Is Habitat More Important than Phylogenetic Relatedness in Determining Gut Bacterial Composition in Sister Lizard Species?

Mauricio Hernández, Doctorado en CB, Centro Tlaxcala de Biología de la Conducta, UATx
Stephanie Hereira-Pacheco, Centro Tlaxcala de Biología de la Conducta, UATx

05 - 05 - 2022

Contents

Alpha Diversity

```
#loading packages
library(qiime2R)
library(hillR)
library(hilldiv)
library(tidyverse)
library(ggpubr)
library(phyloseq)
library(ggpubr)
```

```r
# FUNCTIONAL DIVERSITY
funciones <- read_tsv("../Data/EC_predicted.tsv") %>% arrange(desc(sequence)) %>%
  column_to_rownames(var = "sequence")

tabla <- data.frame(read_q2biom("../Data/feature-table.biom")) %>% rownames_to_column(
    var = "sequence") %>% arrange(desc(sequence)) %>% column_to_rownames(
  var = "sequence") %>% t() %>% as.data.frame()

# Run functional diversity by Hill numbers at different q orders
func_q0 <- hill_func(comm = tabla,traits = funciones, q = 0 )
func_q1 <- hill_func(comm = tabla,traits = funciones, q = 1 )
func_q2 <- hill_func(comm = tabla,traits = funciones, q = 2 )

# Saved as table
Hill_Funct <- cbind(func_q0, func_q1, func_q2)
#write.table(func_q0, file="./hill_funct_q0.txt", sep = "\t")
#write.table(func_q1, file="./hill_funct_q1.txt", sep = "\t")
#write.table(func_q2, file="./hill_funct_q2.txt", sep = "\t")

# PHYLOGENETIC DIVERSITY
# Run phylogenetic diversity by Hill numbers at different q orders
tree <- read_tree("tree_R10.nwk")
phylo_q0 <- hill_phylo(comm = tabla, tree = tree, q = 0) %>% as.data.frame()
phylo_q1 <- hill_phylo(comm = tabla, tree = tree, q = 1) %>% as.data.frame()
phylo_q2 <- hill_phylo(comm = tabla, tree = tree, q = 2) %>% as.data.frame()

Hill_Phylo <- cbind(phylo_q0, phylo_q1, phylo_q2)
#write.table(Hill_Phylo, file="./hill_phylo_div_R10.txt", sep = "\t")

# TAXONOMIC DIVERSITY
otutable <- read.csv(file = "../Data/feature_table.csv", header = TRUE, row.names = 1)

# Run taxonomic diversity by Hill numbers at different q orders
q0 <- hill_taxa(comm = t(otutable), q = 0)
q1 <- hill_taxa(comm = t(otutable), q = 1)
q2 <- hill_taxa(comm = t(otutable), q = 2)

Hill <- cbind(q0, q1, q2)
#write.table(Hill, file="./hill_taxa_numbers.txt", sep = "\t")
```

```r
#TAXONOMIC DIVERSITY
#load file previously obtained - alpha diversity
alpha_div <- read.csv("../Data/Hill_numbers_q012.csv",
                      header = TRUE) %>% dplyr::select(SampleID, q0, q1, q2)

metadata <- read.csv("../Data/metadata.csv", header = TRUE, check.names = F)

alpha_tax <- alpha_div %>% inner_join(metadata, by = c("SampleID"="SampleID"))

# Normality test
#shapiro.test(x =alpha_tax$q0)
#shapiro.test(x =alpha_tax$q1)
#shapiro.test(x =alpha_tax$q2)
```

```
# Paired test (Wilcoxon) (q0)
aeneus <-  subset(alpha_tax, Species_Sceloporus== "Sceloporus aeneus",
                  q0, drop = TRUE)
bicanthalis <- subset(alpha_tax, Species_Sceloporus== "Sceloporus bicanthalis",
                      q0, drop = TRUE)
avsb_q0 <- wilcox.test(x= aeneus, y = bicanthalis)
t.test(aeneus, bicanthalis)
```

```
##
##  Welch Two Sample t-test
##
## data:  aeneus and bicanthalis
## t = -1.1166, df = 14.314, p-value = 0.2826
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -61.90338  19.45893
## sample estimates:
## mean of x mean of y
##  125.2222  146.4444
```

```
# Paired test (Wilcoxon) (q1)
aeneus1 <-  subset(alpha_tax, Species_Sceloporus== "Sceloporus aeneus",
                   q1, drop = TRUE)
bicanthalis1 <- subset(alpha_tax, Species_Sceloporus== "Sceloporus bicanthalis",
                       q1, drop = TRUE)
AvsB_q1 <- wilcox.test(x= aeneus1, y= bicanthalis1)
t.test(aeneus1, bicanthalis1)
```

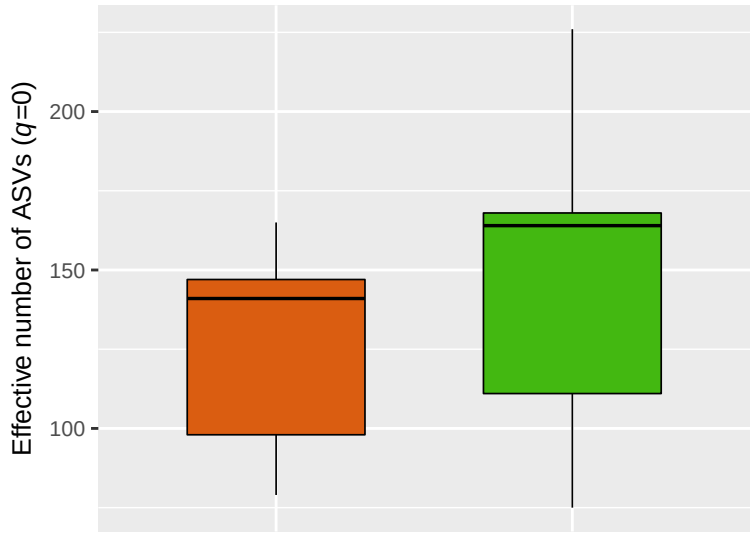
```
##
##  Welch Two Sample t-test
##
## data:  aeneus1 and bicanthalis1
## t = -3.202, df = 13.214, p-value = 0.006816
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -56.21125 -10.96306
## sample estimates:
## mean of x mean of y
##  55.03629  88.62345
```

```
# Paired test (Wilcoxon) (q2)
aeneus2 <-  subset(alpha_tax, Species_Sceloporus== "Sceloporus aeneus",
                   q2, drop = TRUE)
bicanthalis2 <- subset(alpha_tax, Species_Sceloporus== "Sceloporus bicanthalis",
                       q2, drop = TRUE)
AvsB_q2 <- wilcox.test(x= aeneus2, y= bicanthalis2)
t.test(aeneus2, bicanthalis2)
```

```
##
##  Welch Two Sample t-test
##
```

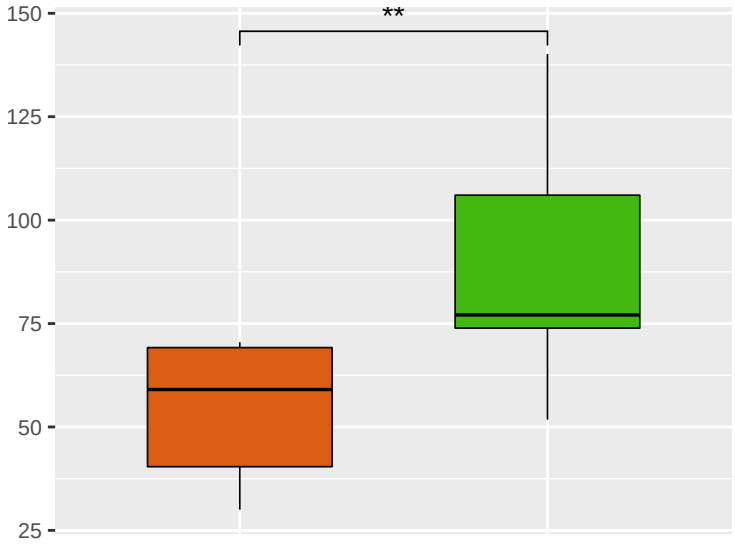
```
## data:  aeneus2 and bicanthalis2
## t = -3.6277, df = 13.293, p-value = 0.002964
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -45.02948 -11.46265
## sample estimates:
## mean of x mean of y
##  28.45872  56.70478
```

```
# Plotting (BOXPLOTS): taxonomic diversity
# q0
my_comparisons_q0 <- list(c("Sceloporus aeneus", "Sceloporus bicanthalis"))
titule0 <- expression(paste("Effective number of ASVs (", italic("q"), "=0)"))
tax_q0 <- ggboxplot(alpha_tax, x= "Species_Sceloporus", y= "q0",
                          color = "black",  width = 0.6, lwd=0.3,
                          order = c("Sceloporus aeneus", "Sceloporus bicanthalis"),
                          fill = c("#DA5D11","#43B811")) +
  labs(x = element_blank(), y = "Effective number of ASVs") +
  theme_gray() + theme(text = element_text (size = 10)) +
  theme(axis.ticks.x = element_blank(),
        axis.text.x = element_blank()) +
 #stat_compare_means(comparisons = my_comparisons_q0, label = "p.signif") +
  ylab(titule0)
print(tax_q0)
```

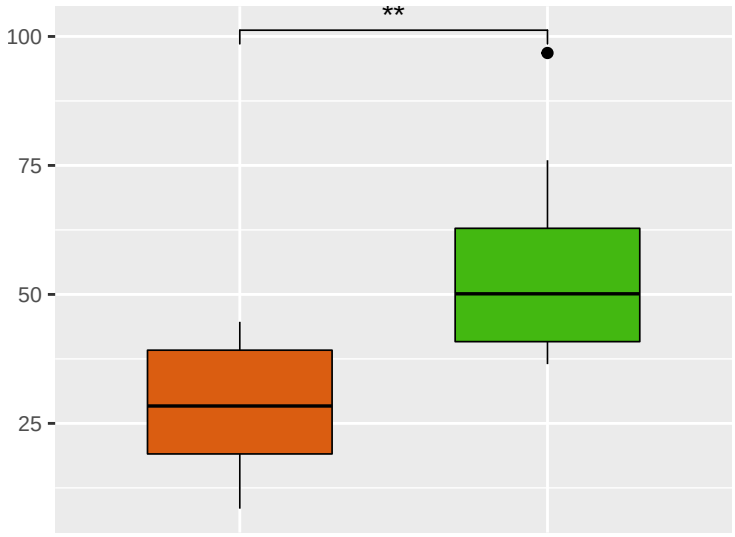



```
#ggsave("tax_q0.jpeg", width=5.5, height=4.5, dpi=300)

# q1
my_comparisons_q1 <- list(c("Sceloporus aeneus", "Sceloporus bicanthalis"))
tax_q1 <- ggboxplot(alpha_tax, x= "Species_Sceloporus", y= "q1",
                    color = "black",  width = 0.6, lwd=0.3,
                    order = c("Sceloporus aeneus", "Sceloporus bicanthalis"),
                    fill = c("#DA5D11","#43B811")) +
  labs(x = element_blank(), y = element_blank()) +
  theme_gray() + theme(text = element_text (size = 10)) +
  theme(legend.position = "none",
```

```
        axis.ticks.x = element_blank(),
        axis.text.x = element_blank()) +
  stat_compare_means(comparisons = my_comparisons_q1, label = "p.signif")
print(tax_q1)
```




```
#ggsave("tax_q1.jpeg", width=5.5, height=4.5, dpi=300)

# q2
my_comparisons_q2 <- list(c("Sceloporus aeneus", "Sceloporus bicanthalis"))
tax_q2 <- ggboxplot(alpha_tax, x= "Species_Sceloporus", y= "q2",
                    color = "black",  width = 0.6, lwd=0.3,
                    order = c("Sceloporus aeneus", "Sceloporus bicanthalis"),
                    fill = c("#DA5D11","#43B811")) +
  labs(x = element_blank(), y = element_blank()) +
  theme_gray() + theme(text = element_text (size = 10)) +
  theme(legend.position = "none",
       axis.ticks.x = element_blank(),
      axis.text.x = element_blank()) +
  stat_compare_means(comparisons = my_comparisons_q2, label = "p.signif")
print(tax_q2)
```




```
#ggsave("tax_q2.jpeg", width=5.5, height=4.5, dpi=300)
```

```
#PHYLOGENETIC DIVERSITY
phylo_div <- read.csv("../Data/Hill_phylo1_q012_R10.csv",
                      header = TRUE) %>% dplyr::select(SampleID, q0, q1, q2)
metadata <- read.csv("../Data/metadata.csv", header = TRUE, check.names = F)
phylo_tax <- phylo_div %>% inner_join(metadata, by = c("SampleID"="SampleID"))

# Normality test
#shapiro.test(x =phylo_tax$q0)
#shapiro.test(x =phylo_tax$q1)
#shapiro.test(x =phylo_tax$q2)


# Paired test (Wilcoxon) (q0)
aeneus_phylo <-  subset(phylo_tax, Species_Sceloporus== "Sceloporus aeneus",
                        q0, drop = TRUE)
bicanthalis_phylo <- subset(phylo_tax, Species_Sceloporus== "Sceloporus bicanthalis",
                            q0, drop = TRUE)
avsb_phylo_q0 <- wilcox.test(x= aeneus_phylo, y = bicanthalis_phylo)
t.test(aeneus_phylo, bicanthalis_phylo)
```

```
##
##  Welch Two Sample t-test
##
## data:  aeneus_phylo and bicanthalis_phylo
## t = -1.4141, df = 15.772, p-value = 0.1768
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.4519682  0.6913719
## sample estimates:
## mean of x mean of y
##  12.39798  13.77828
```

```r
# Paired test (Wilcoxon) (q1)
aeneus_phylo1 <-  subset(phylo_tax, Species_Sceloporus== "Sceloporus aeneus",
                         q1, drop = TRUE)
bicanthalis_phylo1 <- subset(phylo_tax, Species_Sceloporus== "Sceloporus bicanthalis",
                             q1, drop = TRUE)
AvsB_phylo_q1 <- wilcox.test(x= aeneus_phylo1, y= bicanthalis_phylo1)
t.test(aeneus_phylo1, bicanthalis_phylo1)
```

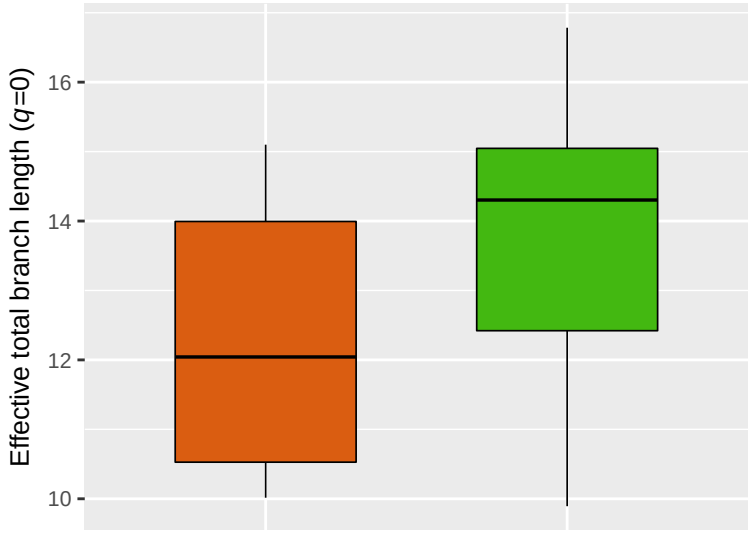
```
##
##  Welch Two Sample t-test
##
## data:  aeneus_phylo1 and bicanthalis_phylo1
## t = -2.9102, df = 10.696, p-value = 0.01457
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.44439171 -0.06090465
## sample estimates:
## mean of x mean of y
##  2.426373  2.679021
```

```r
# Paired test (Wilcoxon) (q2)
aeneus_phylo2 <-  subset(phylo_tax, Species_Sceloporus== "Sceloporus aeneus",
                         q2, drop = TRUE)
bicanthalis_phylo2 <- subset(phylo_tax, Species_Sceloporus== "Sceloporus bicanthalis",
                             q2, drop = TRUE)
AvsB_phylo_q2 <- wilcox.test(x= aeneus_phylo2, y= bicanthalis_phylo2)
t.test(aeneus_phylo2, bicanthalis_phylo2)
```

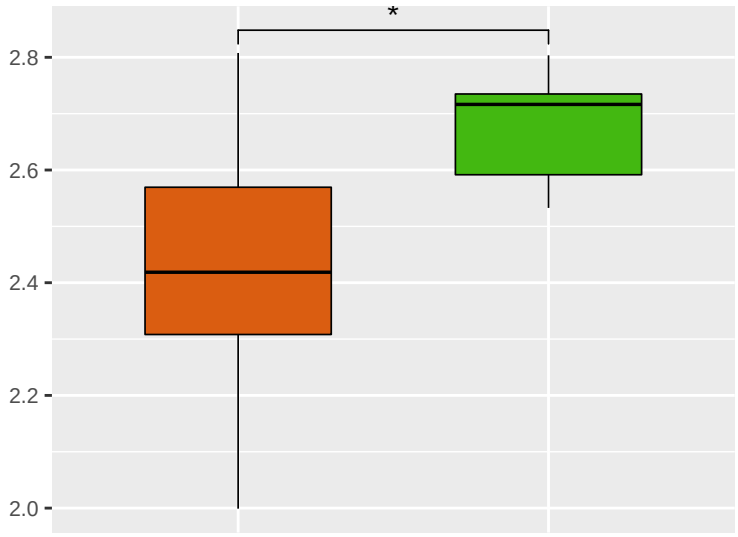
```
##
##  Welch Two Sample t-test
##
## data:  aeneus_phylo2 and bicanthalis_phylo2
## t = -1.2472, df = 9.975, p-value = 0.2408
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1777080  0.0501848
## sample estimates:
## mean of x mean of y
##  1.688266  1.752028
```

```r
# Plotting (BOXPLOTS): phylogenetic diversity
# q0
my_comparisons_q0 <- list(c("Sceloporus aeneus", "Sceloporus bicanthalis"))
tit0 <- expression(paste("Effective total branch length (", italic("q"), "=0)"))
phylo_q0 <- ggboxplot(phylo_tax, x= "Species_Sceloporus", y= "q0",
                      color = "black",  width = 0.6, lwd=0.3,
                      order = c("Sceloporus aeneus", "Sceloporus bicanthalis"),
                      fill = c("#DA5D11","#43B811")) +
  labs(x = element_blank(), y = "Effective total branch length") +
  theme_gray() + theme(text = element_text (size = 10)) +
  theme(axis.ticks.x = element_blank(),
        axis.text.x = element_blank()) +
```

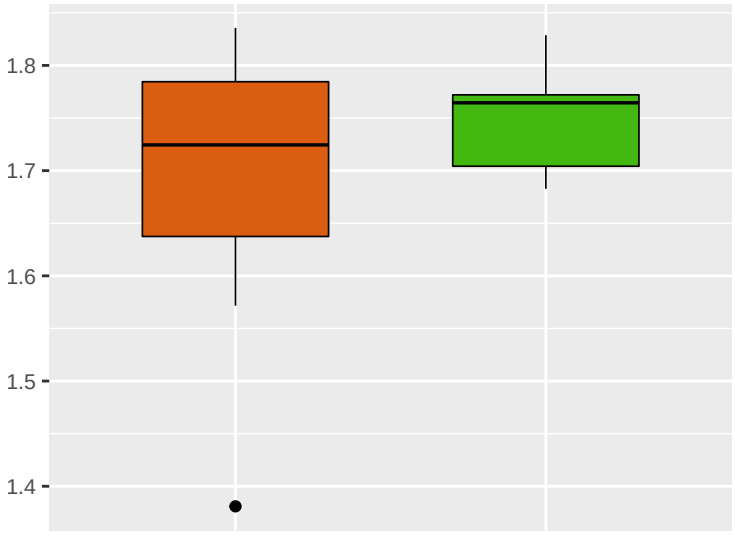
```
  #stat_compare_means(comparisons = my_comparisons_q0, label = "p.signif") +
  ylab(tit0)
print(phylo_q0)
```




```
#ggsave("phylo_q0.jpeg", width=5.5, height=4.5, dpi=300)

# q1
my_comparisons_q1 <- list(c("Sceloporus aeneus", "Sceloporus bicanthalis"))
phylo_q1 <- ggboxplot(phylo_tax, x= "Species_Sceloporus", y= "q1",
                      color = "black",  width = 0.6, lwd=0.3,
                      order = c("Sceloporus aeneus", "Sceloporus bicanthalis"),
                      fill = c("#DA5D11","#43B811")) +
  labs(x = element_blank(), y = element_blank()) +
  theme_gray() + theme(text = element_text (size = 10)) +
  theme(legend.position = "none",
        axis.ticks.x = element_blank(),
        axis.text.x = element_blank()) +
  stat_compare_means(comparisons = my_comparisons_q1, label = "p.signif")
print(phylo_q1)
```




```
#ggsave("phylo_q1.jpeg", width=5.5, height=4.5, dpi=300)

# q2
my_comparisons_q0 <- list(c("Sceloporus aeneus", "Sceloporus bicanthalis"))
phylo_q2 <- ggboxplot(phylo_tax, x= "Species_Sceloporus", y= "q2",
                      color = "black",  width = 0.6, lwd=0.3,
                      order = c("Sceloporus aeneus", "Sceloporus bicanthalis"),
                      fill = c("#DA5D11","#43B811")) +
  labs(x = element_blank(), y = element_blank()) +
  theme_gray() + theme(text = element_text (size = 10)) +
  theme(legend.position = "none",
        axis.ticks.x = element_blank(),
        axis.text.x = element_blank())
 # stat_compare_means(comparisons = my_comparisons_q0, label = "p.signif")
print(phylo_q2)
```




```
#ggsave("phylo_q2.jpeg", width=5.5, height=4.5, dpi=300)
```

```r
# FUNCTIONAL DIVERSITY
Funct_div <- read.csv("../Data/Hill_funct_q012.csv", header = TRUE) %>%
  dplyr::select(SampleID, MD_q0, MD_q1, MD_q2)
metadata <- read.csv("../Data/metadata.csv", header = TRUE, check.names = F)
funct_tax <- Funct_div %>% inner_join(metadata, by = c("SampleID"="SampleID"))

# Normality test
#shapiro.test(x =funct_tax$MD_q0)
#shapiro.test(x =funct_tax$MD_q1)
#shapiro.test(x =funct_tax$MD_q2)

# Paired test (Wilcoxon) (MD_q0)
aeneus_funct <-  subset(funct_tax, Species_Sceloporus== "Sceloporus aeneus",
                        MD_q0, drop = TRUE)
bicanthalis_funct <- subset(funct_tax, Species_Sceloporus== "Sceloporus bicanthalis",
                            MD_q0, drop = TRUE)
avsb_MD_q0_funct <- wilcox.test(x= aeneus_funct, y = bicanthalis_funct)
t.test(aeneus_funct, bicanthalis_funct)
```

```
##
##  Welch Two Sample t-test
##
## data:  aeneus_funct and bicanthalis_funct
## t = -1.7071, df = 13.226, p-value = 0.1112
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -4287.2629   498.8026
## sample estimates:
## mean of x mean of y
##  7528.159  9422.389
```

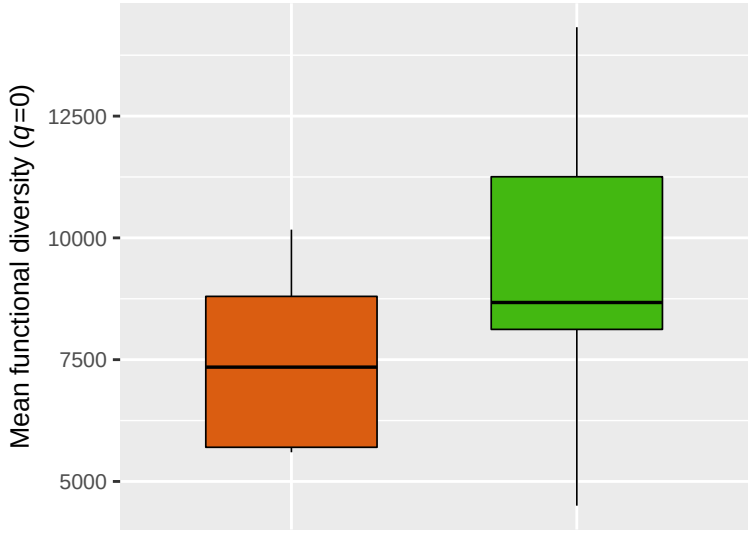
```r
# Paired test (Wilcoxon) (MD_q1)
aeneus_funct1 <-  subset(funct_tax, Species_Sceloporus== "Sceloporus aeneus",
                         MD_q1, drop = TRUE)
bicanthalis_funct1 <- subset(funct_tax, Species_Sceloporus== "Sceloporus bicanthalis",
                             MD_q1, drop = TRUE)
avsb_MD_q1_funct1 <- wilcox.test(x= aeneus_funct1, y = bicanthalis_funct1)
t.test(aeneus_funct1, bicanthalis_funct1)
```

```
##
##  Welch Two Sample t-test
##
## data:  aeneus_funct1 and bicanthalis_funct1
## t = -3.8106, df = 9.6326, p-value = 0.003666
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3556.7579  -923.4681
## sample estimates:
## mean of x mean of y
##  3502.423  5742.536
```

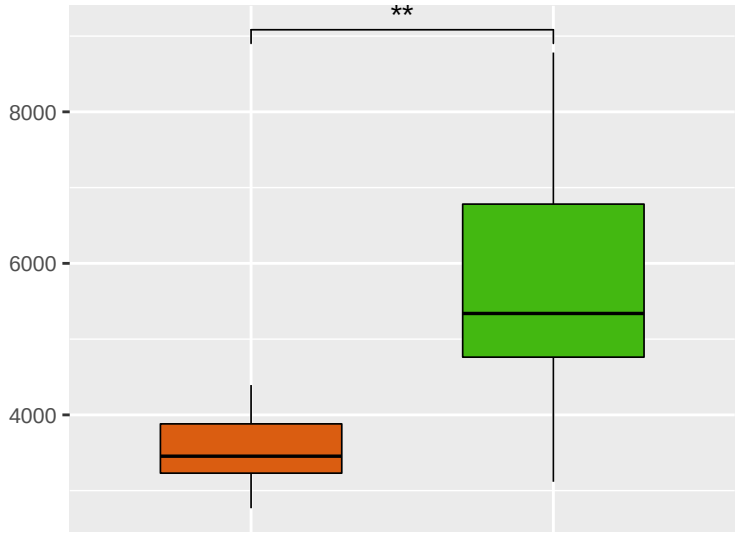
```
# Paired test (Wilcoxon) (MD_q1)
aeneus_funct2 <-  subset(funct_tax, Species_Sceloporus== "Sceloporus aeneus",
                         MD_q2, drop = TRUE)
bicanthalis_funct2 <- subset(funct_tax, Species_Sceloporus== "Sceloporus bicanthalis",
                             MD_q2, drop = TRUE)
avsb_MD_q2_funct2 <- wilcox.test(x= aeneus_funct2, y = bicanthalis_funct2)
t.test(aeneus_funct2, bicanthalis_funct2)
```

```
##
##  Welch Two Sample t-test
##
## data:  aeneus_funct2 and bicanthalis_funct2
## t = -4.1748, df = 9.6844, p-value = 0.002041
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -2821.3439  -852.0863
## sample estimates:
## mean of x mean of y
##  1976.806  3813.521
```

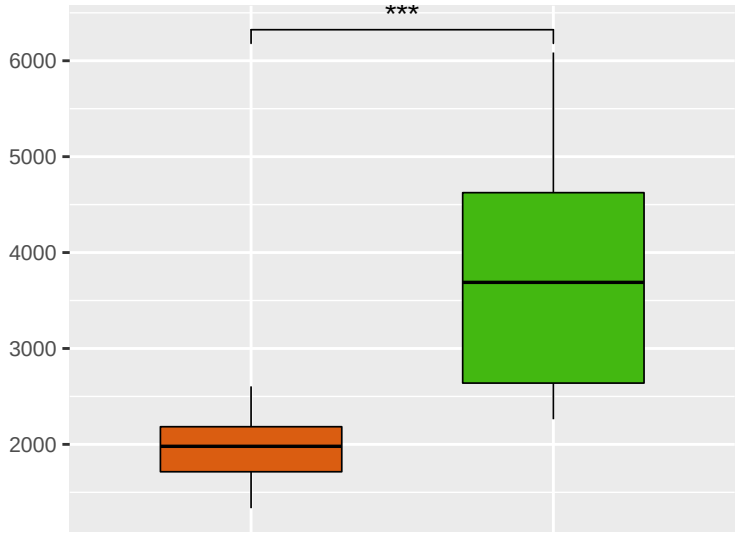
```
# Plotting (BOXPLOTS): functional diversity

# MD_q0
my_comparisons_MD_q0 <- list(c("Sceloporus aeneus", "Sceloporus bicanthalis"))
tittle0 <- expression(paste("Mean functional diversity (", italic("q"), "=0)"))
funct_q0 <- ggboxplot(funct_tax, x= "Species_Sceloporus", y= "MD_q0",
                    color = "black",  width = 0.6, lwd=0.3,
                    order = c("Sceloporus aeneus", "Sceloporus bicanthalis"),
                    fill = c("#DA5D11","#43B811")) +
  labs(x = element_blank(), y = "Mean functional diversity") +
  theme_gray() + theme(text = element_text (size = 10),
                       axis.text.x = element_text(face = "italic")) +
 theme(axis.ticks.x = element_blank(),
   axis.text.x = element_blank()) +
  ylab(tittle0)
 # stat_compare_means(comparisons = my_comparisons_q0, label = "p.signif")
print(funct_q0)
```




```
#ggsave("funct_q0.jpeg", width=5.5, height=4.5, dpi=300)

# MD_q1
my_comparisons_MD_q1 <- list(c("Sceloporus aeneus", "Sceloporus bicanthalis"))
funct_q1 <- ggboxplot(funct_tax, x= "Species_Sceloporus", y= "MD_q1",
                      color = "black",  width = 0.6, lwd=0.3,
                      order = c("Sceloporus aeneus", "Sceloporus bicanthalis"),
                      fill = c("#DA5D11","#43B811")) +
  labs(x = element_blank(), y = element_blank()) +
  theme_gray() + theme(text = element_text (size = 10),
                       axis.text.x = element_text(face = "italic")) +
  theme(legend.position = "none",
        axis.ticks.x = element_blank(),
       axis.text.x = element_blank()) +
  stat_compare_means(comparisons = my_comparisons_MD_q1, label = "p.signif")
print(funct_q1)
```

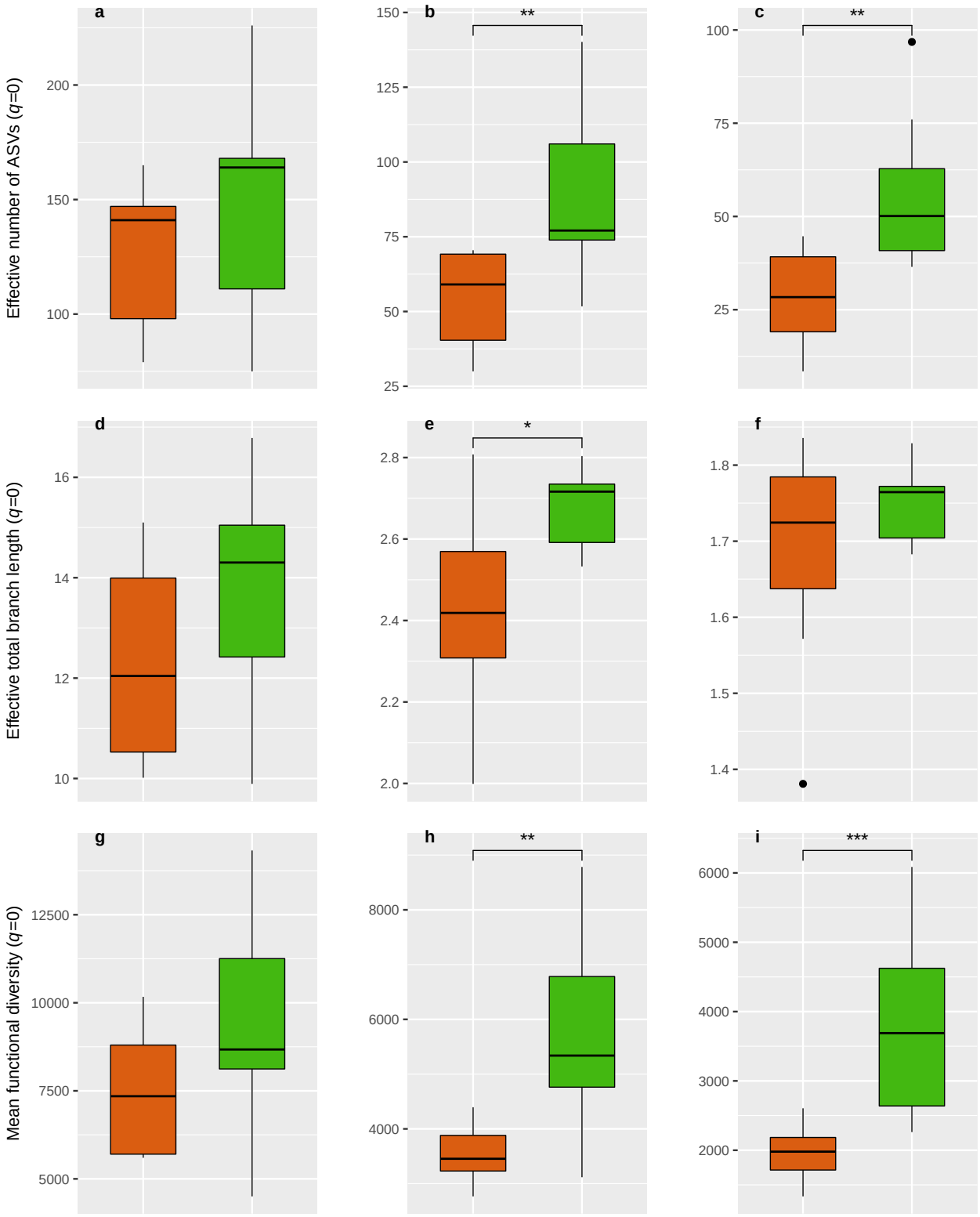


```
#ggsave("funct_q1.jpeg", width=5.5, height=4.5, dpi=300)

# MD_q2
my_comparisons_MD_q2 <- list(c("Sceloporus aeneus", "Sceloporus bicanthalis"))
funct_q2 <- ggboxplot(funct_tax, x= "Species_Sceloporus", y= "MD_q2",
                      color = "black",  width = 0.6, lwd=0.3,
                      order = c("Sceloporus aeneus", "Sceloporus bicanthalis"),
                      fill = c("#DA5D11","#43B811")) +
  labs(x = element_blank(), y = element_blank()) +
  theme_gray() + theme(text = element_text (size = 10),
                       axis.text.x = element_text(face = "italic")) +
  theme(legend.position = "none",
        axis.ticks.x = element_blank(),
       axis.text.x = element_blank()) +
  stat_compare_means(comparisons = my_comparisons_MD_q2, label = "p.signif")
print(funct_q2)
```




```
#ggsave("funct_q2.jpeg", width=5.5, height=4.5, dpi=300)
```

```
library(cowplot)
Tax_Funct_Phylo_Div <- plot_grid(tax_q0, tax_q1, tax_q2,
                                 phylo_q0, phylo_q1, phylo_q2,
                                 funct_q0, funct_q1, funct_q2,
                                nrow = 3, ncol = 3,
                                labels = c("a", "b","c",
                                           "d", "e", "f",
                                           "g", "h", "i"),
                                label_x = 0.3, hjust = 0.1,
                                label_colour = "black",
                                align = "hv",
                                label_size = 10, rel_heights = c(1, 1, 1))

print(Tax_Funct_Phylo_Div)
```



Joined plot

```
#ggsave("Tax_Funct_Phylo_Div.jpeg", width=6.0, height=11.0, dpi=300)
```

Beta diversity - PCA and perMANOVA

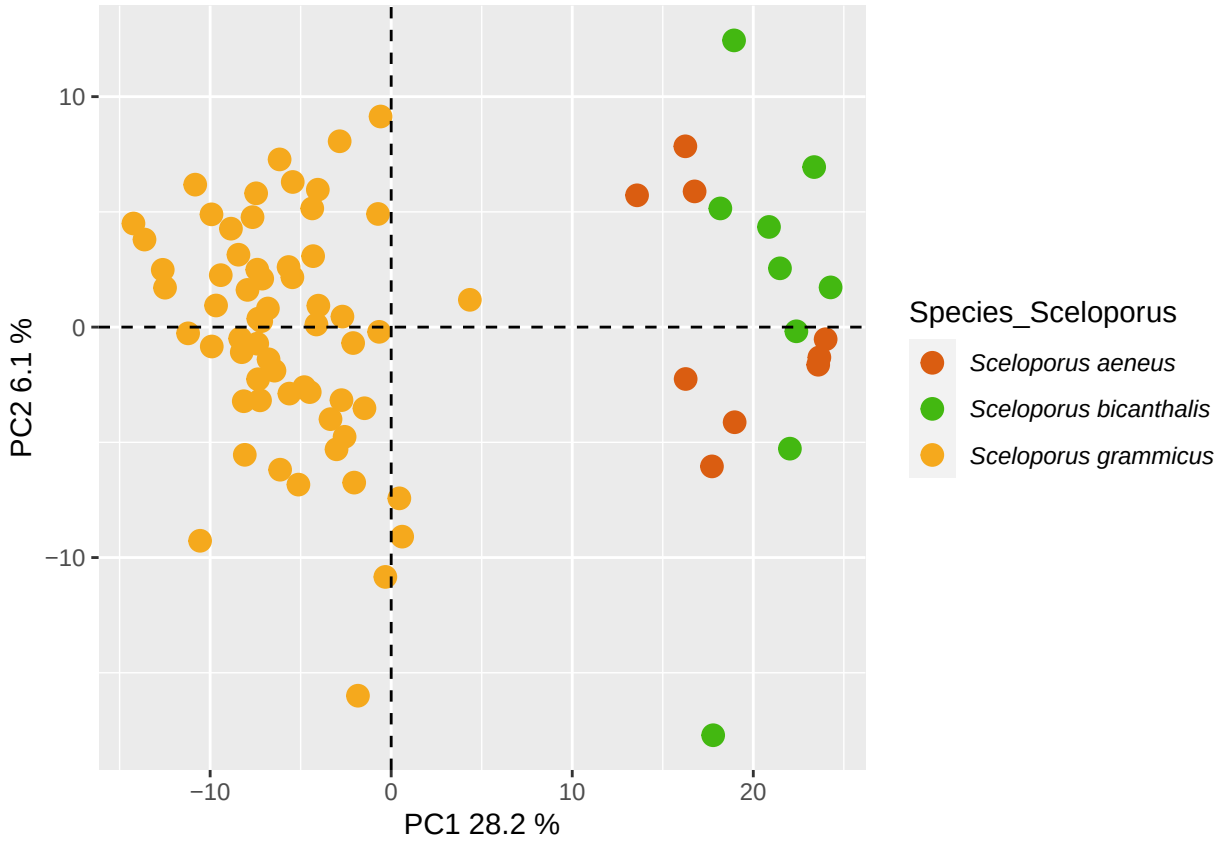
```
## Loading packages
library(tidyverse)
library(compositions)
library(zCompositions)
library(ALDEx2)
library(CoDaSeq)
library(vegan)
library(RVAideMemoire)
library(ggpubr)
library(ALDEx2)
library(ComplexHeatmap)
library(RColorBrewer)
library(circlize)
library(ggplotify)

# Loading files
metadata <- read.csv("../Data/metadata_aeneus_bica_grammicus.csv", row.names = 1,
                     header = TRUE, check.names = F)

otutable <- read.csv("../Data/feature_table.csv", row.names = 1, check.names = F)
```

```
# Transforming data "clr transformation/compositional data"
aldez.clr.transform <- aldex.clr(otutable, mc.samples = 999, denom = "all",
                                 verbose = FALSE, useMC = FALSE)
aldex.clr.transform.data <- t(getMonteCarloSample(aldez.clr.transform, 1))

#write.table(aldex.clr.transform.data, file="./aldex_table.txt", sep = "\t")
```

All species

```
# PCA
metadata <- read.csv("../Data/metadata_aeneus_bica_grammicus.csv",
                     row.names = 1, check.names = F) %>%
  rownames_to_column(var = "SampleID")
otu_table1 <- read.csv("../Data/aldex_table.csv", header = TRUE,
                       row.names = 1, check.names = F)

# Run a PCA with codaSeq.clr
pcx.abund.aldex <- prcomp(otu_table1)

# Labels to PCA axis
pc1 <- paste("PC1", round(sum(pcx.abund.aldex$sdev[1] ^2) /
                            mvar(otu_table1) * 100, 1), "%")

pc2 <- paste("PC2", round(sum(pcx.abund.aldex$sdev[2] ^2) /
                            mvar(otu_table1) * 100, 1), "%")

# Create the base plot with only the arrows
```

```
PCA_SaSbSg_aldex.clr <- ggplot() +
  theme_bw() +
  xlab(pc1) +
  ylab(pc2) +
  theme(axis.text = element_text(colour = "black", size = 14),
        axis.title = element_text(colour = "black", size = 14),
        legend.text = element_text(size = 14),
        legend.title = element_blank(),
        legend.position = "right") +
  theme_gray() + geom_point( #individuals
    data = data.frame(pcx.abund.aldex$x) %>% rownames_to_column(
      var = "SampleID") %>%
      left_join(metadata, by = "SampleID"),
    aes(x=PC1, y=PC2, color = Species_Sceloporus),
    size= 3.5) +
  geom_vline(xintercept = 0, linetype = 2) +   #lines-cross
  geom_hline(yintercept = 0, linetype = 2) +
  scale_color_manual(values = c("#DA5D11","#43B811","#F5A91D")) +
  theme(legend.text = element_text(face = "italic"))# +

print(PCA_SaSbSg_aldex.clr)
```



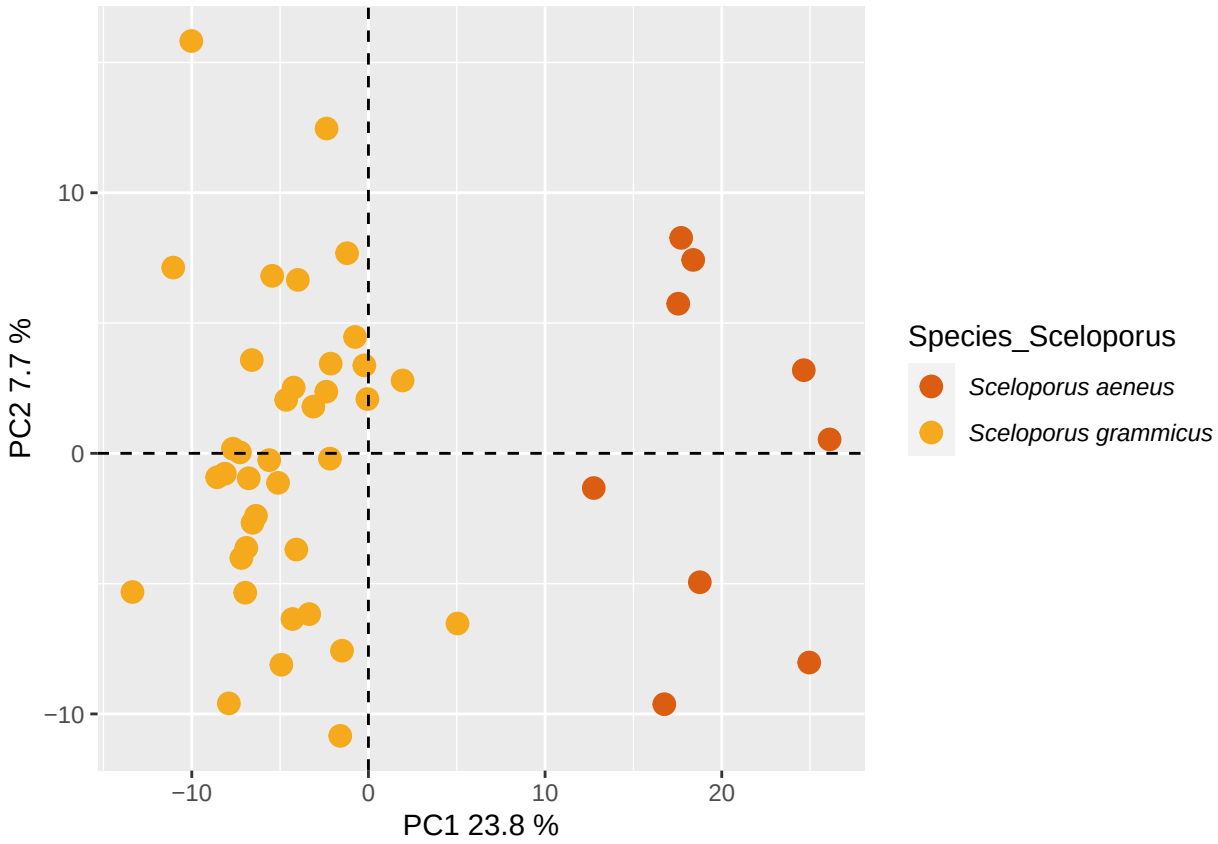

```
ggsave("PCA_SaSbSg_aldex.clr.jpeg", width=5.5, height=5.5, dpi=300)
```

```r
# perMANOVA all Species

Permanova1 <- adonis2(otu_table1 ~ Species_Sceloporus, data = metadata,
                      method = "euclidean", permutations = 999)
print(Permanova1)
```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = otu_table1 ~ Species_Sceloporus, data = metadata, permutations = 999, method = "eu
##                     Df SumOfSqs      R2      F Pr(>F)
## Species_Sceloporus  2     9853 0.26801 13.913  0.001 ***
## Residual           76    26911 0.73199
## Total              78    36764 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Aeneus vs Grammicus

```
# SCELOPORUS AENEUS VERSUS SCELOPORUS GRAMMICUS
otu_table2 <- read.csv(file = "../Data/aldex_Sasg.csv",
                       header = TRUE, row.names = 1, check.names = F)
metadata2 <- read.csv("../Data/metadata_AG.csv", row.names = 1,
                      check.names = F) %>% rownames_to_column(var = "SampleID")

# Run a PCA with codaSeq.clr
pcx.abund.aldex <- prcomp(otu_table2)

# Labels to PCA axis
pc1 <- paste("PC1", round(sum(pcx.abund.aldex$sdev[1] ^2) /
                            mvar(otu_table2) * 100, 1), "%")

pc2 <- paste("PC2", round(sum(pcx.abund.aldex$sdev[2] ^2) /
                            mvar(otu_table2) * 100, 1), "%")

# Create the base plot with only the arrows
PCA_SaSg_aldex.clr <- ggplot() +
  theme_bw() +
  xlab(pc1) +
  ylab(pc2) +
  theme(axis.text = element_text(colour = "black", size = 14), #setting theme
        axis.title = element_text(colour = "black", size = 14),
        legend.text = element_text(size = 14),
        legend.title = element_blank(),
        legend.position = "right") +
  theme_gray() +
  geom_point( #individuals
    data = data.frame(pcx.abund.aldex$x) %>% rownames_to_column(
      var = "SampleID") %>%
      left_join(metadata, by = "SampleID"),
    aes(x=PC1, y=PC2, color = Species_Sceloporus),
    size= 3.5) +
  geom_vline(xintercept = 0, linetype = 2) +
  geom_hline(yintercept = 0, linetype = 2) +
  scale_color_manual(values = c("#DA5D11","#F5A91D")) +
  theme(legend.text = element_text(face = "italic")) # +


print(PCA_SaSg_aldex.clr)
```

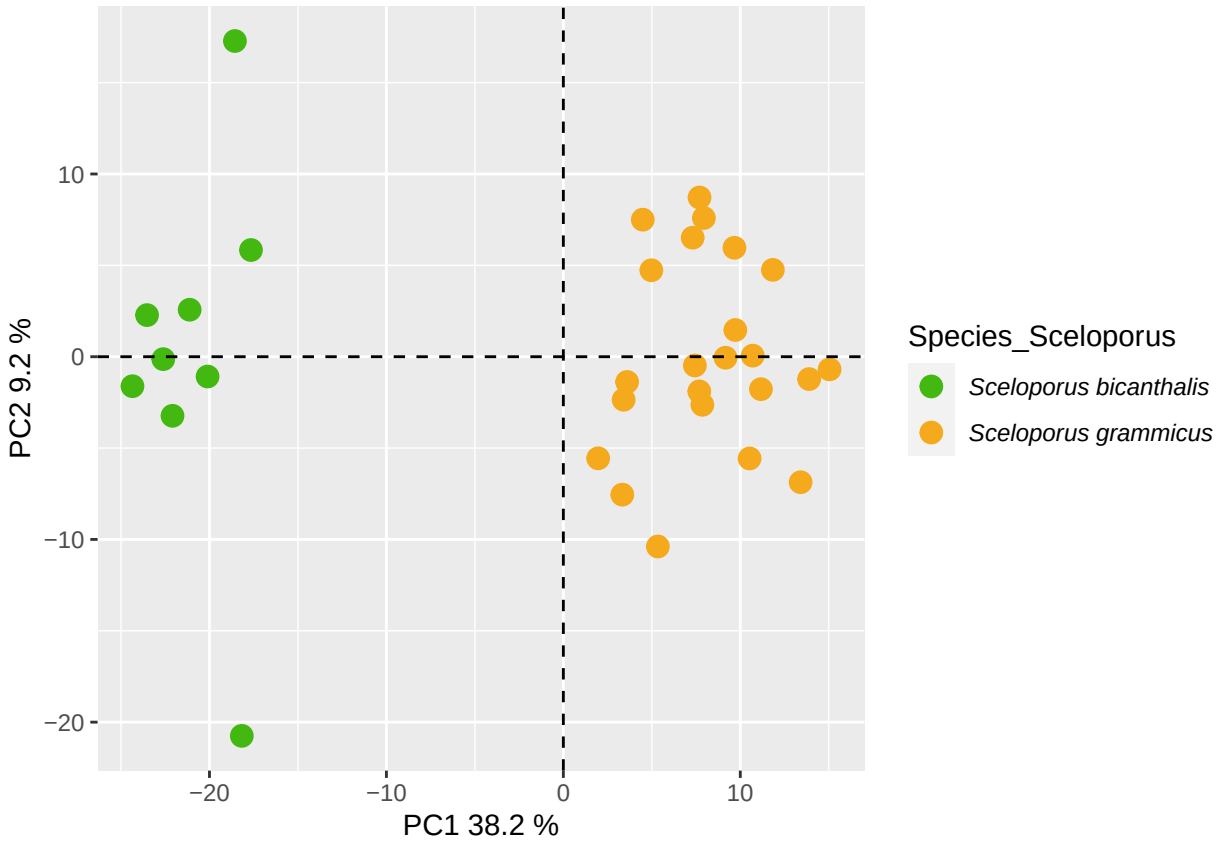



```
#ggsave("PCA_SaSg_aldex.clr.jpeg", width=5.5, height=5.5, dpi=300)

# perMANOVA
Permanova2 <- adonis2(otu_table2 ~ Species_Sceloporus, data = metadata2,
                      method = "euclidean", permutations = 999)
print(Permanova2)
```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = otu_table2 ~ Species_Sceloporus, data = metadata2, permutations = 999, method = "e
##                    Df SumOfSqs      R2      F Pr(>F)
## Species_Sceloporus  1   4396.3 0.20992 11.956  0.001 ***
## Residual           45  16546.2 0.79008
## Total              46  20942.5 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Bicanthalis vs Grammicus

```
# SCELOPORUS BICANTHALIS VERSUS SCELOPORUS GRAMMICUS
otu_table3 <- read.csv(file = "../Data/aldex_SbSg.csv",
                       header = TRUE, row.names = 1, check.names = F)
metadata3 <- read.csv("../Data/metadata_BG.csv", row.names = 1,
                      check.names = F) %>% rownames_to_column(var = "SampleID")

# Run a PCA with codaSeq.clr
pcx.abund.aldex <- prcomp(otu_table3)

# Labels to PCA axis
pc1 <- paste("PC1", round(sum(pcx.abund.aldex$sdev[1] ^2) /
                            mvar(otu_table3) * 100, 1), "%")

pc2 <- paste("PC2", round(sum(pcx.abund.aldex$sdev[2] ^2) /
                            mvar(otu_table3) * 100, 1), "%")

# Create the base plot with only the arrows
PCA_SbSg_aldex.clr <- ggplot() +
  theme_bw() +
  xlab(pc1) +
  ylab(pc2) +
  theme(axis.title = element_text(colour = "black", size = 14),
        legend.text = element_text(size = 14),
        legend.title = element_blank(),
        legend.position = "right") +
  theme_gray() +
  geom_point( #individuals
    data = data.frame(pcx.abund.aldex$x) %>% rownames_to_column(
      var = "SampleID") %>%
      left_join(metadata, by = "SampleID"),
    aes(x=PC1, y=PC2, color = Species_Sceloporus),
    size= 3.5) +
  geom_vline(xintercept = 0, linetype = 2) +   #lines-cross
  geom_hline(yintercept = 0, linetype = 2) +
  scale_color_manual(values = c("#43B811", "#F5A91D")) +
  theme(legend.text = element_text(face = "italic"))# +
print(PCA_SbSg_aldex.clr)
```

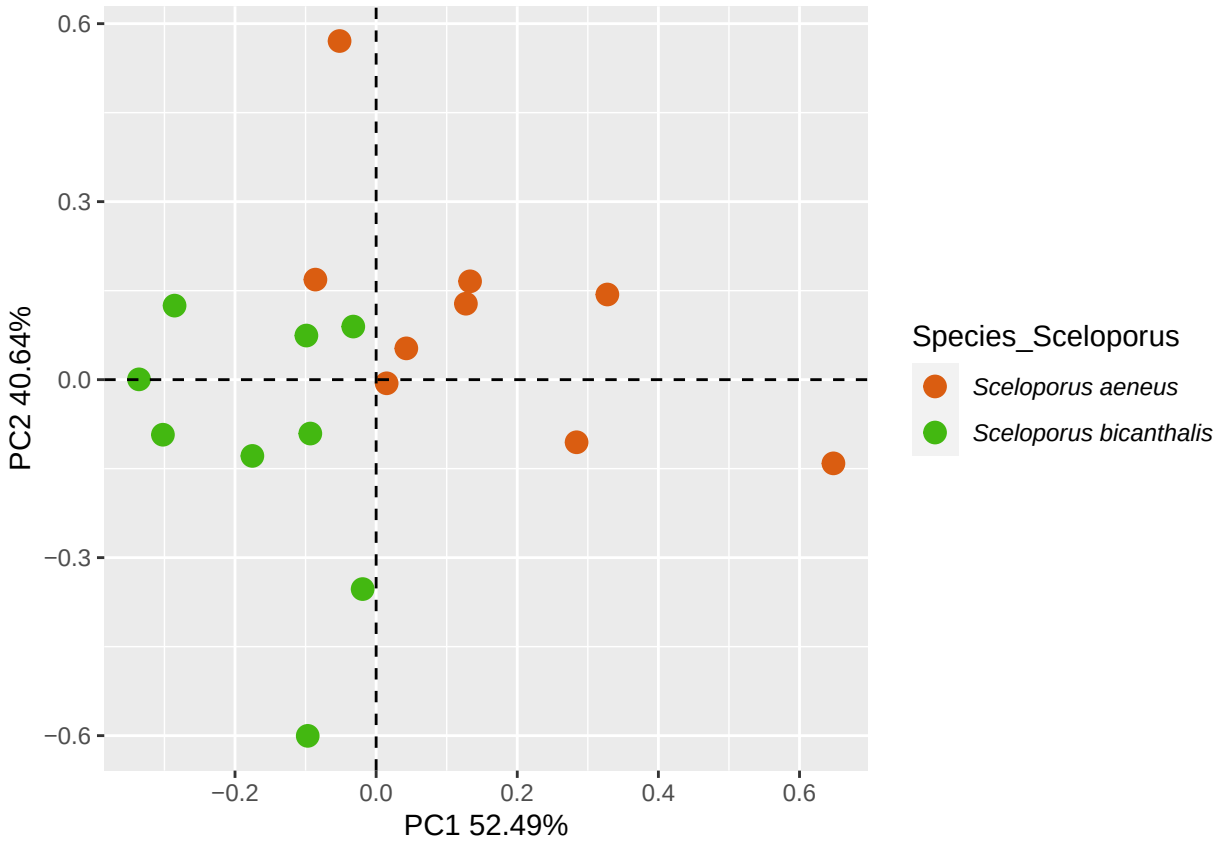



```
# perMANOVA
Permanova3 <- adonis2(otu_table3 ~ Species_Sceloporus, data = metadata3,
                      method = "euclidean", permutations = 999)
print(Permanova3)
```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = otu_table3 ~ Species_Sceloporus, data = metadata3, permutations = 999, method = "eu
##                    Df SumOfSqs      R2      F Pr(>F)
## Species_Sceloporus  1   5497.3 0.36149 16.984  0.001 ***
## Residual           30   9710.2 0.63851
## Total              31  15207.5 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Aeneus vs Bicanthalis (non-core microbiota)

```
metadata <- read.csv("../Data/metadata.csv", row.names = 1,
                     check.names = F) %>% rownames_to_column(var = "SampleID")
PCA <- read.csv("../Data/PCA_Sa_Sb.csv", row.names = 1, check.names = F)
ASVs <- read.csv("../Data/ASVs_PCA123_SaSb.csv", row.names = 1, check.names = F)
taxonomy <- read.csv("../Data/taxonomy.csv", check.names = F) %>% unite(
  taxa, Kingdom:Species, remove = F, sep = ";")

PCA_SaSb_Marzo <- ggplot() +
  theme_bw() +
  xlab("PC1 52.49%") +
  ylab("PC2 40.64%") +
  theme(axis.text = element_text(colour = "black", size = 14),
        axis.title = element_text(colour = "black", size = 14),
        legend.text = element_text(size = 14),
        legend.position = "right") +
  theme_gray() +
  geom_point( #individuals
    data = data.frame(PCA) %>% rownames_to_column(var = "SampleID") %>%
      left_join(metadata, by = "SampleID") %>% dplyr::select(-Tb:-Ts),
    aes(x=PC1, y=PC2, color =Species_Sceloporus),
    size = 3.5) +
  geom_vline(xintercept = 0, linetype = 2) +   #lines-cross
  geom_hline(yintercept = 0, linetype = 2) +
scale_color_manual(values = c("#DA5D11","#43B811")) +
  theme(legend.text = element_text(face = "italic"))
print(PCA_SaSb_Marzo)
```




```
# PERMANOVA SISTER SPECIES (AENEUS VERSUS BICANTHALIS) "Non-core microbiota"

Permanova <- adonis2(PCA ~ Species_Sceloporus, data = metadata,
                     method = "euclidean", permutations = 999)

print(Permanova)
```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = PCA ~ Species_Sceloporus, data = metadata, permutations = 999, method = "euclidean
##                    Df SumOfSqs      R2      F Pr(>F)
## Species_Sceloporus  1  0.67264 0.22421 4.6242  0.002 **
## Residual           16  2.32736 0.77579
## Total              17  3.00000 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# PERMANOVA SISTER SPECIES (AENEUS VERSUS BICANTHALIS) "Core-microbiota"
otu_table4 <- read.csv(file = "../Data/aldex_SaSb.csv",  header = TRUE, row.names = 1,
                       check.names = F)
metadata4 <- read.csv("../Data/metadata_SaSb.csv", row.names = 1, check.names = F) %>%
  rownames_to_column(var = "SampleID")
```

```r
# perMANOVA
Permanova4 <- adonis2(otu_table4 ~ Species_Sceloporus, data = metadata4,
                      method = "euclidean", permutations = 999)
print(Permanova4)
```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = otu_table4 ~ Species_Sceloporus, data = metadata4, permutations = 999, method = "e
##                     Df SumOfSqs      R2      F Pr(>F)
## Species_Sceloporus  1    474.0 0.06461 1.1052  0.322
## Residual           16   6862.4 0.93539
## Total              17   7336.4 1.00000
```

Grammicus populations at low and high altitude

```
# PERMANOVA ALLOPATRIC POPULATION OF SCELOPORUS GRAMMICUS (LOW AND HIGH ZONE)
otutable_Sg <- read.csv("../Data/Sg_feat_table.csv", row.names = 1,
                        check.names = F)
aldez.clr.transform1 <- aldex.clr(otutable_Sg, mc.samples = 999, denom = "all",
                                  verbose = FALSE, useMC = FALSE)
aldex.clr.transform.data <- t(getMonteCarloSample(aldez.clr.transform1, 1))
#write.table(aldex.clr.transform.data, file="./aldex_core_Sg.txt", sep = "\t")

otu_table5 <- read.csv(file = "../Data/aldex_core_Sg.csv",  header = TRUE,
                       row.names = 1, check.names = F)
metadata5 <- read.csv("../Data/metadata_Sg.csv", row.names = 1,
                      check.names = F) %>% rownames_to_column(var = "SampleID")

# perMANOVA
Permanova5 <- adonis2(otu_table5 ~ Species_Sceloporus, data = metadata5,
                      method = "euclidean", permutations = 999)
print(Permanova5)
```

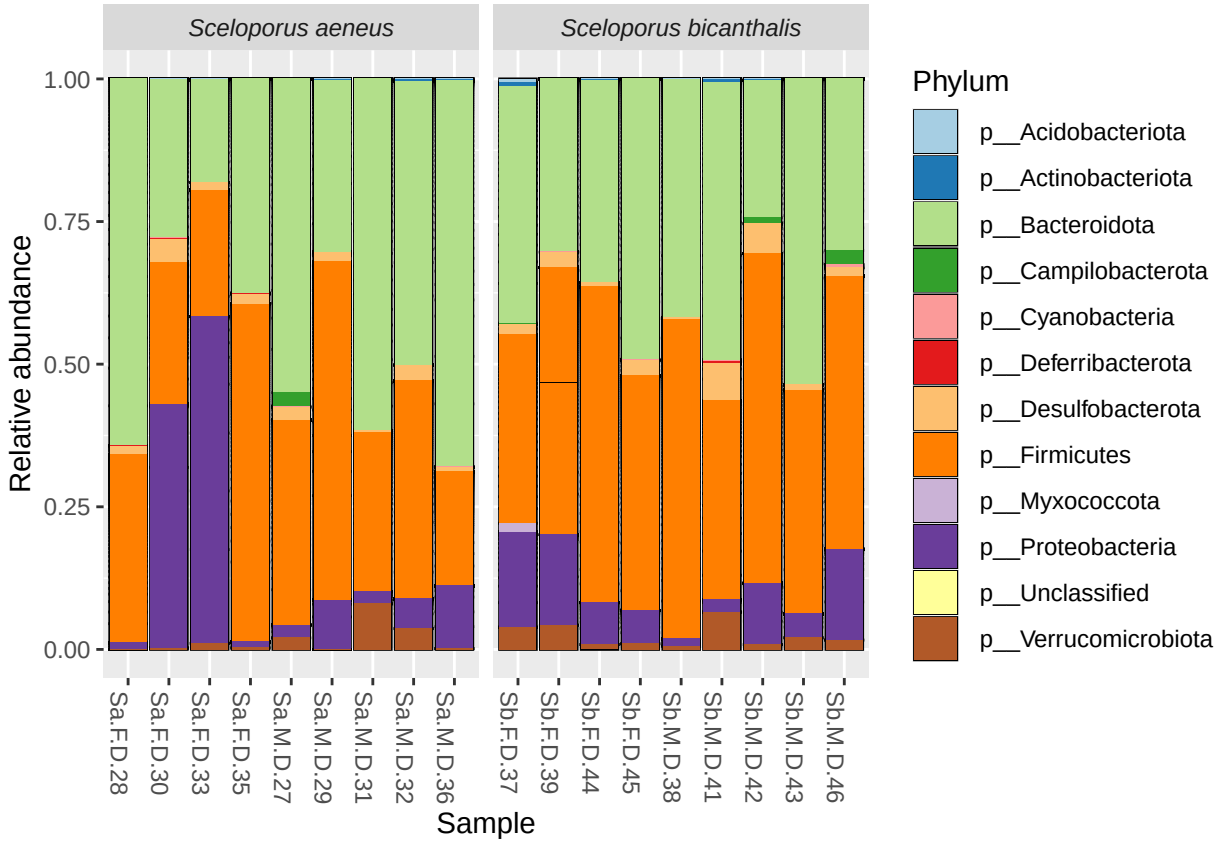
```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = otu_table5 ~ Species_Sceloporus, data = metadata5, permutations = 999, method = "eu
##                     Df SumOfSqs      R2      F Pr(>F)
## Species_Sceloporus  1    118.7 0.02398 1.4494  0.134
## Residual           59   4830.5 0.97602
## Total              60   4949.2 1.00000
```

Barplot Phylum

```
#load packages and files
library(phyloseq)
library(vegan)
library(edgeR)
library(RColorBrewer)
library(scales)
library(viridis)
library(tidyverse)

metadata <- read.csv(file = "../Data/metadata.csv", header = TRUE, row.names = 1)
otu_table <- read.csv("../Data/feature_table2.csv", header = TRUE, row.names = 1)
taxonomy <- read.csv("../Data/taxonomy.csv", header = TRUE, row.names = 1)
#phylo<- read.tree(file = "tree.nwk")

# Crear objeto de categoría phyloseq
SAM <- sample_data(metadata)
TAX <- tax_table(as.matrix(taxonomy))
```

```
OTU <- otu_table(otu_table, taxa_are_rows=TRUE)
#PHY<- phy_tree(phylo)
physeq <- merge_phyloseq(OTU, TAX, SAM)
```

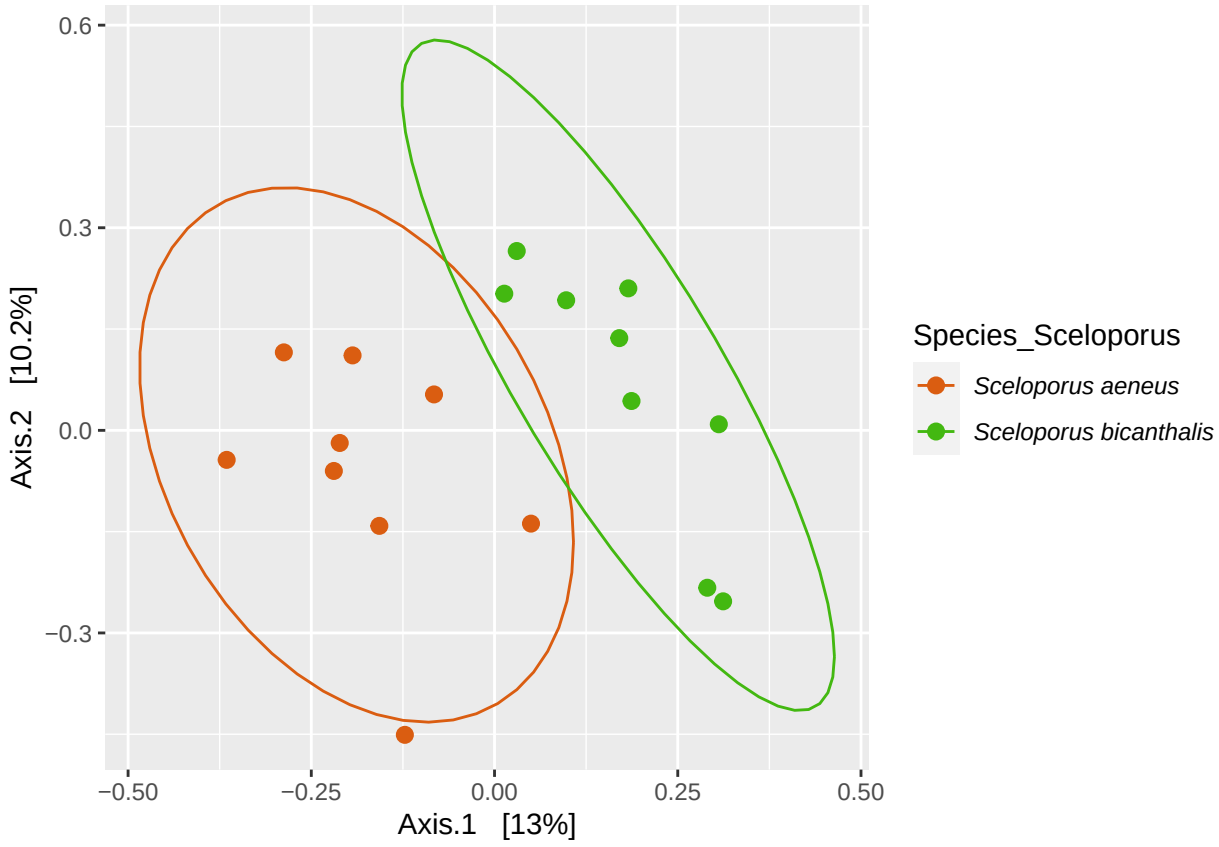
```
# Convert to relative abundance
relative  = transform_sample_counts(physeq = physeq, function(OTU) OTU / sum(OTU))

# Filtering
physeq_sub <- subset_taxa(physeq, !is.na(Kingdom) & !Kingdom %in% c("", "Unassigned"))
physeq_sub <- subset_taxa(physeq, !is.na(Genus) & !Genus %in% c("", "Unassigned"))

# Color palette
paleta <- c(brewer.pal(12, "Paired")[1:12])

## Species
Rel_Abun_Phyl_Species <- plot_bar(physeq = relative, "Sample", fill = "Phylum") +
  facet_grid(~Species_Sceloporus, scales = "free", space = "free") +
  labs(y="Relative abundance") +
  geom_bar(stat = "identity", position="stack", res=300) +
  scale_fill_manual(values = paleta) +
  theme(strip.text.x = element_text(face = "italic")) +
  theme(text = element_text(size = 11))

print(Rel_Abun_Phyl_Species)
```



```
#ggsave("Rel_Abun_Phyl_Species.jpeg", width=8.0, height=4.5, dpi=300)
```

PCoA Bray-Curtis

```
Sceloporus <- ordinate(physeq_sub, "PCoA", "bray")

PCoA_Scel_BrayCuris <- plot_ordination(physeq_sub, Sceloporus, type = "Phylum",
                                       color = "Species_Sceloporus") +
  geom_point(size=2.5) +
  theme_bw() +
  stat_ellipse() +
  theme_gray() +
  scale_color_manual(values = c("#DA5D11","#43B811")) +
  theme(legend.text = element_text(face = "italic"))

print(PCoA_Scel_BrayCuris)
```



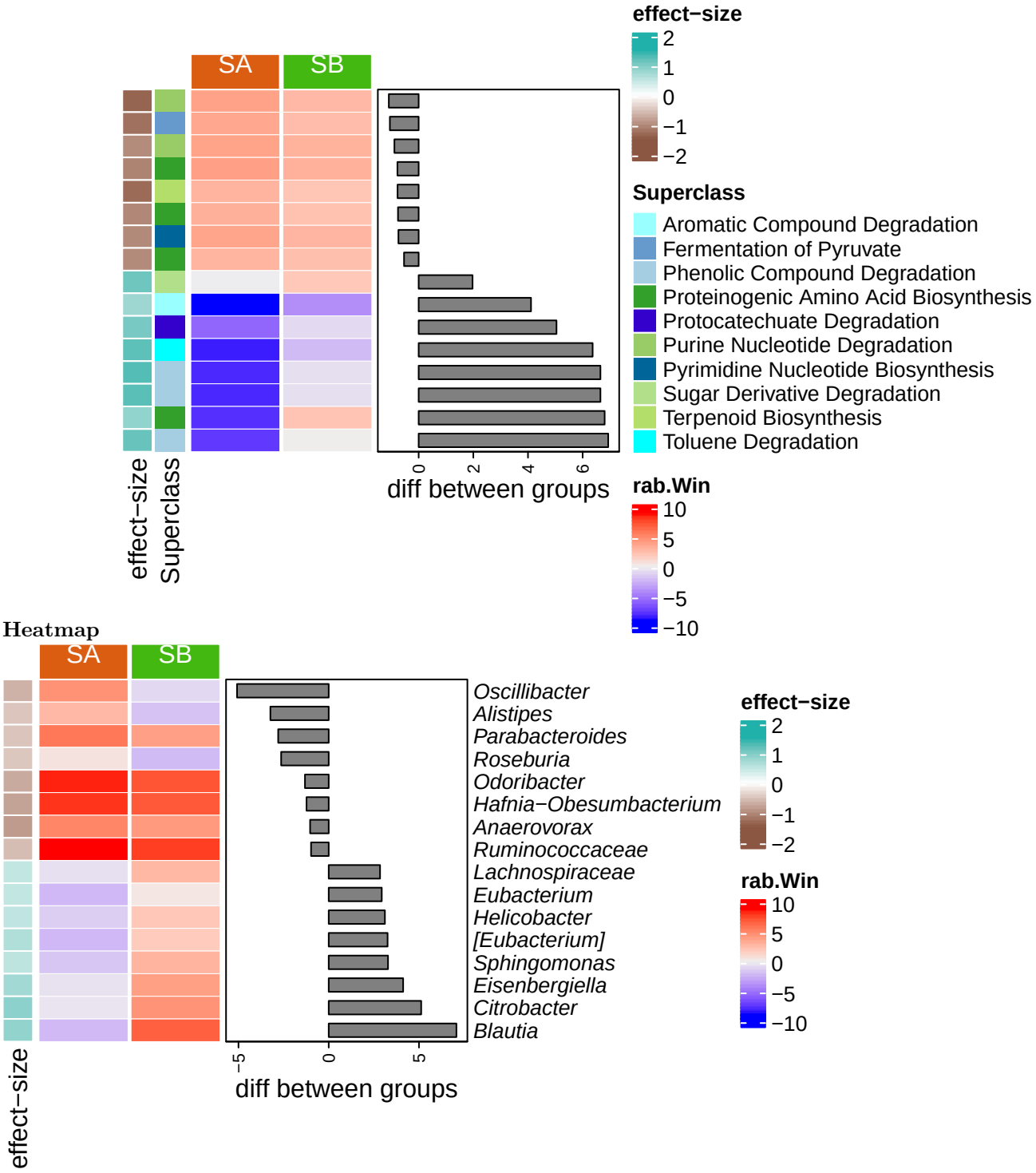

```
#ggsave("PCoA_Scel_BrayCuris.jpeg", width=5.5, height=4.5, dpi=300)
```

ALDEx2 - differential abundances

```
database<- read_tsv("../Data//pathways_abun_by_levels.csv")

ttest_sa_sb_genus_core <- read.delim("../Data/ttest_sa_sb_genus_final_funct_mod.csv")
rownames(ttest_sa_sb_genus_core) <- ttest_sa_sb_genus_core$description

ttest_sa_sb_genus_core2 <- read.delim(
"../Data/ttest_sa_sb_genus_final_mod.csv", row.names=1)

annotation_heatmap <- ttest_sa_sb_genus_core %>%arrange(diff.btw)

annotation_heatmap2 <- ttest_sa_sb_genus_core2%>%  dplyr::select(
  effect, diff.btw) %>% arrange(diff.btw)

data_heatmap <- ttest_sa_sb_genus_core %>%
dplyr::select(rab.win.SA, rab.win.SB, diff.btw) %>% arrange(diff.btw)
color_heatmap= colorRamp2(seq(min(data_heatmap),
                              max(data_heatmap), length = 5), c(
                                "#0000FF","#5499C7", "#DAE7E4", "red", "#FF0000"))

data_heatmap2 <- ttest_sa_sb_genus_core2 %>%
dplyr::select(rab.win.SA, rab.win.SB, diff.btw) %>% arrange(diff.btw)
color_heatmap= colorRamp2(seq(min(data_heatmap),
                              max(data_heatmap), length = 5), c(
                                "#0000FF","#5499C7", "#DAE7E4", "red", "#FF0000"))

#effect annotation
effect_col_fun =colorRamp2(c(-1.5, 0, 1.5), c("lightsalmon4", "white", "lightseagreen"))

ha = HeatmapAnnotation(
  foo = anno_block(gp = gpar(
  fill = c("#800000" ,"#808000" )),
  labels = c("SA", "SB"),
  labels_gp = gpar(col = "white", fontsize = 7, fontface= "bold")))

annEffect = HeatmapAnnotation("effect-size" = annotation_heatmap$effect,
                              which = "row", col = list(
                                "effect-size" = effect_col_fun),
                              show_legend  =T,
                              gp = gpar(col = "white"))
annEffect2 = HeatmapAnnotation("effect-size" = annotation_heatmap2$effect,
                              which = "row", col = list(
                                "effect-size" = effect_col_fun),
                              show_legend  =T,
                              gp = gpar(col = "white"))

#barplot annotation
bardif= rowAnnotation("diff between groups" = anno_barplot(
  annotation_heatmap$diff.btw, width = unit(4, "cm")))

bardif2= rowAnnotation("diff between groups" = anno_barplot(
```

```
  annotation_heatmap2$diff.btw, width = unit(4, "cm")))

labels <- c("Oscillibacter", "Alistipes", "Parabacteroides", "Roseburia",
          "Odoribacter","Hafnia-Obesumbacterium", "Anaerovorax",
          "Ruminococcaceae", "Lachnospiraceae",
          "Eubacterium", "Helicobacter", "[Eubacterium]","Sphingomonas",
          "Eisenbergiella", "Citrobacter", "Blautia")

rownames(data_heatmap2) <- labels

cols_ann <- list('Superclass' = c(
  "Phenolic Compound Degradation"="#A6CEE3",
  "Toluene Degradation"="#00FFFF",
  "Sugar Derivative Degradation"="#B2DF8A",
  "Protocatechuate Degradation"="#3300CC",
  "Proteinogenic Amino Acid Biosynthesis"="#33A02C",
  "Aromatic Compound Degradation"="#99FFFF",
  "Purine Nucleotide Degradation"="#99CC66",
  "Pyrimidine Nucleotide Biosynthesis"="#006699",
  "Fermentation of Pyruvate"="#6699CC",
  "Terpenoid Biosynthesis"="#B3DE69",
  "Carbohydrate Degradation"="#6699FF",
  "Carboxylate Degradation"="#0033CC",
  "Cell Structure Biosynthesis"="#CCEBC5",
  "Cofactor, Carrier, and Vitamin Biosynthesis"="#66FF00",
  "Cofactor, Prosthetic Group, Electron Carrier Degradation"="#00CCFF",
  "Protocatechuate Degradation"="#666699",
  "Fatty Acid and Lipid Biosynthesis"="#66CC33",
  "Fatty Acid and Lipid Degradation"="#000666",
  "Fermentation"="#CC0000",
  "Glycolysis"="#993333",
  "Inorganic Nutrient Metabolism"="#6666FF",
  "Metabolic Regulator Biosynthesis"="#669933",
  "Nucleic Acid Processing"="#FFFF00",
  "Nucleoside and Nucleotide Biosynthesis"="#339933",
  "Nucleoside and Nucleotide Degradation"="#99CCFF",
  "Other"="#000000",
  "Other Biosynthesis"="#069966",
  "Pentose Phosphate Pathways"="#FF6666",
  "Polyprenyl Biosynthesis"="#00FF33",
  "Respiration"="#CC6666",
  "Secondary Metabolite Biosynthesis"="#99CC00",
  "Secondary Metabolite Degradation"="#66CCCC",
  "TCA cycle"="#990033",
  "Protocatechuate Degradation "="#CCFF99"))

colAnn <- HeatmapAnnotation(Superclass = annotation_heatmap$level3,
                            which = 'row',
                            col = cols_ann,
                            show_legend = T)

ht.sa.sb <- ComplexHeatmap::Heatmap(data_heatmap[-3],
                                    width = ncol(data_heatmap)*unit(1, "cm"),
```

```
                                height = ncol(data_heatmap)*unit(2, "cm"),
                                cluster_column_slices = F,
                                cluster_columns = F,
                                column_km = 1,
                                left_annotation =  c(annEffect, colAnn),
                                heatmap_legend_param = list(direction = "vertical"),
                                right_annotation = c(bardif),
                                column_split = rep(c("SA", "SB")),
                                show_heatmap_legend = T,
                                cluster_rows = F,
                                column_title_gp =
                                gpar(fill = c("#DA5D11","#43B811" ),
                                     col="white"),
                                border = F, column_gap = unit(0.5, "mm"),
                                row_dend_side = "left",
                                row_names_side = "right",show_row_names = F,
                                rect_gp = gpar(col = "white", lwd = 0.2),
                                row_names_gp = gpar(fontface ="italic",
                                                    fontsize=10),
                                show_column_names = F, name = "rab.Win")


ht.sa.sb2 <- ComplexHeatmap::Heatmap(data_heatmap2[-3],
                                width = ncol(data_heatmap2)*unit(1, "cm"),
                                height = ncol(data_heatmap2)*unit(2, "cm"),
                                cluster_column_slices = F,
                                cluster_columns = F,
                                column_km = 1,
                                left_annotation =  c(annEffect2),
                                heatmap_legend_param = list(direction = "vertical"),
                                right_annotation = c(bardif2),
                                column_split = rep(c("SA", "SB")),
                                show_heatmap_legend = T,
                                cluster_rows = F,
                                column_title_gp =
                                gpar(fill = c("#DA5D11","#43B811" ),
                                     col="white"),
                                border = F, column_gap = unit(0.5, "mm"),
                                row_dend_side = "left",
                                row_names_side = "right",show_row_names = T ,
                                rect_gp = gpar(col = "white", lwd = 0.2),
                                row_names_gp = gpar(fontface ="italic",
                                                    fontsize=10),
                                show_column_names = F, name = "rab.Win")


ht.sa.sb; ht.sa.sb2
```




```
library(cowplot)

heat <- plot_grid(grid.grabExpr(draw(ht.sa.sb)))
heat2 <- plot_grid(grid.grabExpr(draw(ht.sa.sb2)))
#ggsave(plot = heat, "heat_aldex_sa_sb_taxa.pdf", dpi = 300, width = 10, height = 6 )
```

Barplot functions and taxa

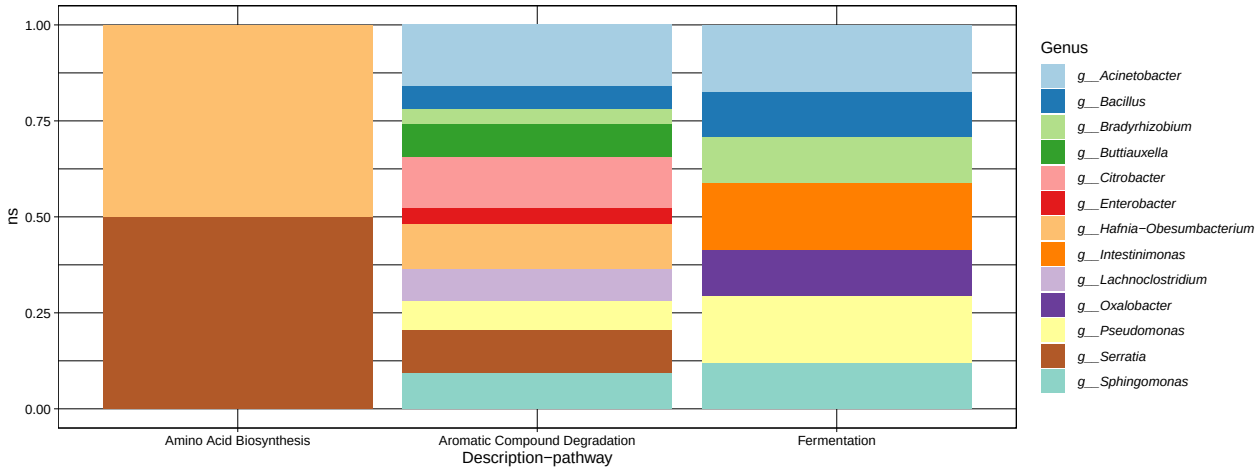
```
#loading files
taxonomy <- read.csv("../Data/taxonomy.csv") %>% rename(ASV=OTUID)
func <- readxl::read_excel("../Data/funciones.xlsx")  %>% replace(is.na(.), "Unassigned")
funciones2 <- func %>% pivot_longer(., cols = EC:EC9,
                                    names_to = "ECs") %>% rename(EC=value)
predic <- read.delim("../Data/EC_TAXA.tsv")
predic_fun<-predic %>% left_join(taxonomy, by = "ASV")

#formatting
dat<-predic_fun %>% group_by(ASV) %>% add_count() %>% dplyr::select(
  EC, ASV, Genus) %>% inner_join(funciones2)
```

```
## Joining, by = "EC"
```

```
dat2 <- dat %>% filter(EC=="EC:1.14.13.127")

dat2 <- dat %>% filter(!Genus=="" & !Genus==" g__uncultured") %>% group_by(
  Genus, `Description-pathway`, Enriched) %>% summarise(ns=n()) %>% arrange(
    `Description-pathway`)
```

```
## `summarise()` has grouped output by 'Genus', 'Description-pathway'. You can
## override using the `.groups` argument.
```

```
paleta <- c(brewer.pal(12, "Paired")[1:12], brewer.pal(12, "Set3")[1:10])

a1 <- dat2 %>% filter(`Description-pathway`==
                      "Aromatic Compound Degradation" & ns>10)
a2 <- dat2 %>% filter(`Description-pathway`==
                      "Fermentation" & ns>1)
a3 <- dat2 %>% filter(`Description-pathway`==
                      "Amino Acid Biosynthesis" & ns>10)

as <- rbind(a1,a2,a3) %>% mutate(log_value=log(ns))

p1 <- as %>%  ggplot(aes(x = `Description-pathway`, y = ns , fill=Genus)) + geom_bar(
  stat = "identity", position = "fill") + theme_linedraw() +
  scale_fill_manual(values = paleta)+theme(legend.text = element_text(face="italic"))

p1
```




```
#ggsave(plot = p1, "plot.jpg", width = 8, height = 4)
```

Venn Diagrams and Net plot

```
#The current script don't show the venn diagram but save the plot in a .tiff file
library(tidyverse)
library(VennDiagram)
library(ggVennDiagram)

# Core microbiota, load files
aeneus_65 <- read.csv("../Data/core_65_aeneus.csv",
                      check.names = F) %>% rownames_to_column(
                        var = "ids")
bicanthalis_65 <- read.csv("../Data/core_65_bicanthalis.csv",
                           check.names = F) %>% rownames_to_column(
                             var = "ids")
grammicus_65 <- read.csv("../Data/core_65_grammicus.csv",
                         check.names = F) %>% rownames_to_column(
                           var = "ids")

# Create Venn Diagramm
venn.plot_60 <- venn.diagram(
  x = list(S_aeneus = aeneus_65$OTUID,
           S_bicanthalis = bicanthalis_65$OTUID,
           S_grammicus = grammicus_65$OTUID),
  category.names = c(
    expression(bold("Sa")),
    expression(bold("Sb")),
    expression(bold("Sg"))),
  filename = "venn_all_species.tiff",
  output = TRUE,
  height = 3000,
  width = 3000,
  resolution = 300,
  compression = "lzw",
```

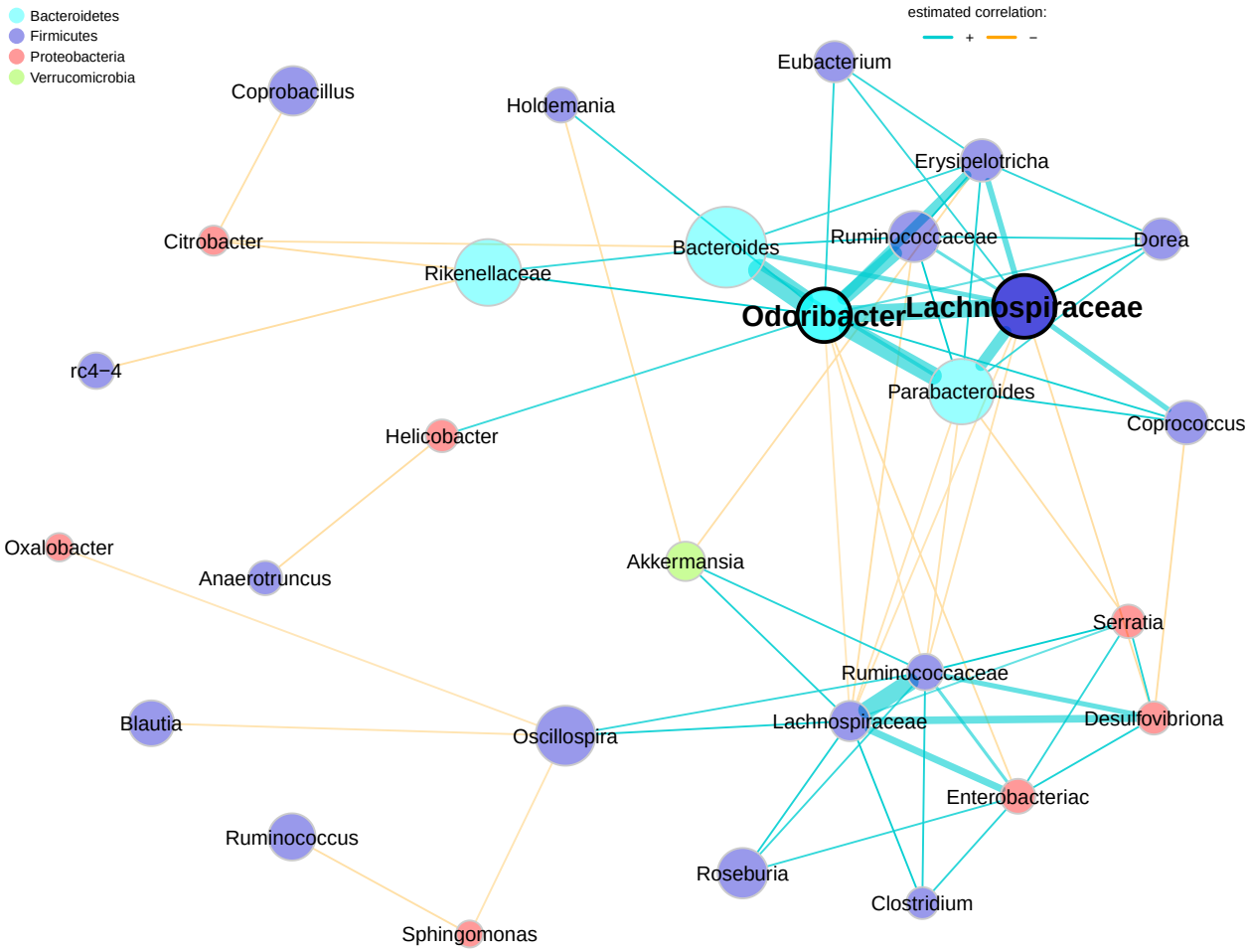
```
  units = "px",
  lwd = 6,
  lty = "blank",
  fill = c("green","red","orange"),
  cex = 1.5,
  fontface = "bold",
  fontfamily = "sans",
  cat.cex = 2,
  cat.fontface = "bold",
  cat.default.pos = "outer",
  # cat.pos = c(115, -125, -50),
  cat.dist = c(0.055, 0.055, 0.055),
  cat.fontfamily = "sans")
```

Net

```
library(tidyverse)
library(phyloseq)
library(NetCoMi)
library(qiime2R)

#loading files and formatting
set.seed(125)
all <- read.delim("../Data/core55_table_L6_gg13_8_97_GBA_mod.csv",
                  check.names = F)

all2 <- all  %>% mutate(ids=paste0("OTU", rownames(.))) %>% column_to_rownames(
  var = "ids") %>% rename(Taxon="#OTU ID")

taxa <- all2 %>% dplyr::select(Taxon) %>%  rownames_to_column(
  var = "Feature.ID")
parse_taxa <- taxa %>% separate(Taxon, c("Kingdom", "Phylum", "Class", "Order", "Family", "Genus"), sep

otu <- all2 %>% dplyr::select(-Taxon)
tax2 <- parse_taxa[match(rownames(otu), rownames(parse_taxa)),] %>% as.matrix()

TAX <- tax_table(tax2)
OTU <- otu_table(otu, taxa_are_rows = T)
ps2 <- phyloseq(OTU, TAX)

amgut_genus <- phyloseq::tax_glom(ps2, taxrank = "Genus")
taxtab <- amgut_genus@tax_table@.Data

# Find undefined taxa (in this data set, unknowns occur only up to Family)
miss_f <- which(taxtab[, "Family"] == "f__")
miss_g <- which(taxtab[, "Genus"] == "g__")

# Number unspecified genera
taxtab[miss_f, "Family"] <- paste0("f__", 1:length(miss_f))
taxtab[miss_g, "Genus"] <- paste0("g__", 1:length(miss_g))
```

```
# Find duplicate genera
dupl_g <- which(duplicated(taxtab[, "Genus"]) |
                  duplicated(taxtab[, "Genus"], fromLast = TRUE))

for(i in seq_along(taxtab)){
  # The next higher non-missing rank is assigned to unspecified genera
  if(i %in% miss_f && i %in% miss_g){
    taxtab[i, "Genus"] <- paste0(taxtab[i, "Genus"],
                                 "(", taxtab[i, "Order"], ")")
  } else if(i %in% miss_g){
    taxtab[i, "Genus"] <- paste0(taxtab[i, "Genus"],
                                 "(", taxtab[i, "Family"], ")")
  }

  # Family names are added to duplicate genera
  if(i %in% dupl_g){
    taxtab[i, "Genus"] <- paste0(taxtab[i, "Genus"], "(", taxtab[i, "Family"], ")")
  }
}

amgut_genus@tax_table@.Data <- taxtab
rownames(amgut_genus@otu_table@.Data) <- taxtab[, "Genus"]

# Get phyla names from the taxonomic table created before

phyla <- as.factor(gsub("p__", "", taxtab[, "Phylum"]))
names(phyla) <- taxtab[, "Genus"]

# Network construction and analysis

net_single3 <- netConstruct(amgut_genus,
                            zeroMethod = "pseudo",
                            sparsMethod = "threshold",
                            thresh=0.3,
                            zeroPar = list(pseudocount = 0.5),
                            normMethod = "clr",
                            measure = "sparcc",
                            verbose = 0)


props_single3 <- netAnalyze(net_single3, clustMethod = "cluster_fast_greedy")

# Compute layout
graph3 <- igraph::graph_from_adjacency_matrix(net_single3$adjaMat1,
                                              weighted = TRUE)
lay_fr <- igraph::layout_with_fr(graph3)

# Note that row names of the layout matrix must match the node names
rownames(lay_fr) <- rownames(net_single3$adjaMat1)
library(viridis)
col2 <- viridis_pal(option = "H", begin = 0.1, end = 0.9)((12))

# Define phylum colors
```

```r
phylcol <- c("cyan", "blue3", "red", "lawngreen", "yellow", "deeppink","green")

#png(filename= "phyla_red_generos_sa_sb.png", width=1648, height=832)

plot(props_single3,
     layout = "spring",
     repulsion = 0.9,
     labelLength = 15,
    # shortenLabels = "none",
     charToRm = "g__",
     labelScale = FALSE,
     rmSingles = TRUE,
     nodeSize = "clr",
     nodeSizeSpread = 4,
     nodeColor = "feature",
     featVecCol = phyla,
     colorVec =  phylcol,
     posCol = "darkturquoise",
     negCol = "orange",
     edgeTranspLow = 0,
     edgeTranspHigh = 40,
     cexNodes = 2,
     cexLabels = 1.2,
     cexHubLabels = 1.7,
     #title1 = "Network on genus level showing phyla with sparcc",
     showTitle = FALSE,
     cexTitle = 2.3)

# Colors used in the legend should be equally transparent as in the plot
phylcol_transp <- NetCoMi:::colToTransp(phylcol, 60)

legend(-1.1, 1.2, cex = 0.8, pt.cex = 2, title = "",
       legend=levels(phyla), col = phylcol_transp, bty = "n", pch = 16)

legend(0.5, 1.15, cex = 0.8, title = "estimated correlation:",
       legend = c("+","-"), lty = 1, lwd = 3, col = c("darkturquoise","orange"),
       bty = "n", horiz = TRUE)
```




```
#dev.off()

#statistics of net
set.seed(122)
summary(props_single3)
```

```
##
## Component sizes
## ‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘
## size: 29 1
##    #:  1 1
## ______________________________
## Global network properties
## ‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘
## Largest connected component (LCC):
##
## Relative LCC size         0.96667
## Clustering coefficient    0.53876
## Moduarity                 0.32147
## Positive edge percentage 69.44444
## Edge density              0.17734
## Natural connectivity      0.07397
## Vertex connectivity       1.00000
```

```
## Edge connectivity         1.00000
## Average dissimilarity*    0.92856
## Average path length**     1.81289
##
## Whole network:
##
## Number of components      2.00000
## Clustering coefficient    0.53876
## Moduarity                 0.32147
## Positive edge percentage 69.44444
## Edge density              0.16552
## Natural connectivity      0.07054
##
##  *Dissimilarity = 1 - edge weight
## **Path length: Units with average dissimilarity
##
## ------------------------------
## Clusters
## - In the whole network
## - Algorithm: cluster_fast_greedy
## ‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘
##
## name: 0  1  2 3 4 5
##    #: 1 12 11 2 2 2
##
## ------------------------------
## Hubs
## - In alphabetical/numerical order
## - Based on empirical quantiles of centralities
## ‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘
##  g__Lachnospiraceae2
##  g__Odoribacter
##
## ------------------------------
## Centrality measures
## - In decreasing order
## - Centrality of disconnected components is zero
## ‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘
## Degree (normalized):
##
## g__Odoribacter         0.48276
## g__Lachnospiraceae     0.41379
## g__Lachnospiraceae2    0.37931
## g__Ruminococcaceae     0.37931
## g__Parabacteroides     0.34483
## g__Erysipelotrichaceae 0.27586
## g__Bacteroides         0.24138
## g__Ruminococcaceae2    0.24138
## g__Enterobacteriaceae  0.24138
## g__Desulfovibrionaceae 0.20690
##
## Betweenness centrality (normalized):
##
## g__Odoribacter      0.48942
```

```
## g__Lachnospiraceae  0.41005
## g__Oscillospira     0.26720
## g__Bacteroides      0.12698
## g__Rikenellaceae    0.07143
## g__Sphingomonas     0.07143
## g__Helicobacter     0.07143
## g__Citrobacter      0.07143
## g__Lachnospiraceae2 0.04497
## g__Parabacteroides  0.03704
##
## Closeness centrality (normalized):
##
## g__Odoribacter         1.30453
## g__Lachnospiraceae2    1.13311
## g__Lachnospiraceae     1.08940
## g__Parabacteroides     1.07652
## g__Ruminococcaceae     1.04809
## g__Bacteroides         1.02331
## g__Ruminococcaceae2    0.99499
## g__Erysipelotrichaceae 0.95624
## g__Enterobacteriaceae  0.91555
## g__Desulfovibrionaceae 0.86718
##
## Eigenvector centrality (normalized):
##
## g__Odoribacter         1.00000
## g__Lachnospiraceae2    0.88595
## g__Parabacteroides     0.81436
## g__Ruminococcaceae2    0.71954
## g__Erysipelotrichaceae 0.71364
## g__Bacteroides         0.67645
## g__Dorea               0.54038
## g__Lachnospiraceae     0.41753
## g__Coprococcus         0.40273
## g__Ruminococcaceae     0.38490
```

Prevalence

```
preval <- readxl::read_excel("../Data/prevalence.xlsx", sheet = "Data_graphic2022") %>%
  rename("S. aeneus"=PrevalenciaSA, "S. bicanthalis"=PrevalenciaSB, "Sceloporusspp."=PrevalenciaTot)

library(tidyverse)
preval_dat<-preval %>% pivot_longer(
  ., cols = c(`S. aeneus`, `S. bicanthalis`, `Sceloporusspp.`),
  names_to = "prev", values_to = "prev_val")%>% pivot_longer(
    ., cols = c(Rel.ab.SA, Rel.ab.SB, Rel.ab.Tot),
 names_to = "relab", values_to = "relab_val") %>% drop_na()

colors <- c("darkorange1", "chartreuse4", "black")

#$preval_dat %>% ggplot(aes(prev_val, relab_val, color= prev_val))+geom_point()
```

```
preval %>% ggplot() + geom_point(aes(`S. aeneus`, Rel.ab.SA,color="S. aeneus"), size=3) +
  geom_point(aes(`S. bicanthalis`, Rel.ab.SB, color="S. bicanthalis"), size=3) +
  geom_point(aes(`Sceloporusspp.`, Rel.ab.Tot,color="Sceloporus spp."), size=3) +
  scale_color_manual(values = colors) +
  geom_text(aes(`S. aeneus`, Rel.ab.SA,label=labels, fontface="italic")) +
  geom_text(aes(`S. bicanthalis`, Rel.ab.SB,label=labels,fontface="italic")) +
  ylab("Average relative abundance") +
  xlab("Prevalence (%)")+ theme(
    legend.title = element_blank(),
    legend.text = element_text(size = 12, face = "italic"),
    legend.position = c(0.1,0.9),
    axis.text = element_text(size = 12, colour = "black"))
```