Data analysis pipeline

Data analysis was performed using R version 4.1.3 via the platform RStudio 1.10.7 (2022.12.0). The packages "tidyverse" and "phyloseq" were pre-installed.

```
#Packages
Library("tidyverse")
Library("phyloseq")


#Data import
otu <- read.table("seqtab.csv", header = T, row.names = 1, sep = ",")
tax <- read.table("tax_table.csv", header = T, row.names = 1, sep = ",")
sam.data <- read.table("SAM.csv", header = T, row.names = 1, sep = ",")


#Format change of the import data
otu.data <- otu_table(t(otu), taxa_are_rows = T)
tax.data <- tax_table(as.matrix(tax))
sam.data <- sample_data
data<- phyloseq(otu.data,tax.data, sam.data)


#Variation in sequencing depth
barplot(sort(colSums(otu_table(data))), las=2, cex.names = 0.5)
> dim(otu_table(data))
[1] 2225    40
> colMeans(otu_table(data))
> mean(colSums(otu_table(data)))
[1] 31011.15


#Remove the OTUs with the lowest frequency
tax_table(data)[,c("Domain")] %>% unique()
> data_sub <- subset_taxa(data,
+                    Domain != ("Unassigned")
> data_sub <-  filter_taxa(data_sub,
+                    function(x) sum(x >= 10) > (1),
+                    prune =  TRUE)
> data_rare <- rarefy_even_depth(data_sub,
+                         sample.size = min(colSums(otu(data_sub))))


#Alpha diversity metrics
source("https://raw.githubusercontent.com/mahendra-mariadassou/phyloseq-
extended/master/load-extra-functions.R")
p <- ggrare(data, step = 50, color = "soil", plot = T, parallel = T, se =
F)
p <- plot_richness(data,
+                x="sample",
+                color="Stage",
+                measures=c("Observed","Shannon"),
+                nrow = 1)
print(p)


#Alpha diversity boxplots
(data_rare_20, fill="Species")
> library(data.table)
```

```
> rich.plus <- dcast(p$data,  samples + Stage ~ variable)
> boxplot(Observed~Stage, data=rich.plus)
> boxplot(Shannon~Stage, data=rich.plus)



#Beta diversity metrics
sample_data(data_species)
rownames(SAM)<- SAM$Barcode
SAM <- as.data.frame(SAM)
> rownames(SAM)<- SAM$Barcode
> sam.data <- sample_data(SAM)
> data<- phyloseq(otu.data,tax.data, sam.data)
> data_species <- tax_glom(data, taxrank = "Species")
> data_species %>% transform_sample_counts(function(x) x/sum(x)) %>%
+    otu_table() %>%
+    t() %>%
+    sqrt() %>%
+    as.data.frame() %>%
+    vegdist(binary=F, method = "bray") -> dist
> ord <- ordinate(data_species,"PCoA",dist)
> ord$vectors


plot_ordination(data_species,
+                ord,
+                color = "Stage",
+                title = "PCoA sqrt Bray curtis",
+                label= "SampleID" ) +
+    geom_point(aes(size=1)) +
+    theme_bw()


plot_ordination(data_species,
+                ord,
+                color = "Stage",
+                title = "PCoA sqrt Bray curtis",
+                label= "barcode" ) +
+    geom_point(aes(size=1)) +
+    theme_bw()


#Scientific Classification - barplot
TopNOTUs <- names(sort(taxa_sums(data_rare), TRUE)[1:20])
data_rare_20   <- prune_species(TopNOTUs, data_rare)
plot_bar(data_rare_20, fill="Order")

> data_rare_20   <- prune_taxa(TopNOTUs, data_species)
> plot_bar(data_rare_20, fill="Order")
> plot_bar(data_rare_20, fill="Genus")
> plot_bar(data_rare_20, fill="Species")
> subset_taxa(data_species, Species=="s__nigrolimitatus")


#tax_table
> View(otu_table(subset_taxa(data_species, Species=="s__species ")))



# wilcoxon test
> wilcox.test(Observed~Stage, data=rich.plus, subset =
Stage=="Laying_partially decomposed"|Stage=="Standing_hint of dryness")

        Wilcoxon rank sum test with continuity correction
```

```
data:  Observed by Stage
W = 98, p-value = 0.2879
alternative hypothesis: true location shift is not equal to 0
> wilcox.test(Observed~Stage, data=rich.plus, subset = Stage=="Laying_parti
ally decomposed"|Stage=="Laying_completely decomposed")

        Wilcoxon rank sum test with continuity correction

data:  Observed by Stage
W = 62, p-value = 0.1063
alternative hypothesis: true location shift is not equal to 0
> wilcox.test(Shannon~Stage, data=rich.plus, subset = Stage=="Laying_partia
lly decomposed"|Stage=="Standing_hint of dryness")

        Wilcoxon rank sum exact test

data:  Shannon by Stage
W = 22, p-value = 0.001553
alternative hypothesis: true location shift is not equal to 0
wilcox.test(Shannon~Stage, data=rich.plus, subset = Stage=="Laying_partiall
y decomposed"|Stage=="Laying_completely decomposed")

        Wilcoxon rank sum exact test

data:  Shannon by Stage
W = 50, p-value = 0.02872
alternative hypothesis: true location shift is not equal to 0



#Taxonomically analysis fungal OTUs by ecological guild
> library(readr)
> sample_guilds <- read_delim("sample_guilds.csv",
+                             delim = ";", escape_double = FALSE, trim_ws =
TRUE)


##FUNGuild
> install.packages("devtools")
> library(readr)
> sample_guilds <- read_delim("sample_guilds.csv",
                             delim = ";", escape_double = FALSE, trim_ws =
TRUE)

> funguild_data[,c(1:42)]<- apply(funguild_data[,c(1:42)], MARGIN = 2,
as.numeric)
> guild <- aggregate(funguild_data[,1:42], by=list(funguild_data$guild),
sum)
> guild<- guild[-1,]
> guild_selected<-guild[which(rowSums(guild[,-1])>5000),]

> guild_selected2<
guild_selected[,which(colnames(guild_selected)%in%sample_data(data_species)
$Barcode)]
> guild_selected2<- data.frame(guild_selected[,1], guild_selected2)

> guild_selected_ord <- guild_selected2[,
c(1,order(sample_data(data_species)$Stage)+1)]
> colnames(guild_selected_ord)[1]<- "Group.1"
> guild_selected_melted<- melt(guild_selected_ord)
> ggplot(guild_selected_melted, aes(x=variable, y=value,
fill=Group.1))+geom_bar(position="fill",stat="identity" )+
```

```
    theme(axis.text.x=element_text(angle=90))
> View(sample_data(data_species))

> rownames(guild_selected2)<- guild_selected2[,1]
> guild_selected2<- guild_selected2[,-1]

> test <- aggregate(t(guild_selected2),
by=list(sample_data(data_species)$Stage), sum)

> test$`Laying partially decomposed`<- as.numeric(test$`Laying partially
decomposed`)
pie(x = test$`Laying partially decomposed`)
pie(x = test$`Laying partially decomposed`, labels = rownames(test))

> test$`Standing hint of dryness`<- as.numeric(test$`Standing hint of
dryness`)
pie(x = test$`Standing hint of dryness`)
pie(x = test$`Standing hint of dryness`, labels = rownames(test))

> test$`Laying highly decomposed`<- as.numeric(test$`Laying highly
decomposed`)
pie(x = test$`Laying highly decomposed `)
pie(x = test$`Laying highly decomposed `, labels = rownames(test))
```