Alpha diversity

Sterling Wright

2023-11-01

```r
library(formatR)
library(knitr)
knitr::opts_chunk$set(tidy.opts = list(width.cutoff = 60), tidy = TRUE)
```

Load the necessary libraries

```r
set.seed(1024)
library(vegan)
library(phyloseq)
library(ggplot2)
library(gridExtra)
library(grid)
library(ggpubr)
```

Load the data from where you have saved the files.

```r
knitr::opts_chunk$set(tidy.opts = list(width.cutoff = 600), tidy = TRUE)
Hungary.species.data <- import_biom("Hungary-species-samples.qza.biom")
metadata <- import_qiime_sample_data("../Library_METADATA.txt")
Hungary.species_phylo <- merge_phyloseq(Hungary.species.data,
    metadata)

plot1 <- plot_richness(Hungary.species_phylo, x = "LibraryMethod",
    measures = c("Observed", "Chao1", "Shannon")) + geom_boxplot() +
    theme_classic() + theme(strip.background = element_blank(),
    axis.text.x.bottom = element_text(angle = -90))
```

```
## Warning in estimate_richness(physeq, split = TRUE, measures = measures): The data you have provided
## any singletons. This is highly suspicious. Results of richness
## estimates (for example) are probably unreliable, or wrong, if you have already
## trimmed low-abundance taxa from the data.
##
## We recommended that you find the un-trimmed data and retry.
```

```r
plot2 <- plot_richness(Hungary.species_phylo, x = "ExtractionType",
    measures = c("Observed", "Chao1", "Shannon")) + geom_boxplot() +
    theme_classic() + theme(strip.background = element_blank(),
    axis.text.x.bottom = element_text(angle = -90))
```

```
## Warning in estimate_richness(physeq, split = TRUE, measures = measures): The data you have provided
## any singletons. This is highly suspicious. Results of richness
## estimates (for example) are probably unreliable, or wrong, if you have already
## trimmed low-abundance taxa from the data.
##
## We recommended that you find the un-trimmed data and retry.
```

```
Gobero.species.data <- import_biom("Gobero-species-samples.qza.biom")
metadata <- import_qiime_sample_data("../Library_METADATA.txt")
species_phylo <- merge_phyloseq(Gobero.species.data, metadata)

plot3 <- plot_richness(species_phylo, x = "LibraryMethod", measures = c("Observed",
    "Chao1", "Shannon")) + geom_boxplot() + theme_classic() +
    # scale_fill_manual(values = c('SingleStranded' =
    # 'blue', 'DoubleStranded' = 'orange')) +
theme(strip.background = element_blank(), axis.text.x.bottom = element_text(angle = -90))
```

```
## Warning in estimate_richness(physeq, split = TRUE, measures = measures): The data you have provided
## any singletons. This is highly suspicious. Results of richness
## estimates (for example) are probably unreliable, or wrong, if you have already
## trimmed low-abundance taxa from the data.
##
## We recommended that you find the un-trimmed data and retry.
```

```
plot4 <- plot_richness(species_phylo, x = "ExtractionType", measures = c("Observed",
    "Chao1", "Shannon")) + geom_boxplot() + theme_classic() +
    theme(strip.background = element_blank(), axis.text.x.bottom = element_text(angle = -90))
```

```
## Warning in estimate_richness(physeq, split = TRUE, measures = measures): The data you have provided
## any singletons. This is highly suspicious. Results of richness
## estimates (for example) are probably unreliable, or wrong, if you have already
## trimmed low-abundance taxa from the data.
##
## We recommended that you find the un-trimmed data and retry.
```

Box plots

Species level plots

Including Plots

```
ggarrange(plot4, plot3, plot2, plot1, labels = c("A", "B", "C", "D"), ncol = 2, nrow = 2)
```

```
Gobero.genus.data <- import_biom("Gobero-genus-samples.qza.biom")
genus_phylo <- merge_phyloseq(Gobero.genus.data, metadata)

plot1 <- plot_richness(genus_phylo, x = "LibraryMethod", measures = c("Observed",
    "Chao1", "Shannon")) + geom_boxplot() + theme_classic() +
    theme(strip.background = element_blank(), axis.text.x.bottom = element_text(angle = -90))
```

```
## Warning in estimate_richness(physeq, split = TRUE, measures = measures): The data you have provided
## any singletons. This is highly suspicious. Results of richness
## estimates (for example) are probably unreliable, or wrong, if you have already
## trimmed low-abundance taxa from the data.
##
## We recommended that you find the un-trimmed data and retry.
```

```
plot2 <- plot_richness(genus_phylo, x = "ExtractionType", measures = c("Observed",
    "Chao1", "Shannon")) + geom_boxplot() + theme_classic() +
    theme(strip.background = element_blank(), axis.text.x.bottom = element_text(angle = -90))
```

```
## Warning in estimate_richness(physeq, split = TRUE, measures = measures): The data you have provided
```
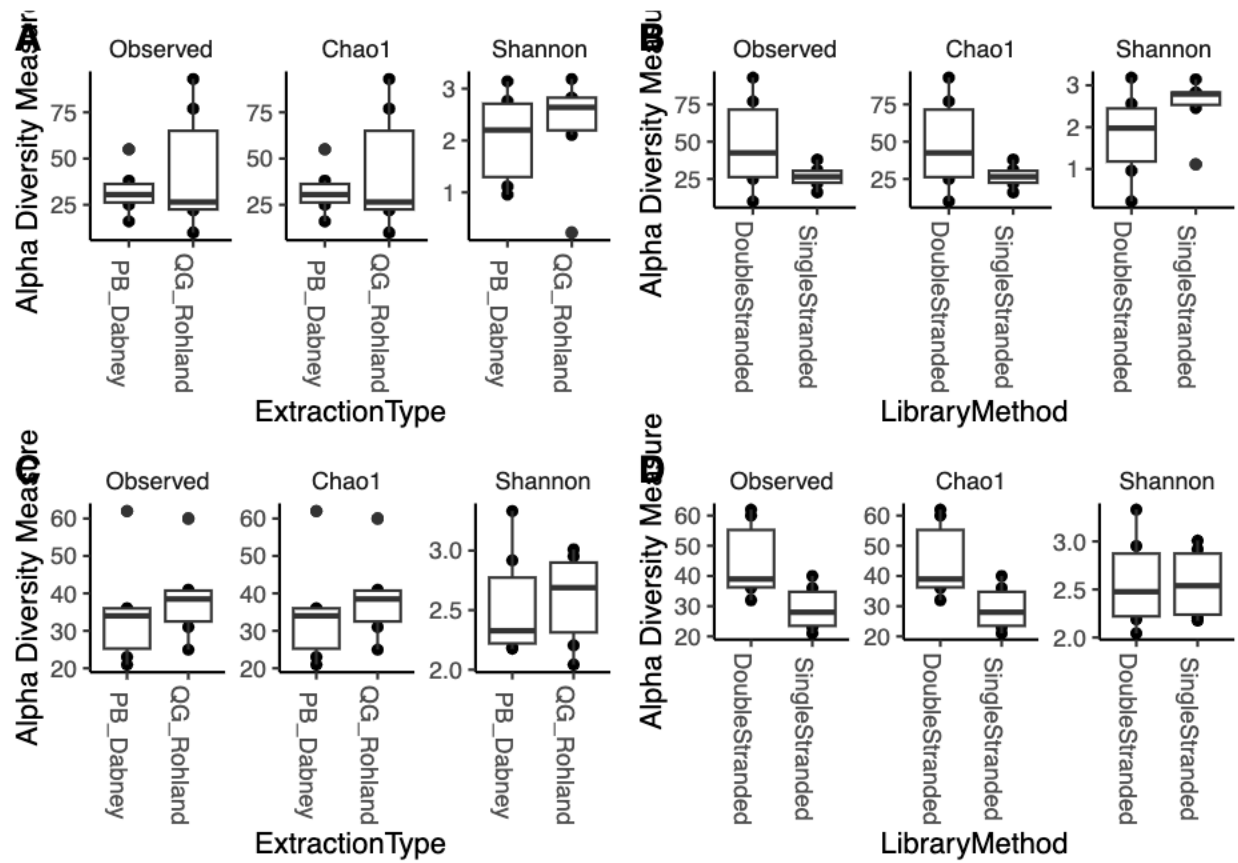
Figure 1: A + B = Hungary, C + D = Niger

```
## any singletons. This is highly suspicious. Results of richness
## estimates (for example) are probably unreliable, or wrong, if you have already
## trimmed low-abundance taxa from the data.
##
## We recommended that you find the un-trimmed data and retry.
```

```r
# Hungary
Hungary.genus.data <- import_biom("Hungary-genus-samples.qza.biom")
Hungary.genus_phylo <- merge_phyloseq(Hungary.genus.data, metadata)

plot3 <- plot_richness(Hungary.genus_phylo, x = "LibraryMethod",
    measures = c("Observed", "Chao1", "Shannon")) + geom_boxplot() +
    theme_classic() + theme(strip.background = element_blank(),
    axis.text.x.bottom = element_text(angle = -90))
```

```
## Warning in estimate_richness(physeq, split = TRUE, measures = measures): The data you have provided
## any singletons. This is highly suspicious. Results of richness
## estimates (for example) are probably unreliable, or wrong, if you have already
## trimmed low-abundance taxa from the data.
##
## We recommended that you find the un-trimmed data and retry.
```

```r
plot4 <- plot_richness(Hungary.genus_phylo, x = "ExtractionType",
    measures = c("Observed", "Chao1", "Shannon")) + geom_boxplot() +
    theme_classic() + theme(strip.background = element_blank(),
    axis.text.x.bottom = element_text(angle = -90))
```

```
## Warning in estimate_richness(physeq, split = TRUE, measures = measures): The data you have provided
## any singletons. This is highly suspicious. Results of richness
## estimates (for example) are probably unreliable, or wrong, if you have already
## trimmed low-abundance taxa from the data.
##
## We recommended that you find the un-trimmed data and retry.
```

```r
ggarrange(plot1, plot2, plot3, plot4, labels = c("A", "B", "C",
    "D"), ncol = 2, nrow = 2)
```