeqfinal.2)$Group,
             conf.level = 0.95, conf.int = TRUE)
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: richness.table$ACE by sample_data(physeqfinal.2)$Group
## W = 2891, p-value = 0.01181
## alternative hypothesis: true location shift is not equal to 0
## 95 percent confidence interval:
## 10.10063 84.99837
## sample estimates:
## difference in location
## 48.68983
```

InvSimpson Wilcoxon rank sum test (Group)

```
wilcox.test(richness.table$InvSimpson ~ sample_data(physeqfinal.2)$Group,
             conf.level = 0.95, conf.int = TRUE)
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: richness.table$InvSimpson by sample_data(physeqfinal.2)$Group
## W = 2478, p-value = 0.4713
## alternative hypothesis: true location shift is not equal to 0
## 95 percent confidence interval:
## -1.555451 2.909300
## sample estimates:
## difference in location
## 0.7052922
```

Shannon Wilcoxon rank sum test (Group)

```
wilcox.test(richness.table$Shannon ~ sample_data(physeqfinal.2)$Group,
             conf.level = 0.95, conf.int = TRUE)
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: richness.table$Shannon by sample_data(physeqfinal.2)$Group
## W = 2566, p-value = 0.2699
## alternative hypothesis: true location shift is not equal to 0
## 95 percent confidence interval:
## -0.03705563 0.12344110
## sample estimates:
```

```
## difference in location
##          0.04568665
```

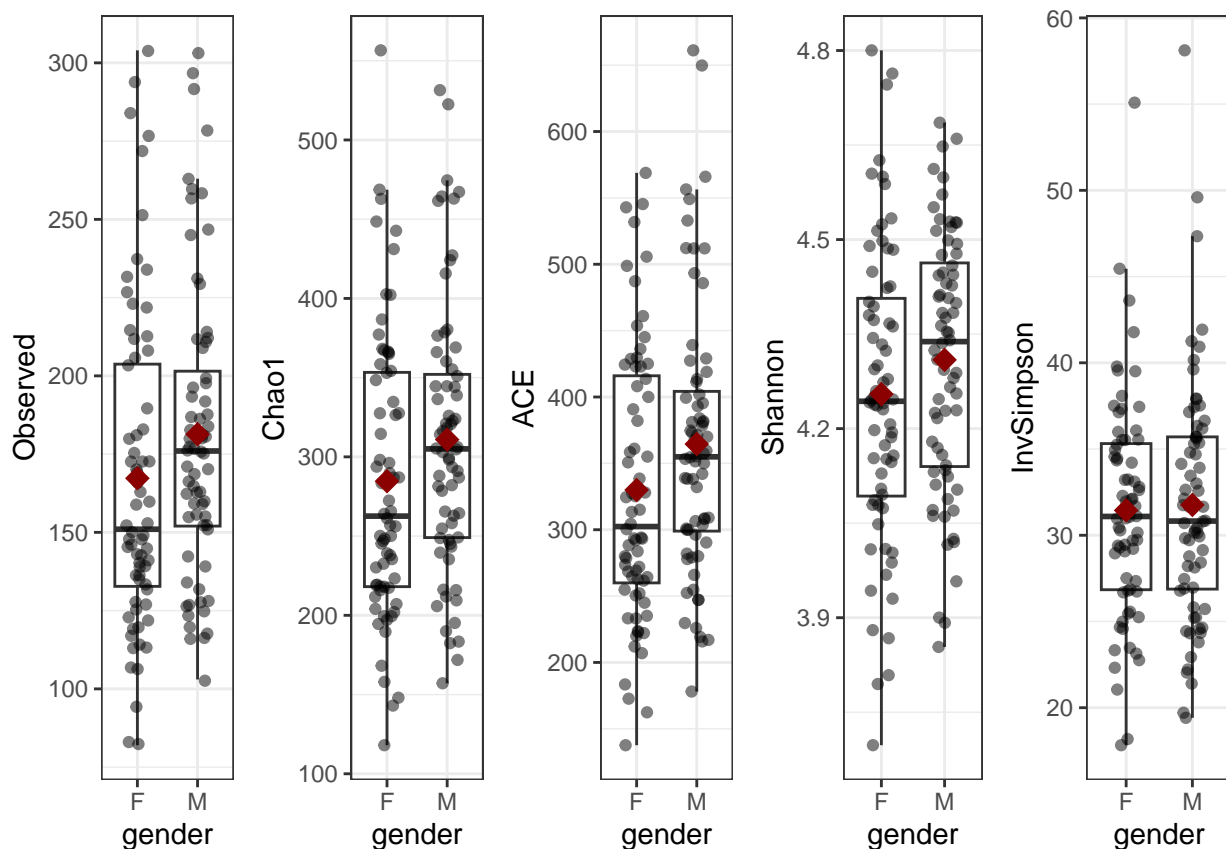
Alpha Diversity for Gender (F vs M)

```
richness.table$gender <- sample_data(physeqfinal.2)$gender
```

Plot Alpha Diversity for Gender (F vs M)

```
theme_set(theme_bw())

grid.arrange(
  ggplot(richness.table, aes(x = gender, y = Observed)) +
    geom_boxplot(outlier.shape = NA) + geom_jitter(width = 0.2, alpha = .5) +
    stat_summary(fun = mean, color = "darkred", geom = "point", shape = 18, size = 4),
  ggplot(richness.table, aes(x = gender, y = Chao1)) +
    geom_boxplot(outlier.shape = NA) + geom_jitter(width = 0.2, alpha = .5) +
    stat_summary(fun = mean, color = "darkred", geom = "point", shape = 18, size = 4),
  ggplot(richness.table, aes(x = gender, y = ACE)) +
    geom_boxplot(outlier.shape = NA) + geom_jitter(width = 0.2, alpha = .5) +
    stat_summary(fun = mean, color = "darkred", geom = "point", shape = 18, size = 4),
  ggplot(richness.table, aes(x = gender, y = Shannon)) +
    geom_boxplot(outlier.shape = NA) + geom_jitter(width = 0.2, alpha = .5) +
    stat_summary(fun = mean, color = "darkred", geom = "point", shape = 18, size = 4),
  ggplot(richness.table, aes(x = gender, y = InvSimpson)) +
    geom_boxplot(outlier.shape = NA) + geom_jitter(width = 0.2, alpha = .5) +
    stat_summary(fun = mean, color = "darkred", geom = "point", shape = 18, size = 4),
  nrow = 1)
```



Calculate mean and standard deviation for each alpha diversity metric

```
# Function to calculate mean and SD, ignoring NaN values
calc_mean_sd <- function(x) {
  x_clean <- x[!is.nan(x)]
  if (length(x_clean) > 0) {
    c(mean = mean(x_clean, na.rm = TRUE),
      sd = sd(x_clean, na.rm = TRUE))
  } else {
    c(mean = NA, sd = NA)
  }
}

# Calculate mean and standard deviation for each index by group, ignoring NaN values
richness_stats <- richness.table %>%
  group_by(gender) %>%
  summarise(
    across(c(Observed, Chao1, ACE, Shannon, InvSimpson),
      list(Mean = ~calc_mean_sd(.)[1],
           SD = ~calc_mean_sd(.)[2]))
  )

## group_by: one grouping variable (gender)
## summarise: now 2 rows and 11 columns, ungrouped

# Print the results
print(richness_stats)
```

```
## # A tibble: 2 x 11
##   gender Observed_Mean Observed_SD Chao1_Mean Chao1_SD ACE_Mean ACE_SD
##   <chr>      <dbl>      <dbl>      <dbl>      <dbl>      <dbl> <dbl>
## 1 F          167.        51.6        285.        89.3        330.   102.
## 2 M          181.        47.3        311.        85.1        365.   103.
## # i 4 more variables: Shannon_Mean <dbl>, Shannon_SD <dbl>,
## #   InvSimpson_Mean <dbl>, InvSimpson_SD <dbl>
# write the results to a CSV file
write.csv(richness_stats, "gender_richness_statistics.csv", row.names = FALSE)
```

Observed Richness Wilcoxon rank sum test (gender)

```
wilcox.test(richness.table$Observed ~ sample_data(physeqfinal.2)$gender,
             conf.level = 0.95, conf.int = TRUE)
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: richness.table$Observed by sample_data(physeqfinal.2)$gender
## W = 1840.5, p-value = 0.04036
## alternative hypothesis: true location shift is not equal to 0
## 95 percent confidence interval:
## -31.9999426 -0.9999979
## sample estimates:
## difference in location
## -16.00004
```

CHAO1 Wilcoxon rank sum test (gender)

```
wilcox.test(richness.table$Chao1 ~ sample_data(physeqfinal.2)$gender,
             conf.level = 0.95, conf.int = TRUE)
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: richness.table$Chao1 by sample_data(physeqfinal.2)$gender
## W = 1891, p-value = 0.06723
## alternative hypothesis: true location shift is not equal to 0
## 95 percent confidence interval:
## -58.677468 1.885731
## sample estimates:
## difference in location
## -28.99363
```

ACE Wilcoxon rank sum test (gender)

```
wilcox.test(richness.table$ACE ~ sample_data(physeqfinal.2)$gender,
             conf.level = 0.95, conf.int = TRUE)
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: richness.table$ACE by sample_data(physeqfinal.2)$gender
```

```
## W = 1859, p-value = 0.0489
## alternative hypothesis: true location shift is not equal to 0
## 95 percent confidence interval:
## -70.8363181 -0.3217077
## sample estimates:
## difference in location
## -35.79709
```

InvSimpson Wilcoxon rank sum test (gender)

```
wilcox.test(richness.table$InvSimpson ~ sample_data(physeqfinal.2)$gender,
             conf.level = 0.95, conf.int = TRUE)
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: richness.table$InvSimpson by sample_data(physeqfinal.2)$gender
## W = 2291, p-value = 0.9289
## alternative hypothesis: true location shift is not equal to 0
## 95 percent confidence interval:
## -2.254604 2.085266
## sample estimates:
## difference in location
## -0.1003698
```

Shannon Wilcoxon rank sum test (gender)

```
wilcox.test(richness.table$Shannon ~ sample_data(physeqfinal.2)$gender,
             conf.level = 0.95, conf.int = TRUE)
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: richness.table$Shannon by sample_data(physeqfinal.2)$gender
## W = 1963, p-value = 0.1293
## alternative hypothesis: true location shift is not equal to 0
## 95 percent confidence interval:
## -0.14171554 0.01754889
## sample estimates:
## difference in location
## -0.06065996
```

Alpha Diversity for BMI (Continuous variable)

```
richness.table$BMI <- sample_data(physeqfinal.2)$BMI
```

Plot Alpha Diversity for BMI (Continuous variable)

```
common_theme <- theme(
  axis.text.x = element_text(angle = 45, hjust = 1),
)
```

```

grid.arrange(
  ggplot(richness.table, aes(x = BMI, y = Observed)) +
    geom_point(alpha = 0.5) +
    geom_smooth(method = "lm", color = "darkred") +
    labs(x = "BMI", y = "Observed") +
    scale_x_continuous(breaks = seq(20, 35, by = 5)) + # Adjust range as needed
    common_theme,

  ggplot(richness.table, aes(x = BMI, y = Chao1)) +
    geom_point(alpha = 0.5) +
    geom_smooth(method = "lm", color = "darkred") +
    labs(x = "BMI", y = "Chao1") +
    scale_x_continuous(breaks = seq(20, 35, by = 5)) +
    common_theme,

  ggplot(richness.table, aes(x = BMI, y = ACE)) +
    geom_point(alpha = 0.5) +
    geom_smooth(method = "lm", color = "darkred") +
    labs(x = "BMI", y = "ACE") +
    scale_x_continuous(breaks = seq(20, 35, by = 5)) +
    common_theme,

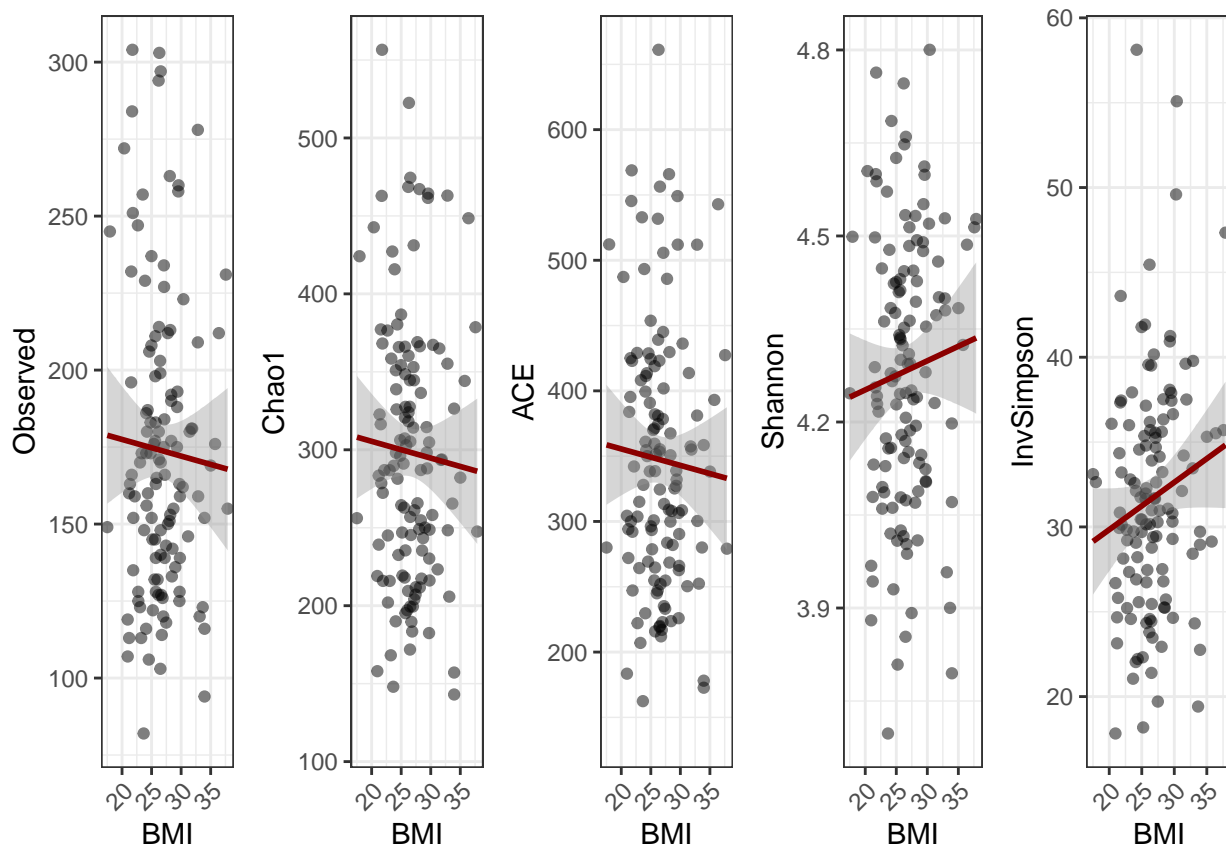
  ggplot(richness.table, aes(x = BMI, y = Shannon)) +
    geom_point(alpha = 0.5) +
    geom_smooth(method = "lm", color = "darkred") +
    labs(x = "BMI", y = "Shannon") +
    scale_x_continuous(breaks = seq(20, 35, by = 5)) +
    common_theme,

  ggplot(richness.table, aes(x = BMI, y = InvSimpson)) +
    geom_point(alpha = 0.5) +
    geom_smooth(method = "lm", color = "darkred") +
    labs(x = "BMI", y = "InvSimpson") +
    scale_x_continuous(breaks = seq(20, 35, by = 5)) +
    common_theme,

  nrow = 1
)

## `geom_smooth()` using formula = 'y ~ x'
## `geom_smooth()` using formula = 'y ~ x'
## `geom_smooth()` using formula = 'y ~ x'
## `geom_smooth()` using formula = 'y ~ x'
## `geom_smooth()` using formula = 'y ~ x'

```



Observed Richness Spearman rank sum test (BMI)

```
cor.test(richness.table$Observed, sample_data(physeqfinal.2)$BMI, method = "spearman")
```

```
##
## Spearman's rank correlation rho
##
## data: richness.table$Observed and sample_data(physeqfinal.2)$BMI
## S = 310899, p-value = 0.9781
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## -0.002499214
```

CHAO1 Richness Spearman rank sum test (BMI)

```
cor.test(richness.table$Chao1, sample_data(physeqfinal.2)$BMI, method = "spearman")
```

```
##
## Spearman's rank correlation rho
##
## data: richness.table$Chao1 and sample_data(physeqfinal.2)$BMI
## S = 329909, p-value = 0.4833
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
```

```
## -0.06379778
```

ACE Richness Spearman rank sum test (BMI)

```
cor.test(richness.table$ACE, sample_data(physeqfinal.2)$BMI, method = "spearman")
```

```
##  
## Spearman's rank correlation rho  
##  
## data: richness.table$ACE and sample_data(physeqfinal.2)$BMI  
## S = 331560, p-value = 0.4474  
## alternative hypothesis: true rho is not equal to 0  
## sample estimates:  
##      rho  
## -0.06912141
```

InvSimpson Richness Spearman rank sum test (BMI)

```
cor.test(richness.table$InvSimpson, sample_data(physeqfinal.2)$BMI, method = "spearman")
```

```
##  
## Spearman's rank correlation rho  
##  
## data: richness.table$InvSimpson and sample_data(physeqfinal.2)$BMI  
## S = 260694, p-value = 0.07825  
## alternative hypothesis: true rho is not equal to 0  
## sample estimates:  
##      rho  
## 0.1593894
```

Shannon Richness Spearman rank sum test (BMI)

```
cor.test(richness.table$Shannon, sample_data(physeqfinal.2)$BMI, method = "spearman")
```

```
##  
## Spearman's rank correlation rho  
##  
## data: richness.table$Shannon and sample_data(physeqfinal.2)$BMI  
## S = 279626, p-value = 0.2792  
## alternative hypothesis: true rho is not equal to 0  
## sample estimates:  
##      rho  
## 0.09834226
```

Alpha Diversity for Age (Continuous variable)

```
richness.table$Age <- sample_data(physeqfinal.2)$age_at_stool_collection
```

Plot Alpha Diversity for Age (Continuous variable)

```
common_theme <- theme(  
  axis.text.x = element_text(angle = 45, hjust = 1),
```



```

)

grid.arrange(
  ggplot(richness.table, aes(x = age, y = Observed)) +
    geom_point(alpha = 0.5) +
    geom_smooth(method = "lm", color = "darkred") +
    labs(x = "age", y = "Observed") +
    scale_x_continuous(breaks = seq(50, 80, by = 10)) + # Adjust range as needed
    common_theme,

  ggplot(richness.table, aes(x = age, y = Chao1)) +
    geom_point(alpha = 0.5) +
    geom_smooth(method = "lm", color = "darkred") +
    labs(x = "age", y = "Chao1") +
    scale_x_continuous(breaks = seq(50, 80, by = 10)) +
    common_theme,

  ggplot(richness.table, aes(x = age, y = ACE)) +
    geom_point(alpha = 0.5) +
    geom_smooth(method = "lm", color = "darkred") +
    labs(x = "age", y = "ACE") +
    scale_x_continuous(breaks = seq(50, 80, by = 10)) +
    common_theme,

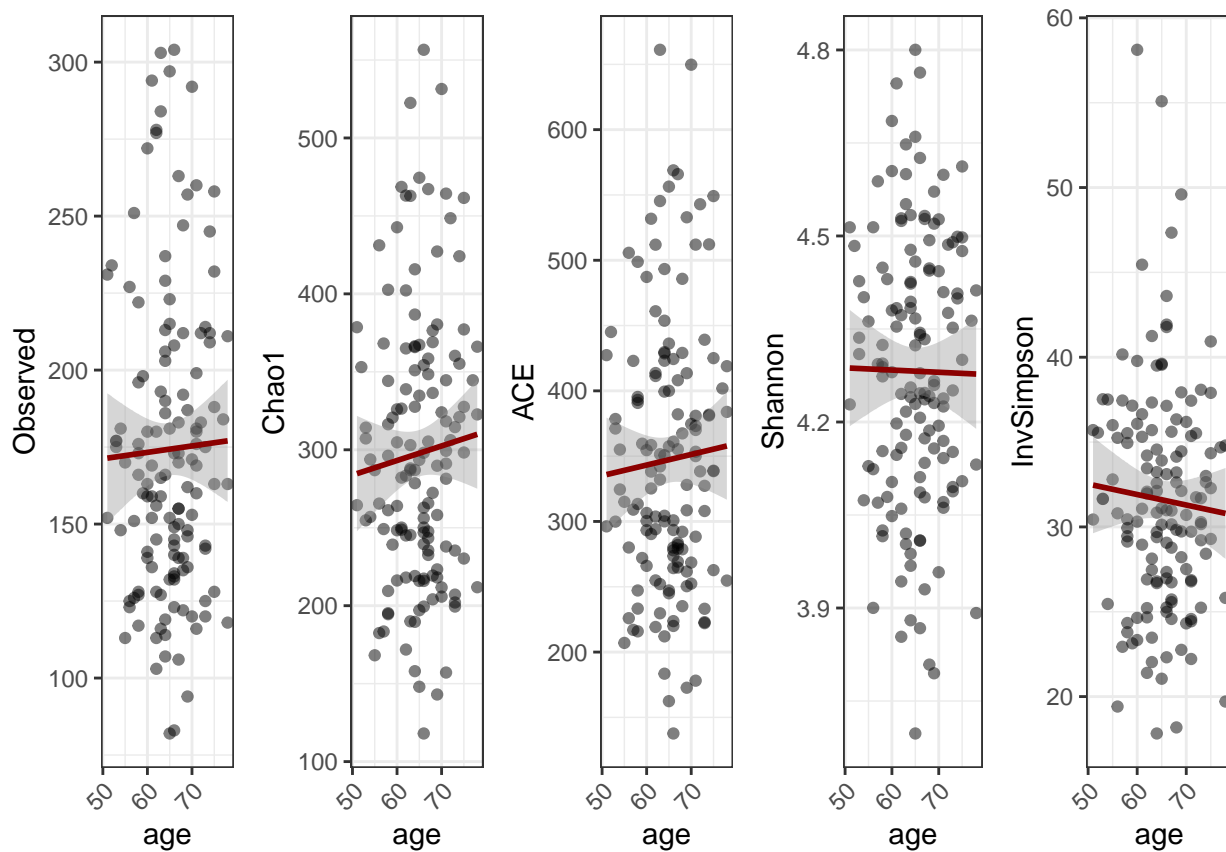
  ggplot(richness.table, aes(x = age, y = Shannon)) +
    geom_point(alpha = 0.5) +
    geom_smooth(method = "lm", color = "darkred") +
    labs(x = "age", y = "Shannon") +
    scale_x_continuous(breaks = seq(50, 80, by = 10)) +
    common_theme,

  ggplot(richness.table, aes(x = age, y = InvSimpson)) +
    geom_point(alpha = 0.5) +
    geom_smooth(method = "lm", color = "darkred") +
    labs(x = "age", y = "InvSimpson") +
    scale_x_continuous(breaks = seq(50, 80, by = 10)) +
    common_theme,

  nrow = 1
)

## `geom_smooth()` using formula = 'y ~ x'
## `geom_smooth()` using formula = 'y ~ x'
## `geom_smooth()` using formula = 'y ~ x'
## `geom_smooth()` using formula = 'y ~ x'
## `geom_smooth()` using formula = 'y ~ x'

```



Observed Richness Spearman rank sum test (age)

```
cor.test(richness.table$Observed,
         sample_data(physeqfinal.2)$age_at_stool_collection, method = "spearman")

##
## Spearman's rank correlation rho
##
## data: richness.table$Observed and sample_data(physeqfinal.2)$age_at_stool_collection
## S = 401491, p-value = 0.625
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## 0.04228955
```

CHAO1 Richness Spearman rank sum test (age)

```
cor.test(richness.table$Chao1,
         sample_data(physeqfinal.2)$age_at_stool_collection, method = "spearman")

##
## Spearman's rank correlation rho
##
## data: richness.table$Chao1 and sample_data(physeqfinal.2)$age_at_stool_collection
## S = 396544, p-value = 0.5317
## alternative hypothesis: true rho is not equal to 0
```

```
## sample estimates:
##      rho
## 0.05409024
```

ACE Richness Spearman rank sum test (age)

```
cor.test(richness.table$ACE,
         sample_data(physeqfinal.2)$age_at_stool_collection, method = "spearman")

##
## Spearman's rank correlation rho
##
## data: richness.table$ACE and sample_data(physeqfinal.2)$age_at_stool_collection
## S = 405517, p-value = 0.7056
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## 0.03268702
```

InvSimpson Richness Spearman rank sum test (age)

```
cor.test(richness.table$InvSimpson,
         sample_data(physeqfinal.2)$age_at_stool_collection, method = "spearman")

##
## Spearman's rank correlation rho
##
## data: richness.table$InvSimpson and sample_data(physeqfinal.2)$age_at_stool_collection
## S = 442380, p-value = 0.5229
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## -0.05524519
```

Shannon Richness Spearman rank sum test (age)

```
cor.test(richness.table$Shannon,
         sample_data(physeqfinal.2)$age_at_stool_collection, method = "spearman")

##
## Spearman's rank correlation rho
##
## data: richness.table$Shannon and sample_data(physeqfinal.2)$age_at_stool_collection
## S = 416514, p-value = 0.9405
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## 0.006454748
```

Distance-based multivariate analysis / beta diversity

CLR transformation followed by Euclidian Distance (a.k.a. Aitchinson Distance)

```
physeqfinal.2.clr <- microbiome::transform(physeqfinal.2, transform = "clr", target = "OTU",
                                          shift = 0, scale = 1)
physeqfinal.2.clr.eucl_dist <- vegdist(t(as.data.frame(as.matrix(otu_table(physeqfinal.2.clr)))),
                                       method = "euclidean", binary = FALSE)

sample_data(physeqfinal.2.clr)$Group <- factor(sample_data(physeqfinal.2.clr)$Group,
                                              levels=c("C" , "P"),
                                              ordered = FALSE)
sample_data(physeqfinal.2.clr)$gender <- factor(sample_data(physeqfinal.2.clr)$gender,
                                              levels=c("F" , "M"),
                                              ordered = FALSE)
```

Run the statistics for group variable

```
set.seed(1337)
adonis.Res.clr <- adonis2(physeqfinal.2.clr.eucl_dist ~
                        sample_data(physeqfinal.2.clr)$Group,
                        perm = 10000,
                        na.action = na.exclude,
                        parallel = 10)

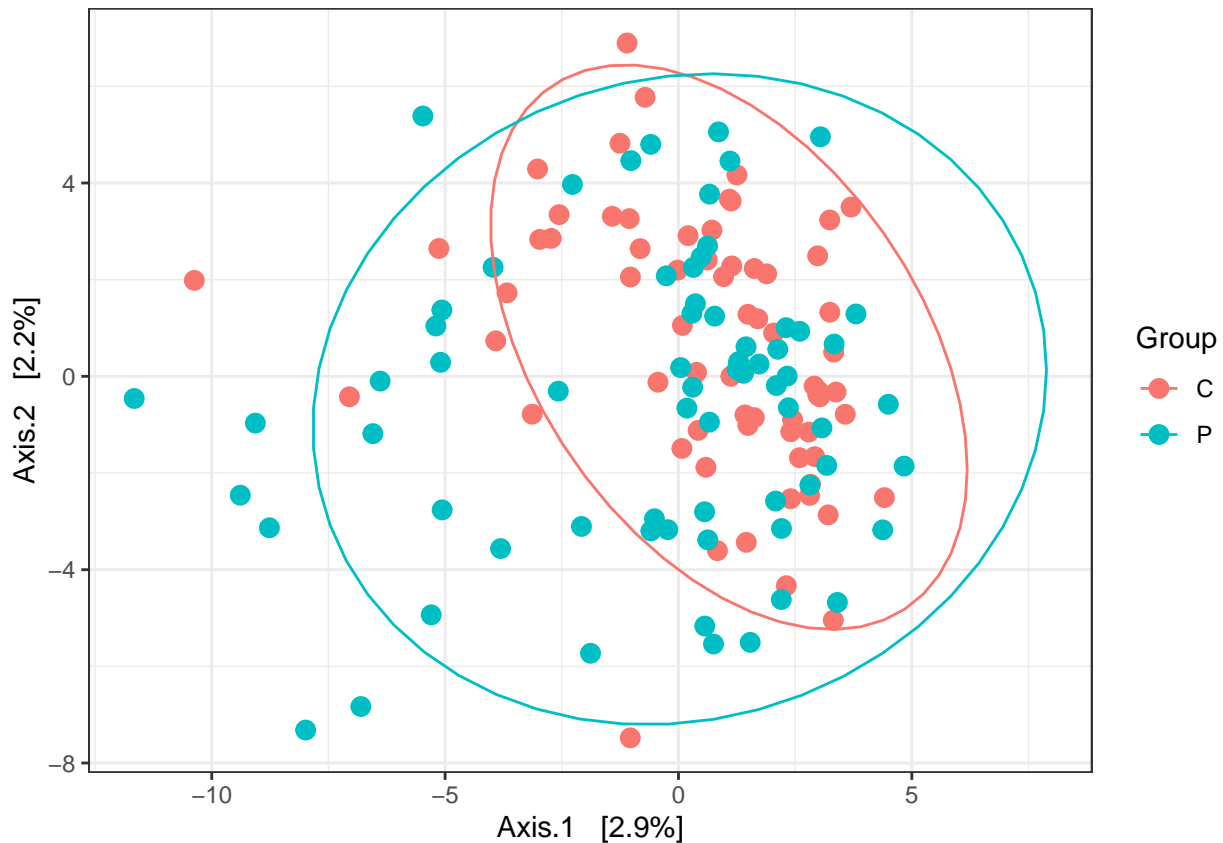
adonis.Res.clr

## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 10000
##
## adonis2(formula = physeqfinal.2.clr.eucl_dist ~ sample_data(physeqfinal.2.clr)$Group, permutations =
##
##           Df SumOfSqs      R2      F Pr(>F)
## sample_data(physeqfinal.2.clr)$Group    1      463 0.00888 1.1999 0.007599 **
## Residual                               134    51665 0.99112
## Total                                   135    52128 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Principal Coordinates Analysis (PCoA) for Group

```
ord_clr <- ordinate(physeqfinal.2.clr, method = "PCoA",
                   distance = "euclidean", binary = FALSE, autotransform = FALSE,
                   trymax = 1000, maxit = 10000, sratmax = 0.999999,
                   previous.best, parallel = 10)

plot_ordination(physeqfinal.2.clr, ord_clr,
                type = "Samples", color = "Group") + geom_point(size = 3) +
  stat_ellipse(level = 0.95)
```



Run the statistics for gender variable

```
set.seed(1337)
adonis.Res.clr <- adonis2(physeqfinal.2.clr.eucl_dist ~
  sample_data(physeqfinal.2.clr)$gender,
  perm = 10000,
  na.action = na.exclude,
  parallel = 10)

adonis.Res.clr

## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 10000
##
## adonis2(formula = physeqfinal.2.clr.eucl_dist ~ sample_data(physeqfinal.2.clr)$gender, permutations = 10000)
##              Df SumOfSqs      R2    F Pr(>F)
## sample_data(physeqfinal.2.clr)$gender  1      410 0.00786 1.0616 0.1859
## Residual                    134    51718 0.99214
## Total                        135    52128 1.00000
```

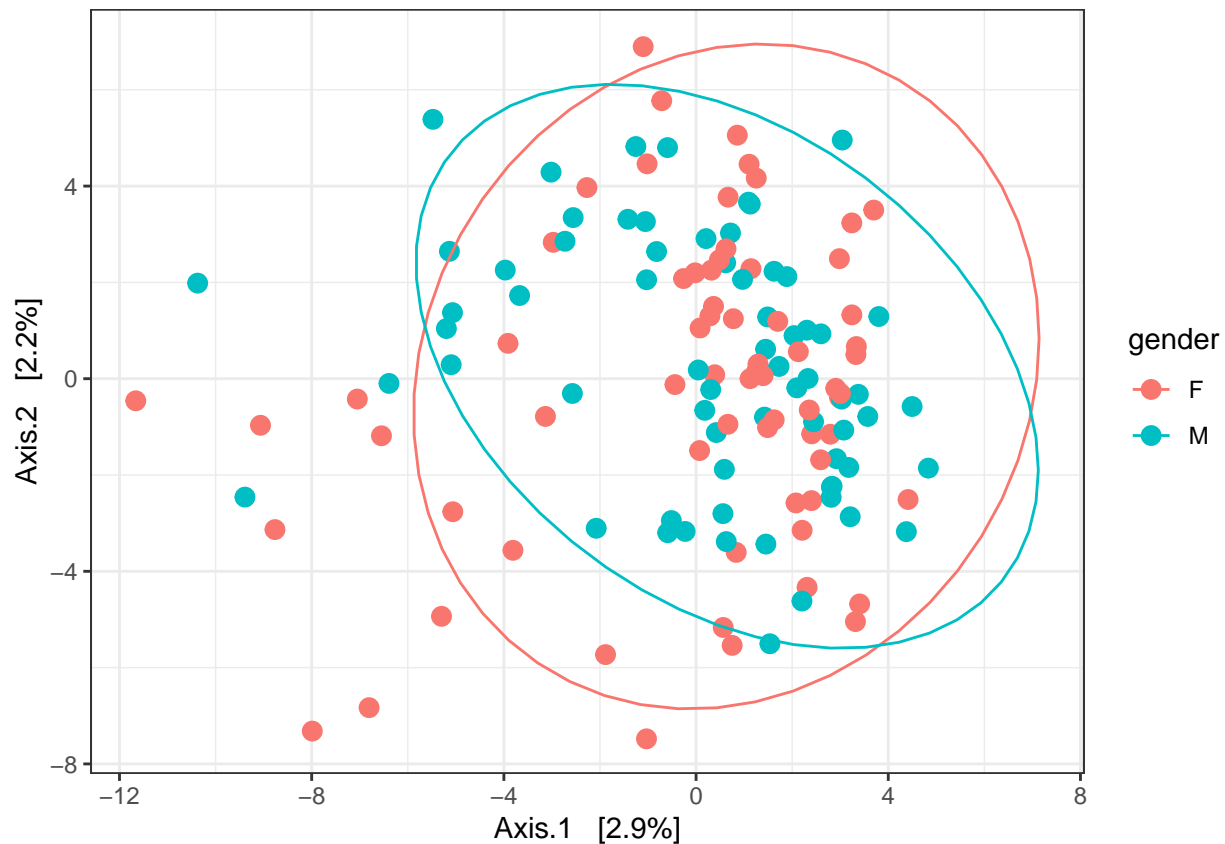
Principal Coordinates Analysis (PCoA) for Gender

```
ord_clr <- ordinate(physeqfinal.2.clr, method = "PCoA",
  distance = "euclidean", binary = FALSE, autotransform = FALSE,
  trymax = 1000, maxit = 10000, sratmax = 0.999999,
```

```

previous.best, parallel = 10)
plot_ordination(physeqfinal.2.clr, ord_clr,
  type = "Samples", color = "gender") + geom_point(size = 3) +
  stat_ellipse(level = 0.95)

```



Run the statistics for BMI variable

```

set.seed(1337)
adonis.Res.clr <- adonis2(physeqfinal.2.clr.eucl_dist ~
  sample_data(physeqfinal.2.clr)$BMI,
  perm = 10000,
  na.action = na.exclude,
  parallel = 10)

```

```
adonis.Res.clr
```

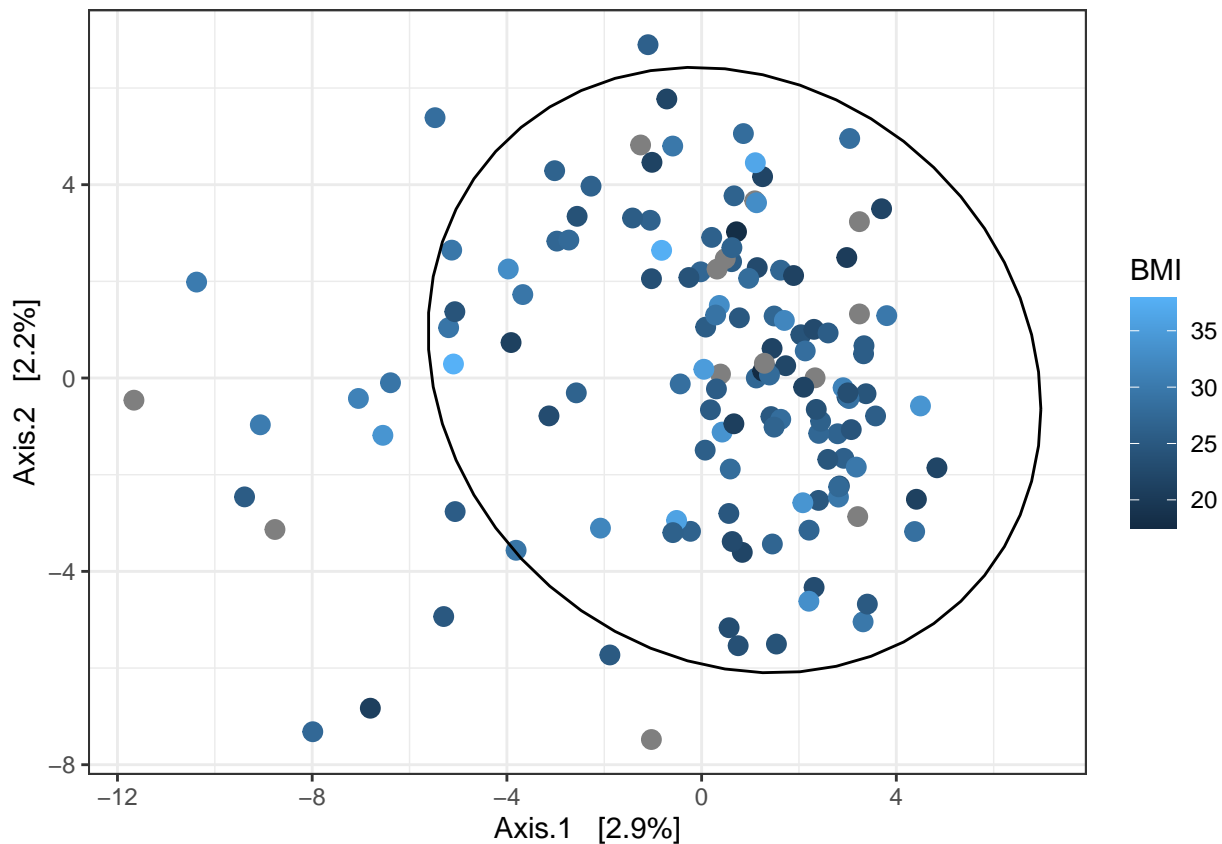
```

## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 10000
##
## adonis2(formula = physeqfinal.2.clr.eucl_dist ~ sample_data(physeqfinal.2.clr)$BMI, permutations = 10000)
##              Df SumOfSqs      R2      F Pr(>F)
## sample_data(physeqfinal.2.clr)$BMI  1      416 0.00885 1.0798 0.1374
## Residual                    121    46613 0.99115
## Total                        122    47029 1.00000

```

Principal Coordinates Analysis (PCoA) for BMI

```
ord_clr <- ordinate(physeqfinal.2.clr, method = "PCoA",
  distance = "euclidean", binary = FALSE, autotransform = FALSE,
  trymax = 1000, maxit = 10000, sratmax = 0.999999,
  previous.best, parallel = 10)
plot_ordination(physeqfinal.2.clr, ord_clr,
  type = "Samples", color = "BMI") + geom_point(size = 3) +
  stat_ellipse(level = 0.95)
```



Run the statistics for Age variable

```
set.seed(1337)
adonis.Res.clr <- adonis2(physeqfinal.2.clr.eucl_dist ~
  sample_data(physeqfinal.2.clr)$age_at_stool_collection,
  perm = 10000,
  na.action = na.exclude,
  parallel = 10)
```

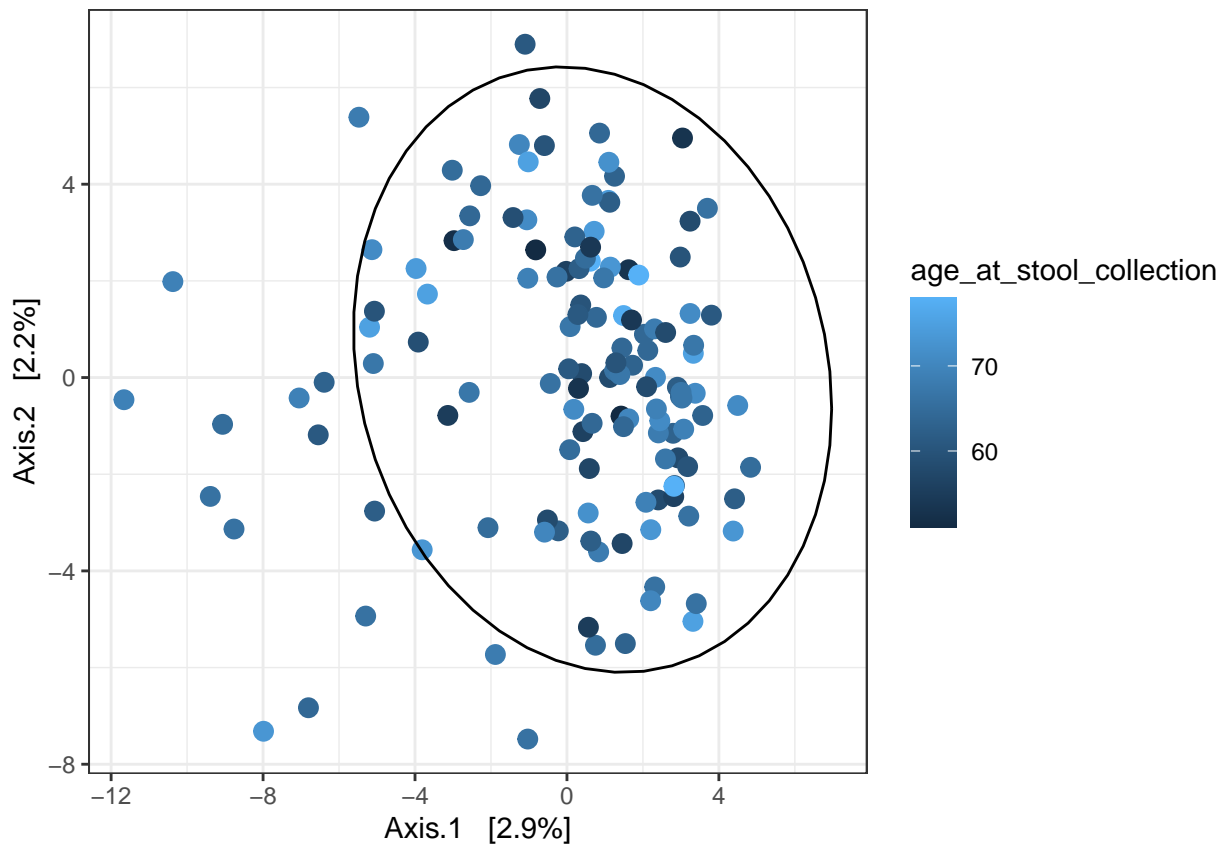
```
adonis.Res.clr
```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 10000
##
## adonis2(formula = physeqfinal.2.clr.eucl_dist ~ sample_data(physeqfinal.2.clr)$age_at_stool_collecti
```

```
##                                     Df SumOfSqs      R2
## sample_data(physeqfinal.2.clr)$age_at_stool_collection    1      375 0.0072
## Residual                                           134    51753 0.9928
## Total                                           135    52128 1.0000
##                                     F Pr(>F)
## sample_data(physeqfinal.2.clr)$age_at_stool_collection 0.9716 0.6278
## Residual
## Total
```

Principal Coordinates Analysis (PCoA) for Age

```
ord_clr <- ordinate(physeqfinal.2.clr, method = "PCoA",
                    distance = "euclidean", binary = FALSE, autotransform = FALSE,
                    trymax = 1000, maxit = 10000, sratmax = 0.999999,
                    previous.best, parallel = 10)
plot_ordination(physeqfinal.2.clr, ord_clr,
                type = "Samples", color = "age_at_stool_collection") + geom_point(size = 3) +
  stat_ellipse(level = 0.95)
```



Session Info for reproducibility

```
sessionInfo()
```

```
## R version 4.4.1 (2024-06-14)
## Platform: x86_64-pc-linux-gnu
## Running under: Ubuntu 22.04.4 LTS
```



```

##
## Matrix products: default
## BLAS: /usr/lib/x86_64-linux-gnu/openblas-pthread/libblas.so.3
## LAPACK: /usr/lib/x86_64-linux-gnu/openblas-pthread/libopenblas-p0.3.20.so; LAPACK version 3.10.0
##
## locale:
## [1] LC_CTYPE=en_GB.UTF-8 LC_NUMERIC=C
## [3] LC_TIME=en_GB.UTF-8 LC_COLLATE=en_GB.UTF-8
## [5] LC_MONETARY=en_GB.UTF-8 LC_MESSAGES=en_GB.UTF-8
## [7] LC_PAPER=en_GB.UTF-8 LC_NAME=C
## [9] LC_ADDRESS=C LC_TELEPHONE=C
## [11] LC_MEASUREMENT=en_GB.UTF-8 LC_IDENTIFICATION=C
##
## time zone: Europe/Helsinki
## tzcode source: system (glibc)
##
## attached base packages:
## [1] stats4 stats graphics grDevices utils datasets methods
## [8] base
##
## other attached packages:
## [1] microbiome_1.26.0 fido_1.1.1
## [3] DESeq2_1.44.0 SummarizedExperiment_1.34.0
## [5] Biobase_2.64.0 MatrixGenerics_1.16.0
## [7] matrixStats_1.3.0 GenomicRanges_1.56.1
## [9] GenomeInfoDb_1.40.1 IRanges_2.38.1
## [11] S4Vectors_0.42.1 BiocGenerics_0.50.0
## [13] vegan_2.6-6.1 lattice_0.22-5
## [15] permute_0.9-7 phyloseq_1.48.0
## [17] dabestr_2023.9.12 coin_1.4-3
## [19] survival_3.7-0 ggridges_0.5.6
## [21] qqplotr_0.0.6 MatrixCorrelation_0.10.0
## [23] energy_1.7-11 corrr_0.4.4
## [25] GGally_2.2.1 patchwork_1.2.0
## [27] cowplot_1.1.3 gridExtra_2.3
## [29] kableExtra_1.4.0 magrittr_2.0.3
## [31] purrr_1.0.2 reshape2_1.4.4
## [33] tidylog_1.1.0 tidyr_1.3.1
## [35] dplyr_1.1.4 RColorBrewer_1.1-3
## [37] ggplot2_3.5.1 BiocParallel_1.38.0
## [39] knitr_1.48
##
## loaded via a namespace (and not attached):
## [1] libcoin_1.0-10 tensorA_0.36.2.1 rstudioapi_0.16.0
## [4] jsonlite_1.8.8 TH.data_1.1-2 modeltools_0.2-23
## [7] farver_2.1.2 rmarkdown_2.27 zlibbioc_1.50.0
## [10] vctrs_0.6.5 multtest_2.60.0 tinytex_0.52
## [13] htmltools_0.5.8.1 S4Arrays_1.4.1 progress_1.2.3
## [16] distributional_0.4.0 plotrix_3.8-4 tidybayes_3.0.6
## [19] Rhdf5lib_1.26.0 SparseArray_1.4.8 rhdf5_2.48.0
## [22] pracma_2.4.4 plyr_1.8.9 sandwich_3.1-0
## [25] zoo_1.8-12 igraph_2.0.3 lifecycle_1.0.4
## [28] iterators_1.0.14 pkgconfig_2.0.3 Matrix_1.6-5
## [31] R6_2.5.1 fastmap_1.2.0 GenomeInfoDbData_1.2.12

```

## [34] digest_0.6.36	colorspace_2.1-1	RSpectra_0.16-2
## [37] labeling_0.4.3	fansi_1.0.6	httr_1.4.7
## [40] abind_1.4-5	mgcv_1.9-1	compiler_4.4.1
## [43] withr_3.0.1	doParallel_1.0.17	gsl_2.1-8
## [46] backports_1.5.0	ggstats_0.6.0	highr_0.11
## [49] MASS_7.3-61	DelayedArray_0.30.1	biomformat_1.32.0
## [52] caTools_1.18.2	tools_4.4.1	ape_5.8
## [55] qqconf_1.3.2	glue_1.7.0	nlme_3.1-165
## [58] rhdf5filters_1.16.0	grid_4.4.1	Rtsne_0.17
## [61] checkmate_2.3.2	cluster_2.1.6	ade4_1.7-22
## [64] generics_0.1.3	gtable_0.3.5	data.table_1.15.4
## [67] hms_1.1.3	xml2_1.3.6	utf8_1.2.4
## [70] XVector_0.44.0	ggdist_3.3.2	foreach_1.5.2
## [73] pillar_1.9.0	stringr_1.5.1	posterior_1.6.0
## [76] robustbase_0.99-3	splines_4.4.1	tidyselect_1.2.1
## [79] locfit_1.5-9.10	Biostrings_2.72.1	arrayhelpers_1.1-0
## [82] svglite_2.1.3	xfun_0.46	DEoptimR_1.1-3
## [85] stringi_1.8.4	UCSC.utils_1.0.0	yaml_2.3.10
## [88] boot_1.3-30	evaluate_0.24.0	codetools_0.2-19
## [91] twosamples_2.0.1	tibble_3.2.1	cli_3.6.3
## [94] pbmcapply_1.5.1	systemfonts_1.1.0	munsell_0.5.1
## [97] Rcpp_1.0.13	coda_0.19-4.1	svUnit_1.0.6
## [100] parallel_4.4.1	prettyunits_1.2.0	opdisDownsampling_1.0.1
## [103] bitops_1.0-8	viridisLite_0.4.2	mvtnorm_1.2-5
## [106] scales_1.3.0	crayon_1.5.3	clisymbols_1.2.0
## [109] rlang_1.1.4	multcomp_1.4-26	