

phyloseq Testing

Ella Crotty

Contents

Lesson Link	1
Setup	1
Phyloseq Plotting Functions	2
Phyloseq Diversity Functions	2
ANOVA	3
Saving plots	3

Lesson Link

- [Phyloseq web page](#)
- [Another phyloseq tutorial](#)

Setup

```
library('phyloseq')
library('ggplot2')

## Warning: package 'ggplot2' was built under R version 4.3.1

library('ape')

## Warning: package 'ape' was built under R version 4.3.1

library("vegan")

## Warning: package 'vegan' was built under R version 4.3.3
## Loading required package: permute
## Loading required package: lattice
## Warning: package 'lattice' was built under R version 4.3.1
## This is vegan 2.6-6.1

library("gridExtra")
```

I cannot find the example data used for this tutorial, so I will just take notes for now.

```

physeq <- readRDS('fish_phyloseq.rds')
print(physeq)

# Rarefied
physeq.rarefied <- readRDS('fish_phyloseq_rarefied.rds')
print(physeq.rarefied)

```

Phyloseq Plotting Functions

- `plot_tree()` takes a phyloseq object and makes a phylogenetic tree out of it
 - `color = "sample"` will add colored dots to the ends of the branches. Any metadata/sample data can be used to color.
 - `label.tips = "taxa_names"` will add taxa names to the end of the branches, this can be used alongside colors
- `plot_bar()` plots the relative abundance of taxa across samples (ex. `plot_bar(phyloseq object, fill = "genus")`)
 - Using the rarefied dataset in this function can help compare samples better
 - To collapse across metadata categories, use `x`: `plot_bar(physeq.rarefied, x='location', fill='genus')`. This example code will lead to one bar per location with total abundances, instead of one bar per sample.
 - You can save this plot as a variable and add ggplot customization
 - To make these graphs less busy, start with `sub <- subset_taxa(phyloseq object, subsetting)` and then use `plot_bar(sub, fill = "smaller taxonomic rank")`. Subsetting looks like `order == "Scombriformes"`, etc. `sub` is now a smaller phyloseq object that can be plotted.

Phyloseq Diversity Functions

Ordination: Taking a phyloseq object and makes it into an ordinated object recognized by downstream phyloseq functions. I'm still not totally sure what it does to the object.

- 2 ways to calculate distance: One is qualitative (presence/absence) and one is quantitative (includes abundance)
 - Qualitative: `jac_dist <- distance(physeq.rarefied, method = "jaccard", binary = TRUE)`
 - Quantitative: `bc_dist <- distance(physeq.rarefied, method = "bray", binary = FALSE)`
- Ordinated & plot distance data:
 - Jaccard: `qual_ord <- ordinate(physeq.rarefied, method="PCoA", distance=jac_dist), plot_jac <- plot_ordination(physeq.rarefied, qual_ord, color="location", title='Jaccard') + theme(aspect.ratio=1) + geom_point(size=4)`
 - Bray-Curtis: `quant_ord <- ordinate(physeq.rarefied, method="PCoA", distance=bc_dist), plot_bc <- plot_ordination(physeq.rarefied, quant_ord, color="location", title='Bray-Curtis') + theme(aspect.ratio=1) + geom_point(size=4)`
- Unifrac is a distance measure incorporating phylogeny. It can be weighted (quantitative) or unweighted (qualitative).
 - Unweighted: `uni_dist <- distance(physeq.rarefied, method= "unifrac")`
 - Weighted: `wuni_dist <- distance(physeq.rarefied, method= "wunifrac")`
- Ordinate and plot Unifrac data:
 - Ordinate, then plot with `plot_ordination()`

```

uni_ord <- ordinate(physeq.rarefied,
                    method="PCoA",
                    distance = uni_dist)

```

```

wuni_ord <- ordinate(physeq.rarefied,
                    method="PCoA",
                    distance = wuni_dist)
plot_uni <- plot_ordination(physeq.rarefied,
                           wuni_ord,
                           color="location",
                           title="Unweighted Unifrac") +
  theme(aspect.ratio=1) +
  geom_point(size=4)
plot_uni

plot_wuni <- plot_ordination(physeq.rarefied,
                            wuni_ord,
                            color="location",
                            title="Weighted Unifrac") +
  theme(aspect.ratio=1) +
  geom_point(size=4)
plot_wuni
- This will return a plot that looks a lot like a PCA plot. Is this a PCA plot? I'm not sure.
- Other characteristics can also be displayed with plot ordination, like: plot_ordination(physeq.rarefied,
  wuni_ord, color="location", shape='temperature', title="Weighted Unifrac") +
  geom_point()

```

ANOVA

- The `adonis` function runs a PERMANOVA to compare OTUs (Operational Taxonomic Units) between metadata fields
- `adonis(wuni_dist ~ sample_data(physeq.rarefied)$location)` will return a table of statistics ending in a p-value

Saving plots

```

plot_bc <- plot_ordination(physeq.rarefied, quant_ord, color="location", title='Bray-Curtis') + theme(a

pdf('ordination_plot.pdf')
## enter saved plot
plot_bc
## close the file
dev.off()

```