

Alpha and Beta Diversity Analysis Tutorial PART 1

Ali Youncha

2024-06-27

Introduction

This tutorial demonstrates how to perform alpha and beta diversity analyses using the **phyloseq** package in R. We will use OTU and TAX tables generated from the MicrobioteRNAPipe pipeline.

First, we install and load the required libraries.

```
knitr::opts_chunk$set(echo = TRUE)
library(phyloseq)
library(ggplot2)
library(vegan)
```

```
## Le chargement a nécessité le package : permute
```

```
## Le chargement a nécessité le package : lattice
```

```
## This is vegan 2.6-4
```

```
library(openxlsx)
library(dplyr)
```

```
##
```

```
## Attachement du package : 'dplyr'
```

```
## Les objets suivants sont masqués depuis 'package:stats':
```

```
##
```

```
##      filter, lag
```

```
## Les objets suivants sont masqués depuis 'package:base':
```

```
##
```

```
##      intersect, setdiff, setequal, union
```

```
library(stringr)
```

Next, we load the sample data from an Excel file and the OTU and TAX tables.

```

otu_table_path <- "C:/Users/inserm/Documents/Ali/diversity_analysis/otu_table_final.csv"
## Path for the OTU table generated from the pipeline
tax_table_path <- "C:/Users/inserm/Documents/Ali/diversity_analysis/tax_table_s339.csv"
## Path for the TAX table generated from the pipeline

sample_data_path <- "C:/Users/inserm/Downloads/samples.xlsx"
## Samples excel file containing all the samples

clin <- read.xlsx(sample_data_path, rowNames = TRUE)

```

Source function is then used to directly call the script that performs the desired analysis.

```

source('C:/Users/inserm/Documents/Ali/diversity_analysis/functions.R')

```

Then, we call process_data function which requires as an input both OTU and TAX tables, as well as the samples excel file. Please make sure that the names of the samples used for your analysis MATCHES with those on the samples.xlsx file, otherwise the function will not work!

Note that: Alpha diversity analysis is able to describe the species diversity (richness) within a functional community, and that Beta diversity analysis is calculated to measure the variation/dissimilarity between different ecosystems or environments

```

CreatePhyloSeqObject <- process_data(otu_table_path, tax_table_path, sample_data_path)

## This function transforms the OTU and TAX table files to a matrix, matches the
## names of the samples file with those in the analysis, then creates a phyloseq
## object and measures alpha and beta diversity values.

```

Then, from the phyloseq object, the OTU and TAX table are then extracted and defined as a dataframe for data visualisation and for future use.

Note that the CreatePhyloSeqObject uses a taxglom function that agglomerates the analysis into the “Genus” taxonomic level.

```

ps_s_genus <- CreatePhyloSeqObject$ps_s_genus

# Extract tax_table and otu_table from the phyloseq object
tax_table_df <- as.data.frame(tax_table(ps_s_genus))
otu_table_df <- as.data.frame(otu_table(ps_s_genus))

```

Using the alpha diversity plotting function to plot the obtained results. This function requires as input the alpha_div variable (which corresponds to the Alpha diversity value calculated for each sample in the CreatePhyloSeqObject function)

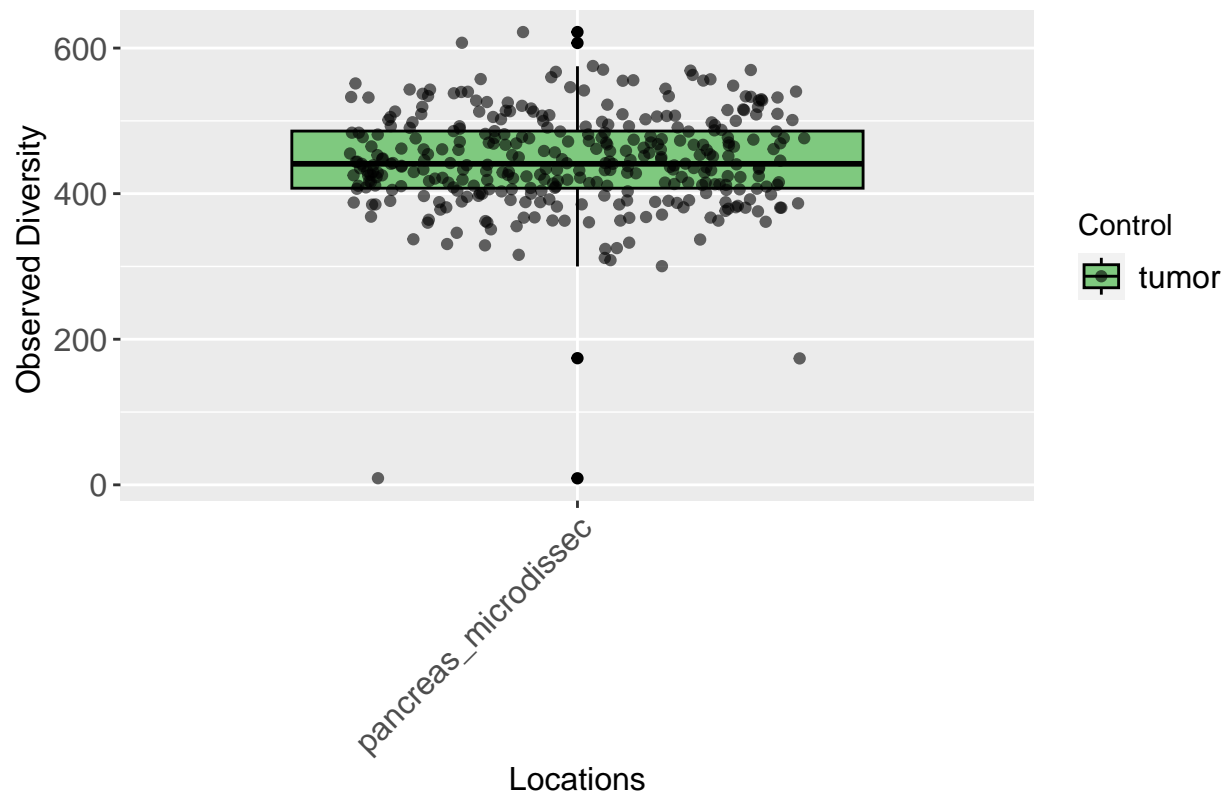
```

alpha_plot <- generate_alpha_diversity_boxplot(CreatePhyloSeqObject$alpha_div)

print(alpha_plot)

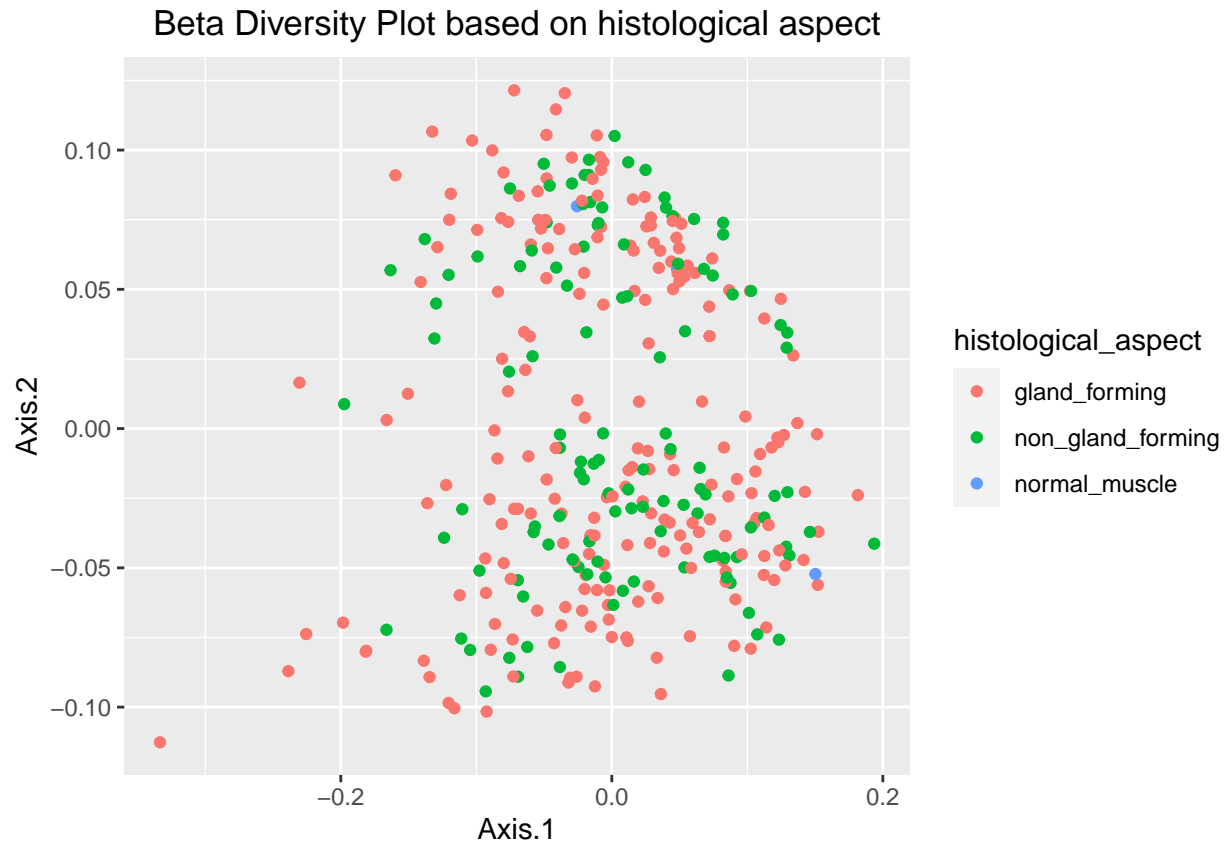
```

diversity Variation Based on Sample Origin (Confidence score 0.1 + 16s + filter read



Using the beta diversity plotting function to plot the obtained results. This function requires as input the beta_div variable (which corresponds to the Beta diversity value calculated for each sample in the CreatePhyloSeqObject function)

```
beta_plot <- generate_beta_diversity_plot(CreatePhyloSeqObject$ps_s_genus)
print(beta_plot)
```



Finally, it is possible to save the alpha and beta plots in a PNG format. For this purpose, ggsave is used.

```
## ggsave is used to save the plot in a png format, comment the code if you don't want to do it.

ggsave("alpha_div_boxplot_microdissec_shannon_species_test.png", plot = alpha_plot,
       width = 10, height = 8)

ggsave("beta_div_plot_microdissec_histolo_purist.png", plot = beta_plot, width = 10,
       height = 8)
```

Refer to the second part of this tutorial (Tutorial_Alpha_beta_part2.Rmd) for the several parameters that can be changed directly on the main script to line up with your desired analysis.