

Appendix S7: results after Qiime Open reference clustering.
Supplementary Materials of "Finding fungi in a needle stack:
high alpha and low beta-diversity of foliar endophytic
Ascomycetes revealed by metabarcoding in Corsican pine
forests".

Adrien Taudiere*

CEFE - Centre d'Ecologie Fonctionnelle et Evolutive, Montpellier: France

November 9, 2017

Abstract

Plant leaves host highly diverse communities of foliar endophytic fungi (FEF). Compared to the other compartments of the plant microbiome, FEF diversity is poorly known. We here document the communities of FEF associated with the endemic Corsican black pine *Pinus nigra* subsp. *laricio* at three sites across its natural range and examine the effect of tree age and light exposure on FEF composition. Metabarcoding using next-generation sequencing provided 8243608 Ascomycota ITS2 sequences clustered into 642 FEF operational taxonomic units (OTUs). Site is the main determinant to explain the diversity and composition of FEF communities. Tree age somewhat affects FEF community composition, whereas needle location (shade vs canopy) has no effect. Results are robust against the various options of the bioinformatic pipeline specifically developed. This study provides the first picture of FEF diversity in a Mediterranean island and underlines the complementarity of forest massifs for fungal conservation.

Key words: foliar endophyte; fungi; community ecology; metabarcoding; *Cyclaneusma minus*, *Pinus nigra* subsp. *laricio*, Mediterranean, endemism, environmental sequencing

To set the filter parameter, see directly section 'Choice of filter parameters' [2.1](#).
To read a summary of this appendix, see directly section 'Summary' [7](#).

*adrien.taudiere@zaclys.net

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 4 |
| 1.1 | R requirements | 4 |
| 1.2 | System and session informations | 4 |
| 1.3 | Some usefull functions | 6 |
| 2 | Data | 6 |
| 2.1 | Choice of filter parameters | 6 |
| 2.2 | Load and convert loading | 6 |
| 2.2.1 | Otu table | 6 |
| 2.2.2 | Taxonomy | 6 |
| 2.2.3 | Add FUNguild information to taxonomy Table | 7 |
| 2.2.4 | Representative sequences | 7 |
| 2.2.5 | Samples information | 8 |
| 2.2.6 | Create the phyloseq object | 8 |
| 2.2.7 | Characteristics of the phyloseq data | 8 |
| 2.3 | Filter sample by number of sequences | 8 |
| 2.4 | Filter OTUs by number of samples | 8 |
| 2.5 | Filter OTUs by number of sequences | 10 |
| 2.6 | Summary of filtration workflow | 11 |
| 3 | Simple description of the dataset | 11 |
| 3.1 | Number of sequences and OTUs by samples | 11 |
| 3.2 | Number of sequences and samples for each OTUs | 12 |
| 3.3 | Distribution of sequences in the taxonomy | 14 |
| 3.4 | Focus on the 30 more abundant OTUs (number of sequences) | 14 |
| 3.5 | Focus on the 30 more frequent OTUs (number of samples) | 16 |
| 4 | Number of sequences and OTUs in function of putative ecology (using FUNGuild software; Nguyen et al, 2015) | 18 |
| 5 | Distribution of fungal endophytic alpha-biodiversity | 19 |
| 5.1 | Local diversity = Diversity by sites | 19 |
| 5.2 | Diversity by age of tree | 19 |
| 5.3 | Diversity by elevation of the sample | 20 |
| 5.4 | Which factor affect diversity? | 20 |
| 6 | Effect of site, age and elevation on fungal endophytic beta-diversity | 28 |
| 6.1 | Venn diagramm | 28 |
| 6.2 | Venn diagramm for OTUs present in at least 3 samples | 28 |
| 6.3 | Ordination | 28 |
| 6.4 | Permanova on sites, host ages and elevation | 35 |
| 6.5 | Permanova on sites, host ages and individual trees | 35 |
| 6.6 | Differences in abundances and OTUs number by Order. | 36 |
| 6.7 | Differences in abundances for each OTUs | 40 |
| 6.7.1 | Pairwise comparison of the OTUs composition by sites | 40 |
| 6.7.2 | Pairwise comparison of Order composition by sites | 43 |
| 6.8 | Distribution of OTUs abundance in the fungal phylogeny | 43 |

| | | |
|----------|-----------------------------|-----------|
| 7 | Summary | 46 |
| 7.1 | Filtering summary | 46 |
| 7.2 | Alpha diversity | 46 |
| 7.3 | Beta diversity | 47 |

1 Introduction

This supplementary material presents the ecological analysis of endophytic fungal communities in *Pinus nigra* subsp. *laricio*, an endemic species of Corsica. The dataset analysed here was computed using Qiime Open reference clustering (see main article and Sup. Mat. 1 for more details).

1.1 R requirements

First, set the working directory. In this directory, there is data folder and a R script "functions_for_phyloseq.R".

```
setwd("~/Nextcloud/GitHub/FEF_paper/")
```

Then, we may need to install packages.

```
# install.packages(c('ape', 'biom', 'optparse', 'RColorBrewer', 'randomForest', 'vegan',
#                   'VennDiagram', 'venneuler', 'xtable', 'schoRsch', 'ape',
#                   'ips', 'adegenet', 'mvabund', 'rCharts', 'networkD3', 'data.tree'))
#
# # Upgrade Bioconductor to the latest version available for this version of R
# source("http://bioconductor.org/biocLite.R")
# biocLite(c("multtest", "DECIPHER", "edgeR", "phyloseq", "DESeq2", "metagenomeSeq"))
#
# require(devtools)
# install_github('ramnathu/rCharts')
# install_github("timelyportfolio/d3treeR")
```

```
## May be needed under windows
Sys.setenv(JAVA_HOME = "C:\\Program Files\\Java\\jdk1.8.0_73")

#Load the packages.
lapply(list("ggplot2", "phyloseq", "cluster", "plyr", "VennDiagram",
            "circlize", "xtable", "schoRsch", "DESeq2", "mvabund",
            "edgeR", "phangorn", "DECIPHER", "ips", "adegenet", "multtest",
            "networkD3", "treemap", "data.tree", "d3treeR", "venneuler",
            "gridExtra"), library,
       character.only = TRUE)
library(vegan)
```

1.2 System and session informations

This document was created with R version 3.4.2 (2017-09-28) on Linux the 2017-11-09 15:16:39. See below for more information.

```
sessionInfo()

## R version 3.4.2 (2017-09-28)
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: Ubuntu 16.04.3 LTS
##
## Matrix products: default
## BLAS: /usr/lib/libblas/libblas.so.3.6.0
## LAPACK: /usr/lib/lapack/liblapack.so.3.6.0
##
```

```

## locale:
## [1] LC_CTYPE=fr_FR.UTF-8          LC_NUMERIC=C
## [3] LC_TIME=fr_FR.UTF-8          LC_COLLATE=fr_FR.UTF-8
## [5] LC_MONETARY=fr_FR.UTF-8      LC_MESSAGES=fr_FR.UTF-8
## [7] LC_PAPER=fr_FR.UTF-8         LC_NAME=fr_FR.UTF-8
## [9] LC_ADDRESS=fr_FR.UTF-8       LC_TELEPHONE=fr_FR.UTF-8
## [11] LC_MEASUREMENT=fr_FR.UTF-8   LC_IDENTIFICATION=fr_FR.UTF-8
##
## attached base packages:
## [1] parallel stats4 grid stats graphics grDevices utils
## [8] datasets methods base
##
## other attached packages:
## [1] vegan_2.4-4 lattice_0.20-35
## [3] permute_0.9-4 gridExtra_2.2.1
## [5] venneuler_1.1-0 rJava_0.9-8
## [7] d3treeR_0.1 data.tree_0.7.0
## [9] treemap_2.4-2 networkD3_0.4
## [11] multtest_2.32.0 adegenet_2.1.0
## [13] ade4_1.7-8 ips_0.0-7
## [15] XML_3.98-1.9 colorspace_1.3-2
## [17] DECIPHER_2.4.0 RSQLite_2.0
## [19] Biostrings_2.44.2 XVector_0.16.0
## [21] phangorn_2.2.0 ape_4.1
## [23] edgeR_3.18.1 limma_3.32.5
## [25] mvabund_3.12.3 DESeq2_1.16.1
## [27] SummarizedExperiment_1.6.3 DelayedArray_0.2.7
## [29] matrixStats_0.52.2 Biobase_2.36.2
## [31] GenomicRanges_1.28.4 GenomeInfoDb_1.12.2
## [33] IRanges_2.10.3 S4Vectors_0.14.3
## [35] BiocGenerics_0.22.0 schoRsch_1.4
## [37] xtable_1.8-2 circlize_0.4.1
## [39] VennDiagram_1.6.17 futile.logger_1.4.3
## [41] plyr_1.8.4 cluster_2.0.6
## [43] phyloseq_1.20.0 ggplot2_2.2.1
## [45] knitr_1.17
##
## loaded via a namespace (and not attached):
## [1] backports_1.1.0 Hmisc_4.0-3
## [3] fastmatch_1.1-0 igraph_1.1.2
## [5] lazyeval_0.2.0 sp_1.2-5
## [7] splines_3.4.2 BiocParallel_1.10.1
## [9] gridBase_0.4-7 digest_0.6.12
## [11] foreach_1.4.3 htmltools_0.3.6
## [13] viridis_0.4.0 gdata_2.18.0
## [15] magrittr_1.5 checkmate_1.8.3
## [17] memoise_1.1.0 readr_1.1.1
## [19] annotate_1.54.0 gmodels_2.16.2
## [21] blob_1.1.0 dplyr_0.7.2
## [23] RCurl_1.95-4.8 jsonlite_1.5
## [25] genefilter_1.58.1 bindr_0.1
## [27] brew_1.0-6 survival_2.41-3
## [29] iterators_1.0.8 glue_1.1.1
## [31] gtable_0.2.0 zlibbioc_1.22.0
## [33] seqinr_3.4-5 Rook_1.1-1
## [35] shape_1.4.3 scales_0.5.0
## [37] futile.options_1.0.0 DBI_0.7
## [39] Rcpp_0.12.12 viridisLite_0.2.0
## [41] htmlTable_1.9 foreign_0.8-69
## [43] bit_1.1-12 spdep_0.6-15
## [45] Formula_1.2-2 tweedie_2.2.5
## [47] htmlwidgets_0.9 DiagrammeR_0.9.1
## [49] RColorBrewer_1.1-2 acepack_1.4.1
## [51] pkgconfig_2.0.1 nnet_7.3-12
## [53] deldir_0.1-14 locfit_1.5-9.1
## [55] rlang_0.1.2 reshape2_1.4.2
## [57] AnnotationDbi_1.38.2 visNetwork_2.0.1
## [59] munsell_0.4.3 tools_3.4.2
## [61] downloader_0.4 evaluate_0.10.1
## [63] biomformat_1.4.0 stringr_1.2.0
## [65] bit64_0.9-7 purrr_0.2.3
## [67] bindrcpp_0.2 nlme_3.1-131
## [69] mime_0.5 rstudioapi_0.6
## [71] compiler_3.4.2 rgexf_0.15.3
## [73] tibble_1.3.4 statmod_1.4.30
## [75] geneplotter_1.54.0 stringi_1.1.5
## [77] highr_0.6 Matrix_1.2-11
## [79] LearnBayes_2.15 GlobalOptions_0.0.12
## [81] data.table_1.10.4 bitops_1.0-6
## [83] httpuv_1.3.5 R6_2.2.2
## [85] latticeExtra_0.6-28 gridSVG_1.5-1
## [87] codetools_0.2-15 lambda.r_1.1.9
## [89] boot_1.3-20 MASS_7.3-47
## [91] gtools_3.5.0 assertthat_0.2.0
## [93] rhdf5_2.20.0 GenomeInfoDbData_0.99.0
## [95] mgcv_1.8-22 expm_0.999-2
## [97] hms_0.3 influenceR_0.1.0
## [99] quadprog_1.5-5 rpart_4.1-11
## [101] tidyr_0.7.1 coda_0.19-1
## [103] shiny_1.0.5 base64enc_0.1-3

```

1.3 Some usefull functions

The function `as.binaryOtuTable` converts a phyloseq object into a phyloseq object with binary (*i.e.* 0/1) OTU table. It allows to suppress effect due to the number of sequences wich may be the result of a lot of molecular artefact (Lindhal et al., 2013).

`funky.color` and `transpa` allow to create nice color palette.

`accu_plot` allows to plot accumulation curves in fonction of a factor in samples data (`@sam_data` of phyloseq object).

`otu_circle` uses the package `circlize` to plot circle of OTUs/sequences distributions in samples. `sankey_phyloseq` is an alternative using Sankey plot.

`phyloseq_to_edgeR`, wrote by Paul J. McMurdie, converts phyloseq OTU count data into DGEList for edgeR package.

`plot_deseq2_phyloseq` and `plot_edgeR_phyloseq` plot the result of differential analysis of count data (using either the package DESeq2 or edgeR).

```
source(file = "functions_for_phyloseq.R")
```

2 Data

2.1 Choice of filter parameters

```
#Choose the dataset folder
data_folder <- "Open_ref"

#Choose the minimum number of sequences by sample.
N_sam_min <- 20000

#Choose the minimum number of samples by OTU.
N_otu_sam_min <- 1

#Choose the minimum number of sequences by OTU.
N_seq_otu_min <- 5
```

2.2 Load and convert loading

2.2.1 Otu table

```
#Import biom data
dataBiom <- import_biom(paste("data/", data_folder, "/otu_table.biom", sep=""))
```

2.2.2 Taxonomy

```

#Import taxonomy data
taxRDP_brut <- readLines(paste("data/", data_folder, "/tax_assignments.txt", sep=""))
taxRDP_brut <- gsub(";", "\t", taxRDP_brut)
taxRDP_brut <- gsub("\"", "", taxRDP_brut)
taxRDP_brut <- gsub("\\(", "\t", taxRDP_brut)
taxRDP_brut <- gsub("*_", "\t", taxRDP_brut)
taxRDP_brut <- read.table(textConnection(taxRDP_brut), sep = "\t", fill = TRUE)

# Format taxonomy for phyloseq
taxRDP <- taxRDP_brut[match(taxa_names(dataBiom), taxRDP_brut[, 1]),
                      c(1, 3, 5, 7, 9, 11, 13, 15)]
taxRDP <- tax_table(as.matrix(taxRDP))
taxa_names(taxRDP) <- taxa_names(dataBiom)
colnames(taxRDP) <- c("Species Hypothesis", "Domain", "Phylum", "Class", "Order", "Family",
                     "Genus", "Species")

```

2.2.3 Add FUNguild information to taxonomy Table

```

taxRDP2 <- as.data.frame(taxRDP)
funguild <- read.delim(paste("data/", data_folder, "/FUNGUILD.guilds.txt", sep = ""))

match_interm <- match(rownames(taxRDP2), funguild$OTU_ID)

taxRDP2$Trophic_Mode <- NA
taxRDP2$Trophic_Mode <- as.character(funguild$Trophic.Mode)[match_interm]
taxRDP2$Guild <- NA
taxRDP2$Guild <- as.character(funguild$Guild)[match_interm]
taxRDP2$Confidence_Ranking <- NA
taxRDP2$Confidence_Ranking <- as.character(funguild$Confidence.Ranking)[match_interm]
taxRDP2$Growth_Morphology <- NA
taxRDP2$Growth_Morphology <- as.character(funguild$Growth.Morphology)[match_interm]
taxRDP2$Trait <- NA
taxRDP2$Trait <- as.character(funguild$Trait)[match_interm]

taxRDP2 <- tax_table(as.matrix(taxRDP2))
taxa_names(taxRDP2) <- taxa_names(dataBiom)
colnames(taxRDP2) <- c("Species Hypothesis", "Domain", "Phylum", "Class", "Order", "Family", "Genus",
                     "Trophic_Mode", "Guild", "Confidence_Ranking", "Growth_Morphology",
                     "Trait")

```

2.2.4 Representative sequences

```

map_endo <-
  import_qiime(map = "data/map_qiimedata.txt")

## Processing map file...

map_endo <- map_endo[order(rownames(map_endo)),]

```

2.2.5 Samples information

```
repset <- import_qiime(refseqfilename = paste("data/", data_folder, "/seq.fasta",
                                             sep = ""))

## Processing Reference Sequences...

taxa_names(repset) <- unlist(strsplit(taxa_names(repset),
                                     split = " "))[seq(1, 2*length(repset), by = 2)]
```

2.2.6 Create the phyloseq object

```
data_all <- merge_phyloseq(dataBiom, repset, taxRDP2)

sample_data(data_all) <- map_endo

data_all@tax_table[data_all@tax_table == ""] <- NA
```

2.2.7 Characteristics of the phyloseq data

```
data_all

## phyloseq-class experiment-level object
## otu_table() OTU Table: [ 4373 taxa and 80 samples ]
## sample_data() Sample Data: [ 80 samples by 6 sample variables ]
## tax_table() Taxonomy Table: [ 4373 taxa by 13 taxonomic ranks ]
## refseq() DNASTringSet: [ 4373 reference sequences ]
```

The data are made of 8.398038×10^6 sequences representing 4373 OTUs allocate to 80 samples.

2.3 Filter sample by number of sequences

```
N_sam_min

## [1] 20000
```

If we discard samples with less than 2×10^4 sequences, we keep 72 on the 80 samples (90%).

```
barplot(sort(sample_sums(data_all)))
abline(h = N_sam_min)
data.f1 <- prune_samples(sample_sums(data_all) > N_sam_min, data_all)
data.f1 <- prune_taxa(taxa_sums(data.f1) >= 1, data.f1)
```

2.4 Filter OTUs by number of samples

First, we can visualize the number of OTUs in a given number of samples (Figure 2.2).

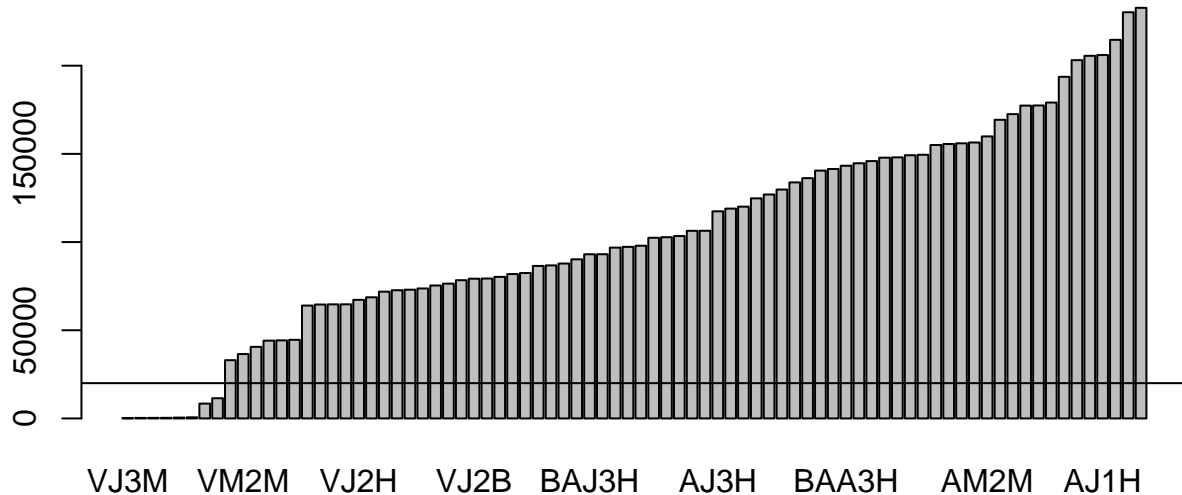


Figure 2.1: Number of sequences by sample. Horizontal line indicates the filtering parameter.

```
df_nbOtu_sample <- data.frame("Nb of OTUs" = table(rowSums(as.binaryOtuTable(
  data.f1@otu_table)) [table(rowSums(as.binaryOtuTable(data.f1@otu_table)) > 1],
  "Nb samples" = as.numeric(names(table(rowSums(as.binaryOtuTable(data.f1@otu_table))
    [table(rowSums(as.binaryOtuTable(data.f1@otu_table)) > 1]))))
```

```
g <- ggplot(df_nbOtu_sample, aes(y = Nb.of.OTUs.Freq, x = Nb.samples))
g + geom_point(size = 4, col = rgb(0.1, 0.1, 0.1, 0.5)) +
  scale_y_continuous(trans = 'log10') +
  geom_smooth(size = 2, col = rgb(0.1, 0.1, 0.1, 0.5)) +
  geom_vline(xintercept= N_otu_sam_min)
```

```
## 'geom_smooth()' using method = 'loess'
```

```
summary(df_nbOtu_sample$Nb.samples)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      1.00  18.75   36.50   36.50  54.25   72.00
```

```
N_otu_sam_min
```

```
## [1] 1
```

```
data.f2 <- prune_taxa(rowSums(as.binaryOtuTable(data.f1@otu_table) >=
  N_otu_sam_min, data.f1)
```

If we discard OTUs present in less than 1 sample, we keep 4359 on the 4359 OTUs (100%).

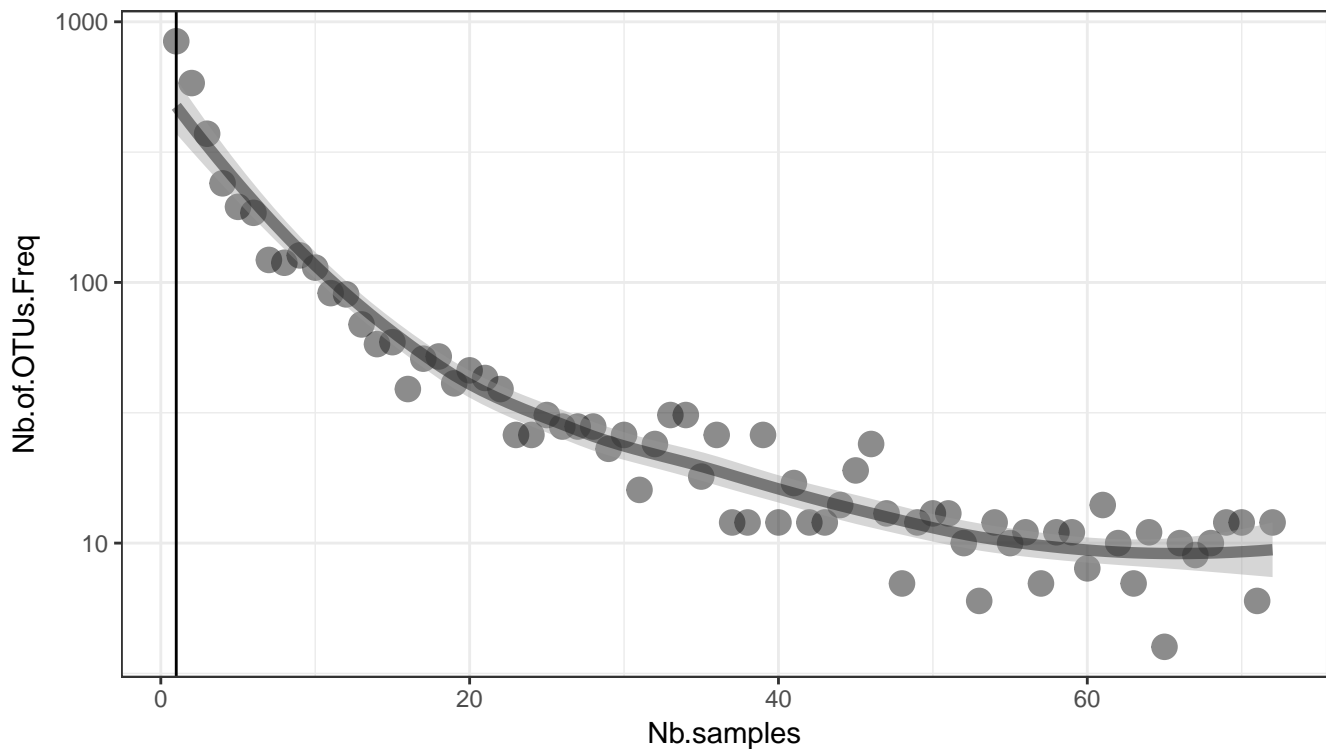


Figure 2.2: Number of OTU present in a given number of samples. Vertical bar illustrates the filtering parameter.

2.5 Filter OTUs by number of sequences

We can visualize the number of sequences by OTU (Figure 2.3).

```
df_nbseq_Otu <- data.frame("Nb of sequences by OTUs" = rowSums(data.f2@otu_table))
g <- ggplot(df_nbseq_Otu, aes(x = Nb.of.sequences.by.OTUs))
g + geom_histogram(size = 2, col = rgb(0.8, 0.8, 0.8, 0.3)) +
  scale_x_continuous(trans = 'log10') +
  geom_vline(xintercept= N_seq_otu_min)
```

'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.

```
summary(df_nbseq_Otu[, 1])
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##         1         5       22   1922   124    773785
```

```
N_seq_otu_min
```

```
## [1] 5
```

If we discard OTUs with less than 1 sequences, we keep 3382 on the 4373 OTUs (77.34%).

```
data.f3 <- prune_taxa(rowSums(data.f2@otu_table) >= N_seq_otu_min, data.f2)
```

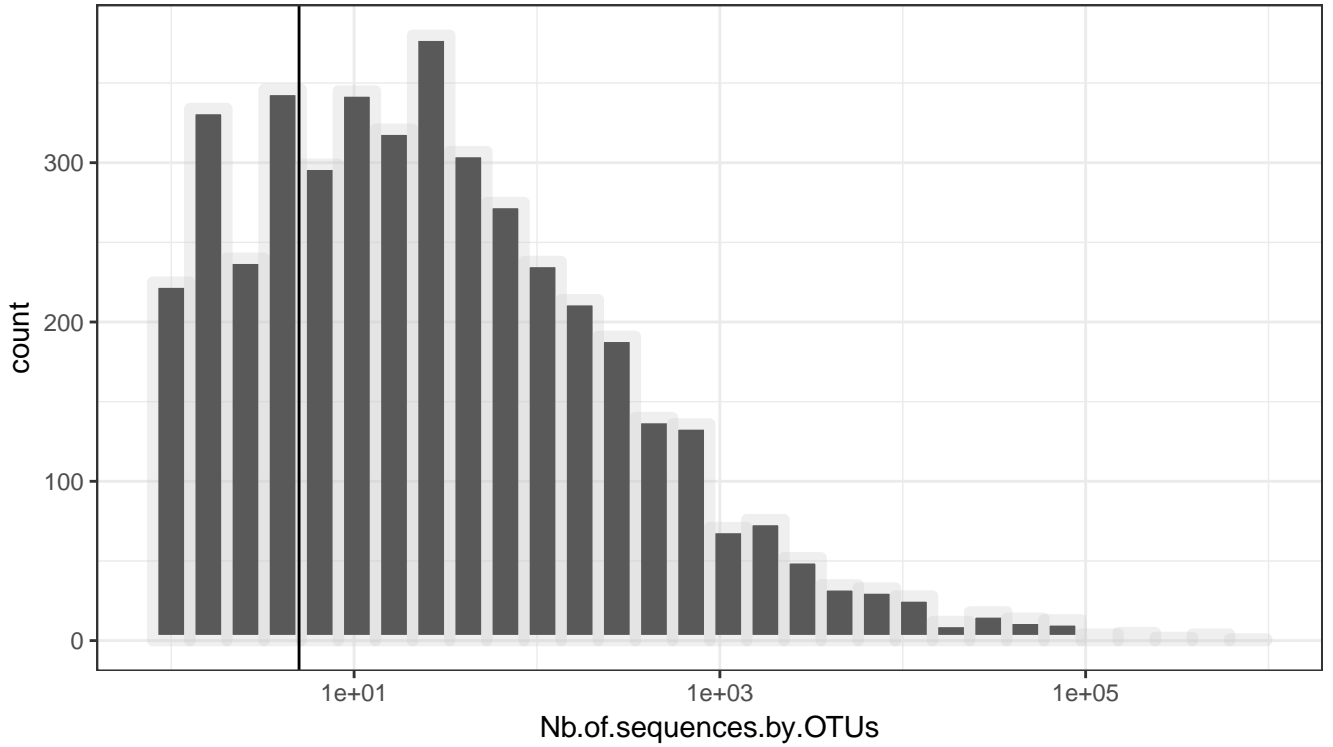


Figure 2.3: Number of sequences by OTU (log10 transformed). Horizontal bar illustrates the filtering parameter.

2.6 Summary of filtration workflow

The filtered data are made of 8.373567×10^6 sequences representing 3382 OTUs allocate to 72 samples.

| | Nb.of.OTUs | Nb.of.samples | Nb.of.sequences |
|--|------------|---------------|-----------------|
| No filter | 4373 | 80 | 8398038.00 |
| Nb of sequences by sample ≥ 20000 | 4359 | 72 | 8375892.00 |
| Nb of sample by OTUs ≥ 1 | 4359 | 72 | 8375892.00 |
| Nb of sequences by OTUs ≥ 5 | 3382 | 72 | 8373567.00 |

Table 1: Number of OTUs, samples and sequences after filtering

3 Simple description of the dataset

3.1 Number of sequences and OTUs by samples

```
df_nbseq_nbotu <- data.frame("Nb of sequences by samples" = colSums(data.f3@otu_table),
                             "Nb of OTUs by samples" =
                               colSums(as.binaryOtuTable(data.f3)@otu_table))

g <- ggplot(df_nbseq_nbotu, aes(x = Nb.of.OTUs.by.samples,
                               y = Nb.of.sequences.by.samples))
g + geom_point(size = 3, col = rgb(0.1, 0.1, 0.1, 0.5)) +
  scale_y_continuous(trans = 'log10') +
```

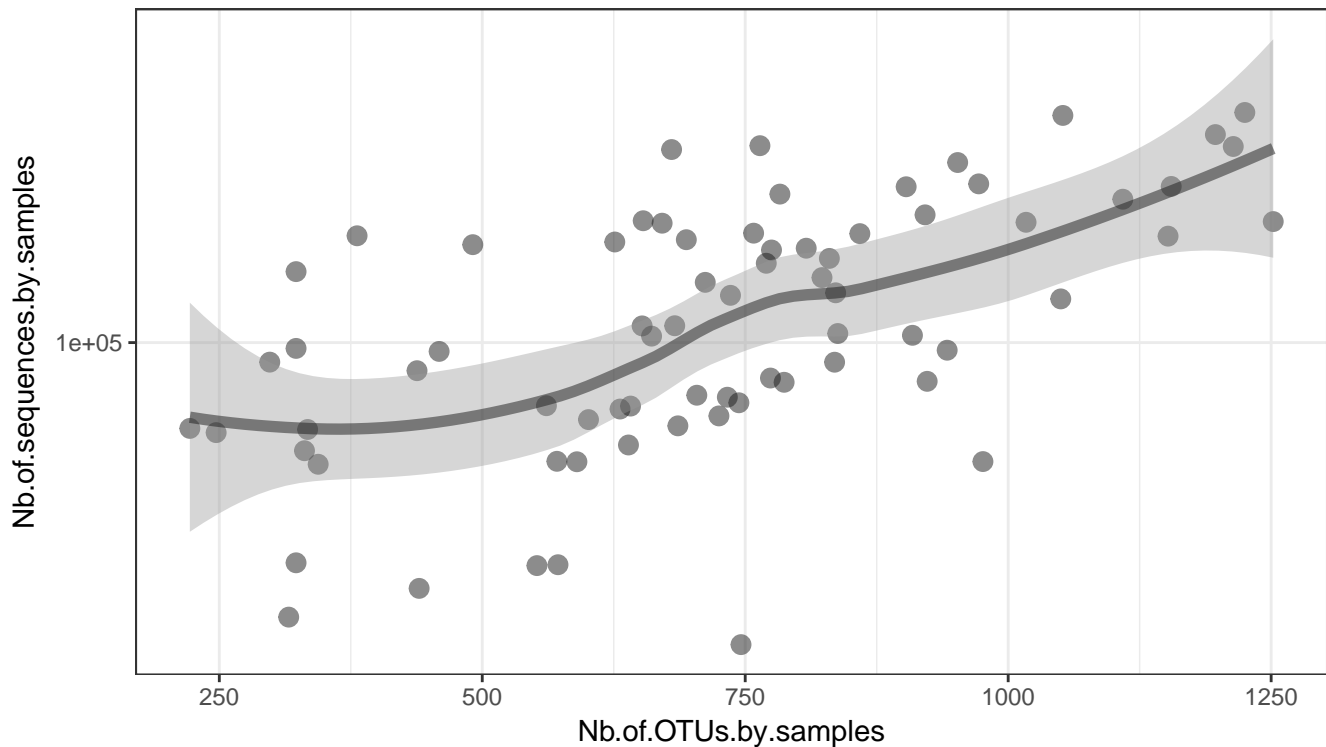


Figure 3.1: Number of OTUs by sample in fonction of the number of sequences by sample (log10 axe). The tendency is represented by the line obtained from loess (Local Polynomial Regression Fitting).

```
geom_smooth(size = 2, col = rgb(0.1, 0.1, 0.1, 0.5))

## 'geom_smooth()' using method = 'loess'

ggplot(as.data.frame(data.f3@refseq@ranges), aes(x = width)) + geom_density() +
  ylab("Reference sequences length")
```

3.2 Number of sequences and samples for each OTUs

```
df_nbseq_nbsam <- data.frame("Nb of sequences by OTUs" = rowSums(data.f3@otu_table)
                             [rowSums(data.f3@otu_table) > 0],
                             "Nb of samples by OTUs" =
                               rowSums(as.binaryOtuTable(data.f3)@otu_table)
                               [rowSums(data.f3@otu_table) > 0])

g <- ggplot(df_nbseq_nbsam, aes(y = Nb.of.samples.by.OTUs,
                               x = Nb.of.sequences.by.OTUs))
g + geom_point(size = 3, col = rgb(0.1, 0.1, 0.1, 0.5)) +
  scale_x_continuous(trans = 'log10') +
  geom_smooth(size = 2, col = rgb(0.1, 0.1, 0.1, 0.5), method = "gam",
             formula = y ~ s(x, bs = "cs"))
```

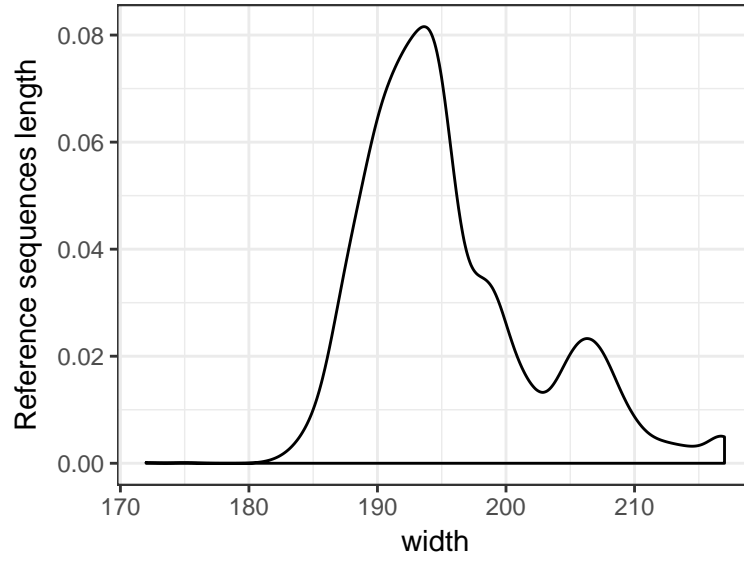


Figure 3.2: Distribution of reference sequences length.

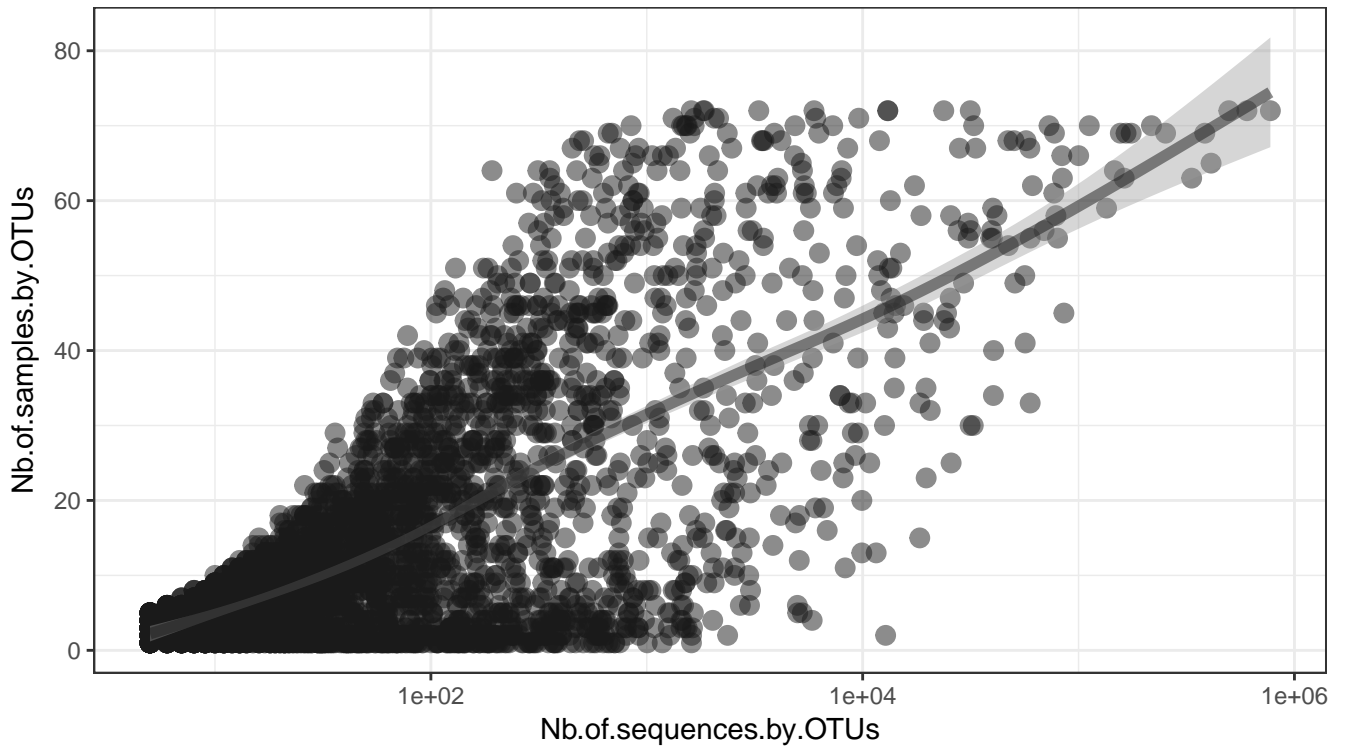


Figure 3.3: Number of sequences by OTUs (log10 axe) in fonction of the number of samples where OTUs were found. The tendency is represented by the line obtain from gam (Generalized additive models with integrated smoothness estimation).

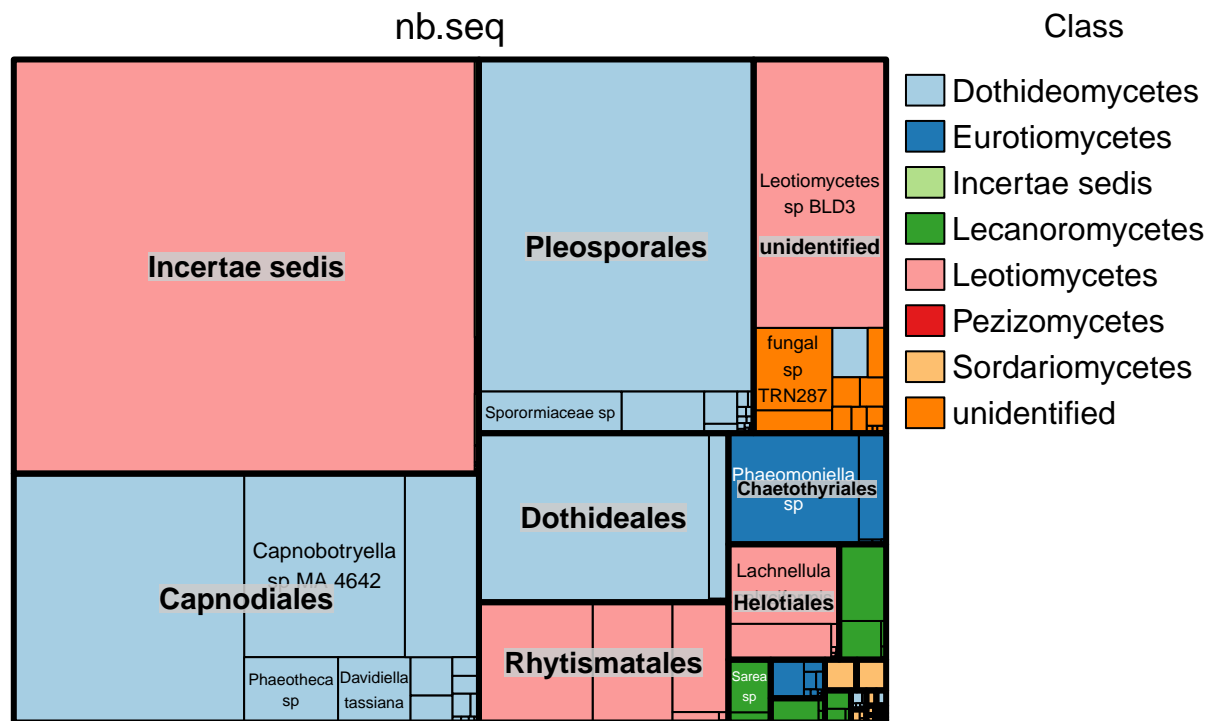


Figure 3.4: Distribution of the number of sequences in the Ascomycota taxonomy. Colors represent Class, bold lines delimit Order and thick line delimit species.

3.3 Distribution of sequences in the taxonomy

```
df3 <- data.frame(as.data.frame(data.f3@tax_table), nb.seq = rowSums(data.f3@otu_table))
tm <- treemap(df3, index = c("Order", "Species"), vSize = "nb.seq", vColor = "Class",
  type = "categorical", palette = "Paired")
# For an interactive version in html
# d3tree(tm)
```

3.4 Focus on the 30 more abundant OTUs (number of sequences)

```
the30mostfrequents <- sort(decreasing = T, rowSums(data.f3@otu_table))[1:30]
barplot(the30mostfrequents, horiz = T, cex.names = 0.4, las = 2)
```

```
print(xtable(df_the30mostfrequent[, c(1:8, 12)], auto = T,
  caption = "Taxonomie of the 30 more
  abundant OTUs (number of sequences)",
  size = "\\tiny", include.rownames = FALSE)
```

| Domain | Phylum | Class | Order | Family | Genus | Species | Trophic_Mode | Trait |
|--------|--------------|-----------------|-----------------|---------------------|----------------|---------------------------|--------------|-------|
| Fungi | Ascomycota | Leotiomycetes | Incertae sedis | Incertae sedis | Cyclaneusma | Cyclaneusma minus | - | - |
| Fungi | Ascomycota | Leotiomycetes | Incertae sedis | Incertae sedis | Cyclaneusma | Cyclaneusma minus | - | - |
| Fungi | Ascomycota | Leotiomycetes | Incertae sedis | Incertae sedis | Cyclaneusma | Cyclaneusma minus | - | - |
| Fungi | Ascomycota | Dothideomycetes | Pleosporales | | | | - | - |
| Fungi | Ascomycota | Dothideomycetes | Dothideales | Dothioraceae | unidentified | Dothioraceae sp | - | - |
| Fungi | Ascomycota | Dothideomycetes | Capnodiales | Incertae sedis | Capnobotryella | Capnobotryella sp MA 4642 | Saprotroph | NULL |
| Fungi | Ascomycota | Dothideomycetes | Pleosporales | | | | - | - |
| Fungi | Ascomycota | Dothideomycetes | Capnodiales | | | | - | - |
| Fungi | Ascomycota | Dothideomycetes | Pleosporales | | | | - | - |
| Fungi | Ascomycota | Leotiomycetes | unidentified | unidentified | unidentified | Leotiomycetes sp BLD3 | - | - |
| Fungi | Ascomycota | Leotiomycetes | unidentified | unidentified | unidentified | Leotiomycetes sp BLD3 | - | - |
| Fungi | Ascomycota | Dothideomycetes | Pleosporales | | | | - | - |
| Fungi | Ascomycota | Dothideomycetes | Capnodiales | Mycosphaerellaceae | Phaeothecoidea | Phaeothecoidea sp | Saprotroph | NULL |
| Fungi | | | | | | | - | - |
| Fungi | Ascomycota | Leotiomycetes | Rhytismatales | Rhytismataceae | Lophodermium | Lophodermium conigenum | Pathotroph | NULL |
| Fungi | | | | | | | - | - |
| Fungi | Ascomycota | Eurotiomycetes | Chaetothyriales | Herpotrichiellaceae | Phaeomoniella | Phaeomoniella sp | Saprotroph | NULL |
| Fungi | Ascomycota | Dothideomycetes | Capnodiales | | | | - | - |
| Fungi | | | | | | | - | - |
| Fungi | Ascomycota | Leotiomycetes | Incertae sedis | Incertae sedis | Cyclaneusma | Cyclaneusma minus | - | - |
| Fungi | Ascomycota | Leotiomycetes | unidentified | unidentified | unidentified | Leotiomycetes sp BLD3 | - | - |
| Fungi | Ascomycota | Leotiomycetes | Incertae sedis | Incertae sedis | Cyclaneusma | Cyclaneusma minus | - | - |
| Fungi | unidentified | unidentified | unidentified | unidentified | unidentified | fungal sp TRN287 | - | - |
| Fungi | Ascomycota | Eurotiomycetes | Chaetothyriales | Herpotrichiellaceae | Phaeomoniella | Phaeomoniella sp | Saprotroph | NULL |
| Fungi | Ascomycota | Dothideomycetes | Capnodiales | | | | - | - |
| Fungi | | | | | | | - | - |
| Fungi | Ascomycota | Leotiomycetes | Rhytismatales | Rhytismataceae | Lophodermium | | Pathotroph | NULL |

Table 2: Taxonomie of the 30 more abundant OTUs (number of sequences)

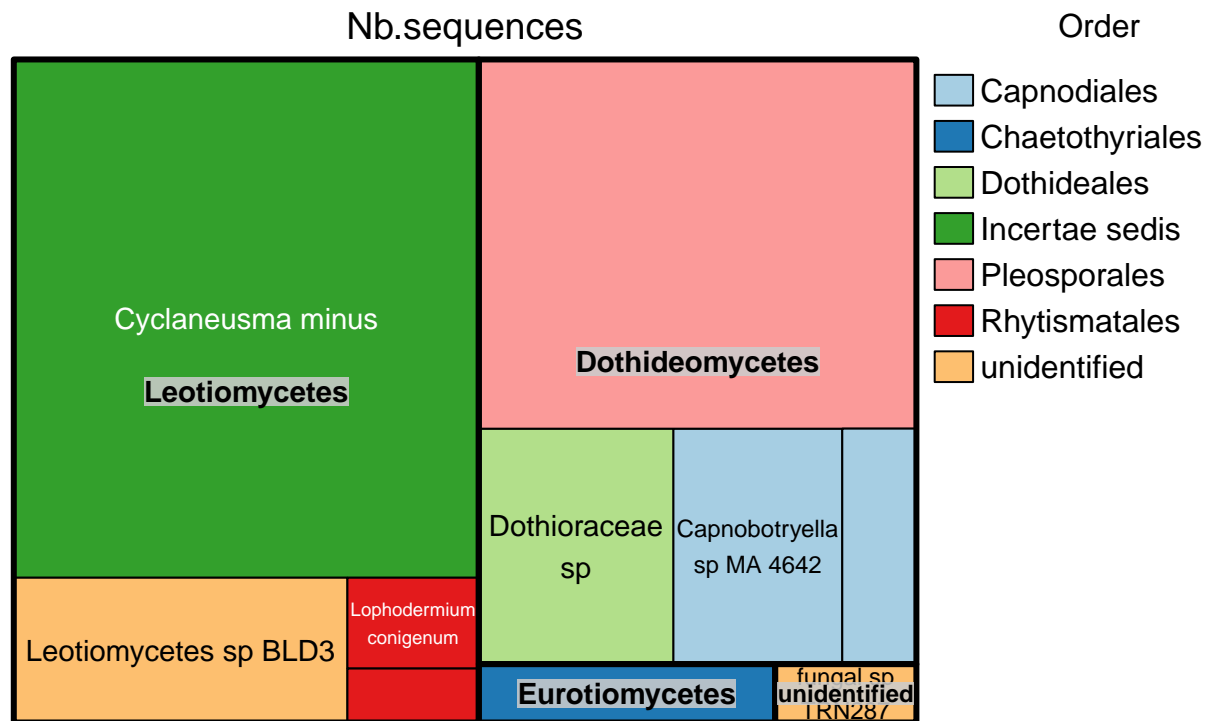


Figure 3.5: Number of sequences of the 30 most abundant OTUs (number of sequences). Colors indicate Order, bold lines delimit Class and thick lines delimit species.

```
treemap(df_the30mostfrequents, index = c("Class", "Species"), vSize = "Nb.sequences",
        vColor = "Order", type = "categorical", palette = "Paired")
```

3.5 Focus on the 30 more frequent OTUs (number of samples)

```
the30mostfrequents_samp <- sort(decreasing = T,
                                rowSums(as.binaryOtuTable(data.f3@otu_table))[1:30])
barplot(the30mostfrequents_samp, horiz = T, cex.names = 0.4, las = 2)
```



```
print(xtable(df_the30mostfrequents_samp[, c(1:8, 12)], auto = T,
  caption = "Taxonomie of the 30 more frequent OTUs (number of samples)",
  size = "\\tiny", include.rownames = FALSE)
```

[illegible]

Table 3: Taxonomie of the 30 more frequent OTUs (number of samples)

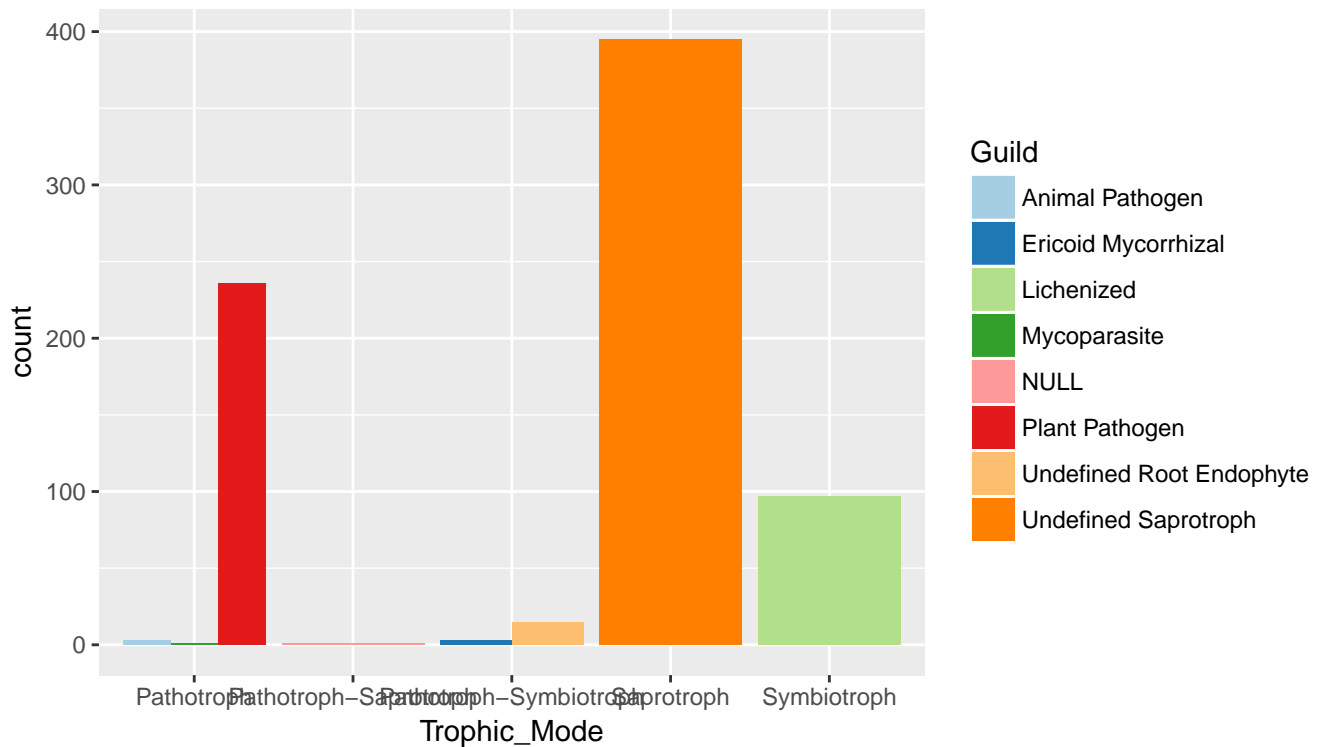


Figure 4.1: Distribution of OTUs into functional Guild.

4 Number of sequences and OTUs in function of putative ecology (using FUNGuild software; Nguyen et al, 2015)

```
tabPutativeEcology <- apply(data.f3@tax_table, 2, function(x) table(x))
tabPutativeEcology_percent <- apply(data.f3@tax_table, 2, function(x)
  round(table(x)/dim(data.f3@tax_table)[1]*100, 3))
sum(data.f3@otu_table[data.f3@tax_table[, "Trophic_Mode"] == "-"]) /
  sum(data.f3@otu_table)*100

## [1] 81.36235

tmdata <- as.data.frame(data.f3@tax_table[data.f3@tax_table[, "Trophic_Mode"] != "-"])
tmdata$Nb.sequences <- rowSums(data.f3@otu_table[data.f3@tax_table[, "Trophic_Mode"] != "-"])
tmdata$Nb.OTU <- rep(1, length(tmdata$Nb.sequences))
```

```
ggplot(tmdata) + geom_bar(aes(x = Trophic_Mode, fill=Guild), position = "dodge") +
  scale_fill_discrete("Paired") + theme_grey()
```

```
ggplot(tmdata, stat = "identity") +
  geom_bar(aes(x = Trophic_Mode, weight = Nb.sequences, fill = Guild), position = "dodge") +
  scale_fill_discrete("Paired") + scale_y_log10() + theme_grey()
```

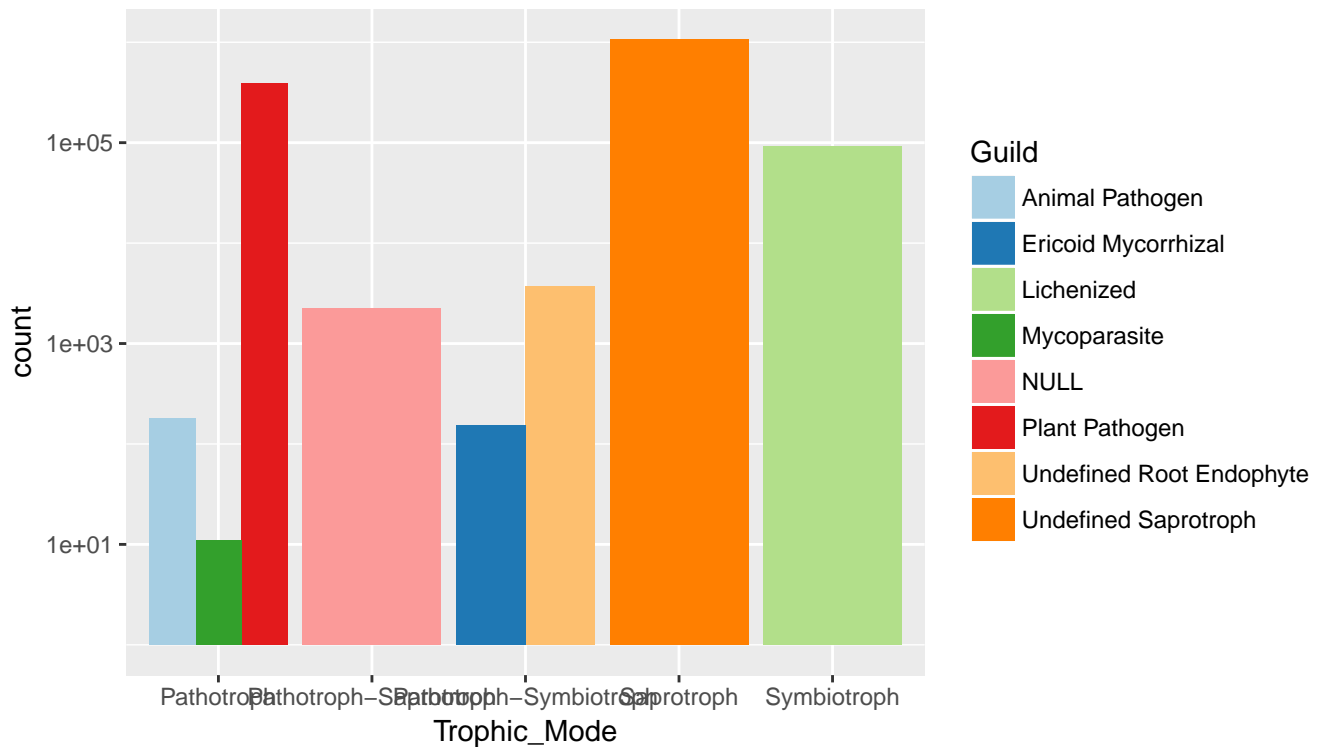


Figure 4.2: Distribution of sequences (log10 transformed) into functional Guild.

5 Distribution of fungal endophytic alpha-biodiversity

5.1 Local diversity = Diversity by sites

```
accu_plot(data.f3, "Sites", nbSeq = FALSE)
```

```
accu_plot(data.f3, "Sites", step = 5000)
```

```
measures_index <- c("Observed", "Chao1", "Shannon", "Simpson")
p <- plot_richness(data.f3, x = "Sites", color = "Sites", measures = measures_index)
p + geom_boxplot(data = p$data, alpha = 0.5)
```

5.2 Diversity by age of tree

```
accu_plot(data.f3, "Age", nbSeq = FALSE)
```

```
measures_index <- c("Observed", "Chao1", "Shannon", "Simpson")
p <- plot_richness(data.f3, x = "Age", color = "Sites", measures = measures_index)
p + geom_boxplot(data = p$data, aes(x = p$data$Age, y = value, color = NULL),
  alpha = 0.5)
```

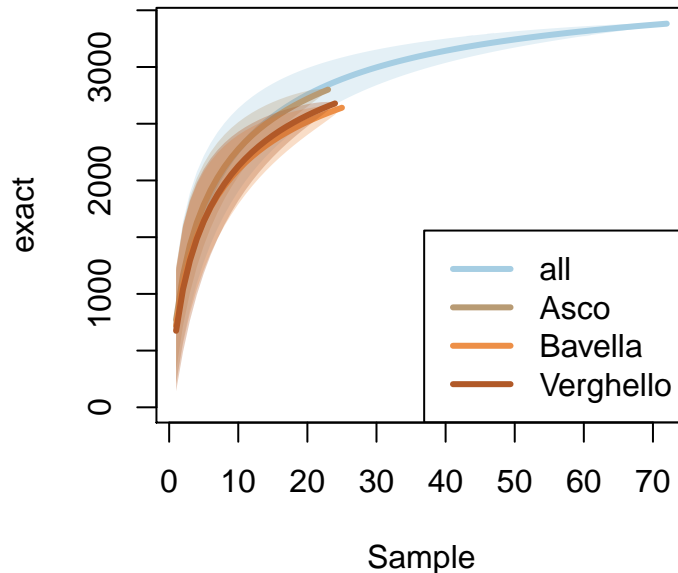


Figure 5.1: Rarefaction curves for each site. Note that if singletons were removed, these curves are biased.

5.3 Diversity by elevation of the sample

```
accu_plot(data.f3, "Elevation", nbSeq = FALSE)
```

```
measures_index <- c("Observed", "Chao1", "Shannon", "Simpson")
p <- plot_richness(data.f3, x = "Elevation", color = "Sites", measures = measures_index)
p + geom_boxplot(data = p$data, aes(x = p$data$Elevation, y = value, color = NULL),
  alpha = 0.5)
```

5.4 Which factor affect diversity?

```
## Uneven sequencing depth may have an impact
readNumbers = apply(t(data.f3@otu_table), 1, sum)

otuHill <- renyi(t(data.f3@otu_table), scale = c(0, 1, 2), hill = T)

hill.1 = otuHill$"0"
hill.2 = otuHill$"1"
hill.3 = otuHill$"2"

hill.1.m1 = lm(hill.1 ~ sqrt(readNumbers) + data.f3@sam_data$Sites +
  data.f3@sam_data$Age + data.f3@sam_data$Elevation)
```

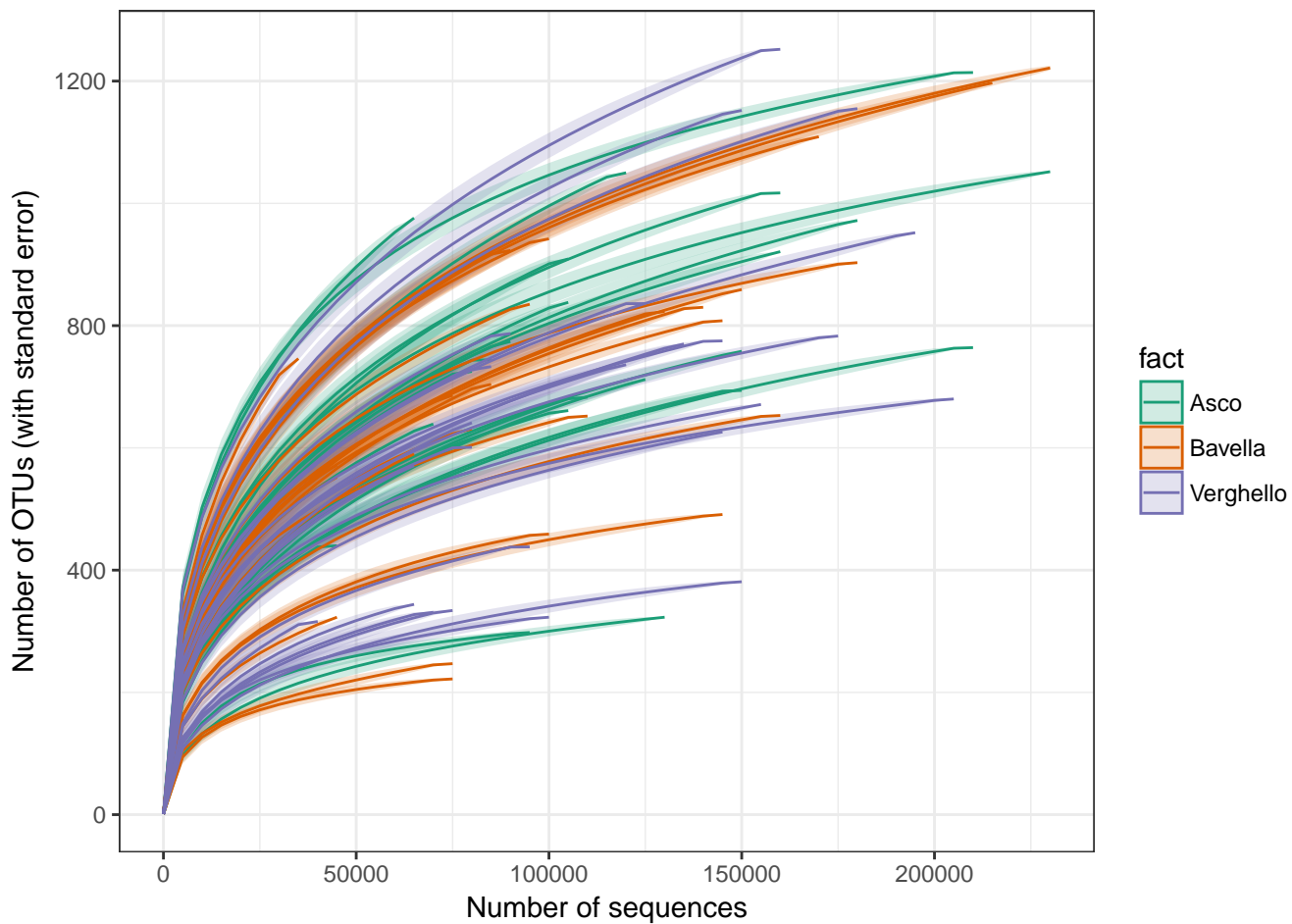


Figure 5.2: Rarefaction curves for each sample using sequences number on x-axes. Note that if singletons were removed, these curves are biased.

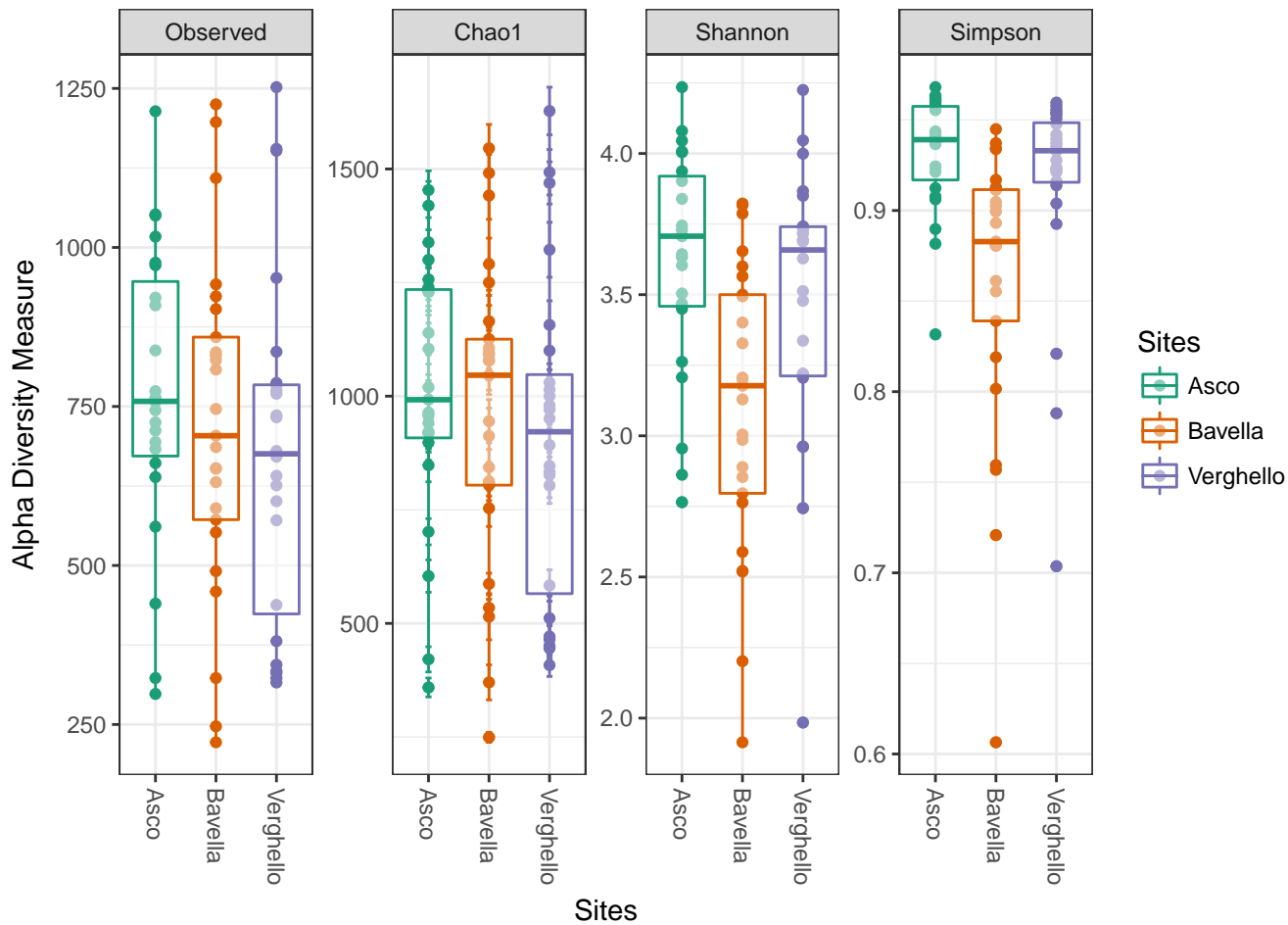


Figure 5.3: Diversity of each sites

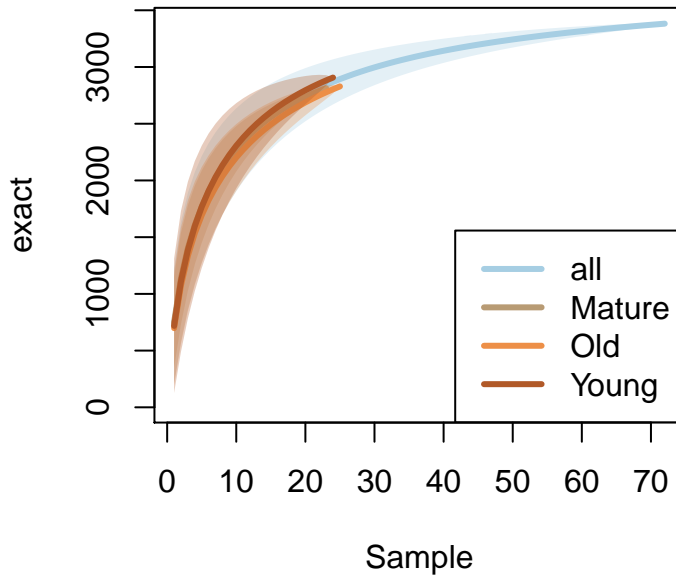


Figure 5.4: Rarefaction curves for each host age. Note that if singletons were removed, these curves are biased.

```
hill.2.m1 = lm(hill.2 ~ sqrt(readNumbers) + data.f3@sam_data$Sites +
               data.f3@sam_data$Age + data.f3@sam_data$Elevation)
hill.3.m1 = lm(hill.3 ~ sqrt(readNumbers) + data.f3@sam_data$Sites +
               data.f3@sam_data$Age + data.f3@sam_data$Elevation)
```

| | Estimate | Std. Error | t value | Pr(> t) |
|-----------------------------------|-------------|-------------|------------|-----------|
| (Intercept) | 29.4189474 | 133.5044144 | 0.2203594 | 0.8262927 |
| sqrt(readNumbers) | 2.2375997 | 0.3396355 | 6.5882379 | 0.0000000 |
| data.f3@sam_data\$SitesBavella | -10.1118810 | 59.3953793 | -0.1702469 | 0.8653530 |
| data.f3@sam_data\$SitesVerghello | -84.7689405 | 59.4710379 | -1.4253819 | 0.1589059 |
| data.f3@sam_data\$AgeOld | -25.8215383 | 59.1438094 | -0.4365890 | 0.6638788 |
| data.f3@sam_data\$AgeYoung | -89.9361492 | 60.5620816 | -1.4850241 | 0.1424466 |
| data.f3@sam_data\$ElevationMiddle | 54.0275700 | 59.8951283 | 0.9020361 | 0.3704200 |
| data.f3@sam_data\$ElevationTop | -3.7752802 | 59.0829784 | -0.0638979 | 0.9492507 |

Table 4: Summary of the linear model of species richness (Hill number with $q = 0$)

Post-hoc Tukey tests among the three experimental treatments with partial residuals, after accounting for differential sequencing success.

```
tuk1 <- TukeyHSD(aov(lm(hill.1 ~ sqrt(readNumbers))$residuals ~ data.f3@sam_data$Site))
tuk2 <- TukeyHSD(aov(lm(hill.2 ~ sqrt(readNumbers))$residuals ~ data.f3@sam_data$Site))
tuk3 <- TukeyHSD(aov(lm(hill.3 ~ sqrt(readNumbers))$residuals ~ data.f3@sam_data$Site))
```

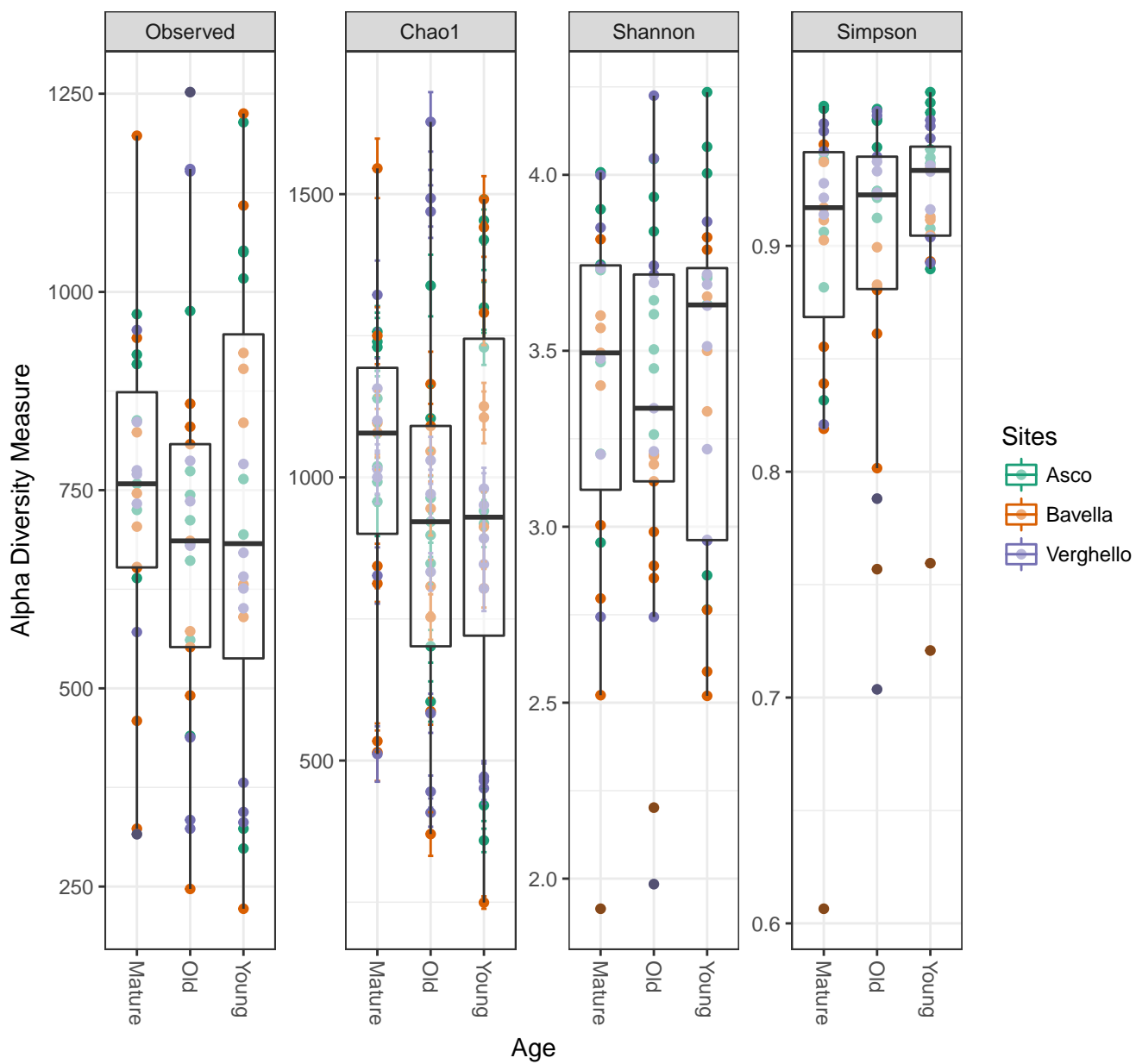


Figure 5.5: Diversity in function of tree age. Color represent sites.

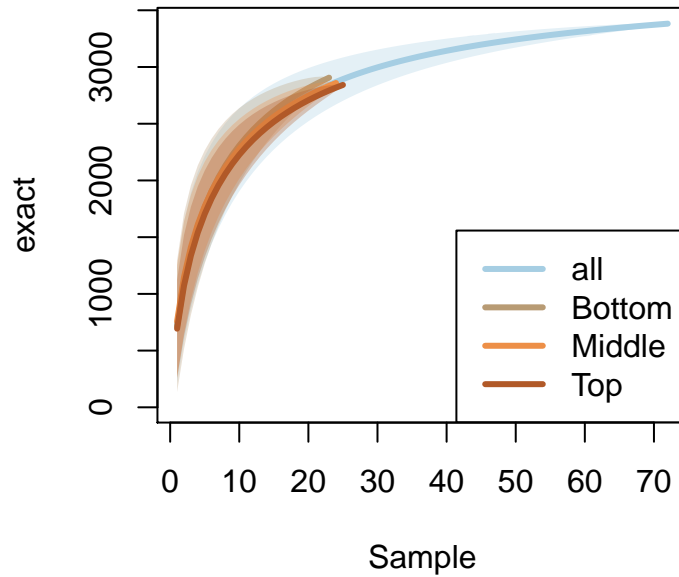


Figure 5.6: Rarefaction curves for each elevation. Notes that if singletons were removed, these curves are biased.

```
ggplot(data = df) + geom_linerange(aes(ymax = xSup, ymin = xInf, x = y), size = 2) +
  geom_point(aes(x=y, y=x), size=4, shape=21, fill="white") +
  coord_flip() + theme_gray() + geom_hline(yintercept = 0) +
  ylab("Differences in mean levels") + xlab("")
```

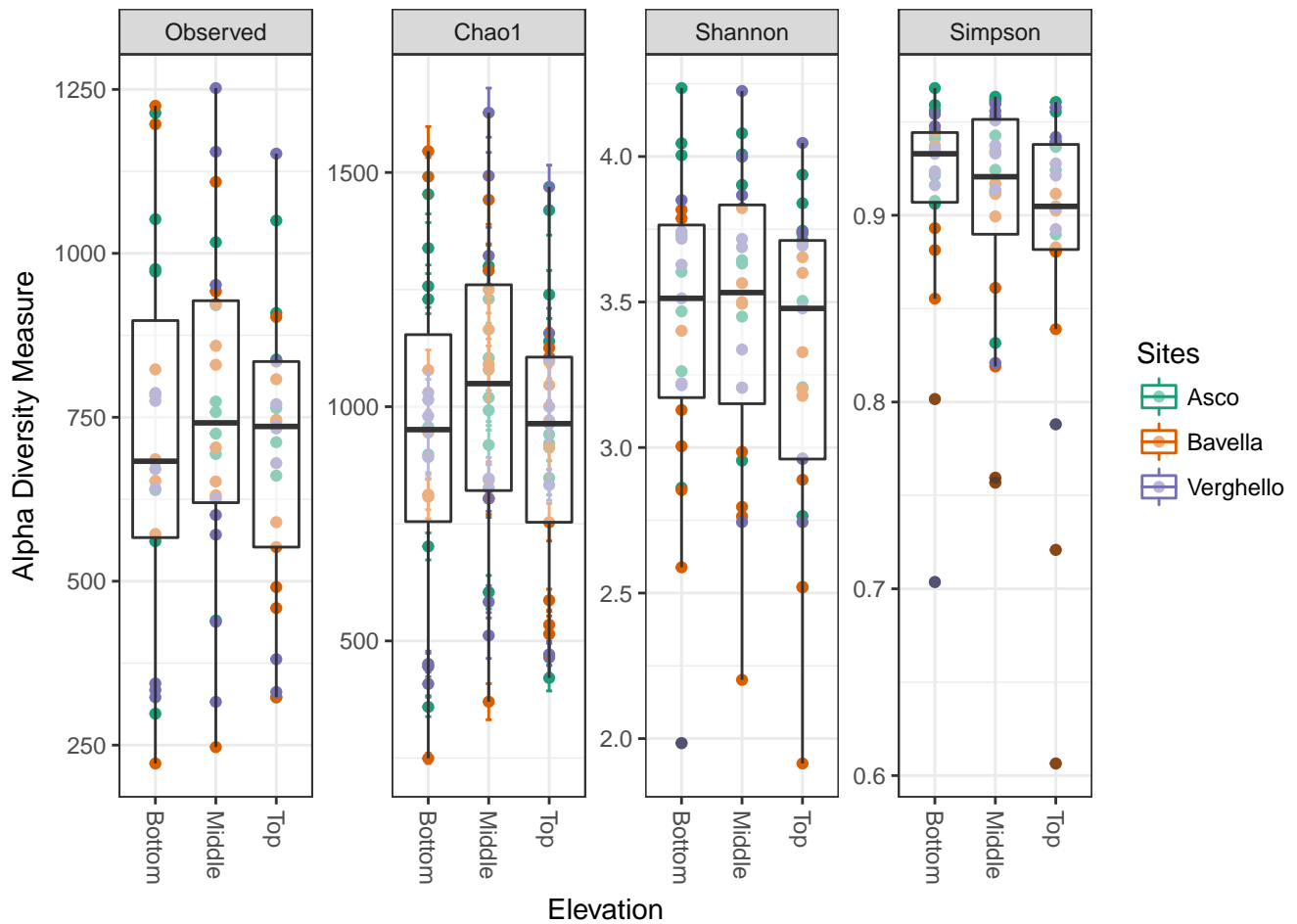


Figure 5.7: Diversity in function of elevation. Color represent sites.

| | Estimate | Std. Error | t value | Pr(> t) |
|-----------------------------------|-------------|------------|------------|-----------|
| (Intercept) | 13.8331728 | 8.1620095 | 1.6948244 | 0.0949697 |
| sqrt(readNumbers) | 0.0784018 | 0.0207642 | 3.7758197 | 0.0003517 |
| data.f3@sam_data\$SitesBavella | -13.4455488 | 3.6312331 | -3.7027501 | 0.0004463 |
| data.f3@sam_data\$SitesVerghello | -4.3814745 | 3.6358586 | -1.2050729 | 0.2326122 |
| data.f3@sam_data\$AgeOld | -0.6822508 | 3.6158530 | -0.1886832 | 0.8509381 |
| data.f3@sam_data\$AgeYoung | -1.1411514 | 3.7025613 | -0.3082059 | 0.7589265 |
| data.f3@sam_data\$ElevationMiddle | 2.4734946 | 3.6617861 | 0.6754886 | 0.5017988 |
| data.f3@sam_data\$ElevationTop | -2.1846248 | 3.6121340 | -0.6048017 | 0.5474492 |

Table 5: Summary of the linear model of the exponential of Shannon’s entropy index (Hill number with $q = 1$)

| | Estimate | Std. Error | t value | Pr(> t) |
|-----------------------------------|------------|------------|------------|-----------|
| (Intercept) | 6.7623603 | 3.4389473 | 1.9664042 | 0.0535919 |
| sqrt(readNumbers) | 0.0321194 | 0.0087487 | 3.6713377 | 0.0004940 |
| data.f3@sam_data\$SitesBavella | -7.5793958 | 1.5299687 | -4.9539547 | 0.0000056 |
| data.f3@sam_data\$SitesVerghello | -2.3673181 | 1.5319176 | -1.5453299 | 0.1271973 |
| data.f3@sam_data\$AgeOld | -0.0838357 | 1.5234885 | -0.0550288 | 0.9562870 |
| data.f3@sam_data\$AgeYoung | 0.5107140 | 1.5600219 | 0.3273762 | 0.7444518 |
| data.f3@sam_data\$ElevationMiddle | 0.5160196 | 1.5428418 | 0.3344605 | 0.7391257 |
| data.f3@sam_data\$ElevationTop | -1.7834158 | 1.5219216 | -1.1718185 | 0.2456139 |

Table 6: Summary of the linear model of inverse of Simpson’s concentration index (Hill number with $q = 2$)

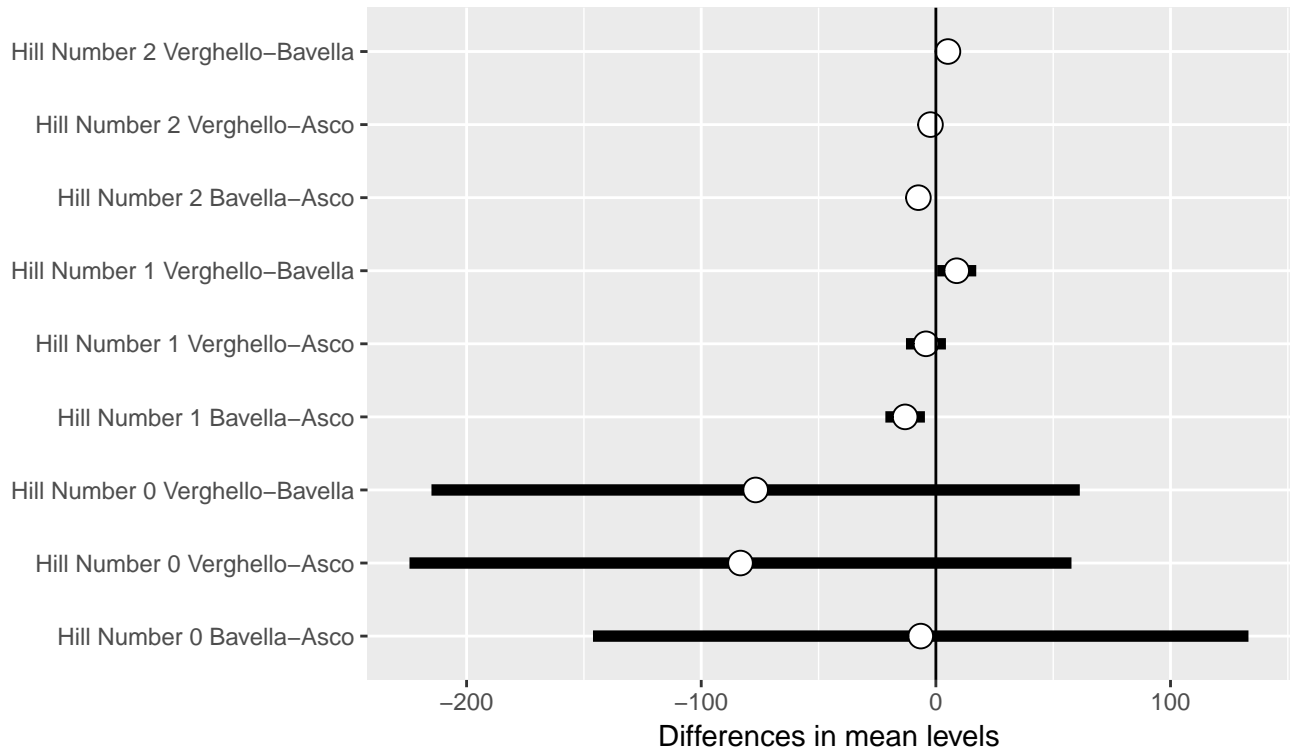


Figure 5.8: Results of the Tukey HSD testing for differences in mean Hill numbers among pairs of modalities

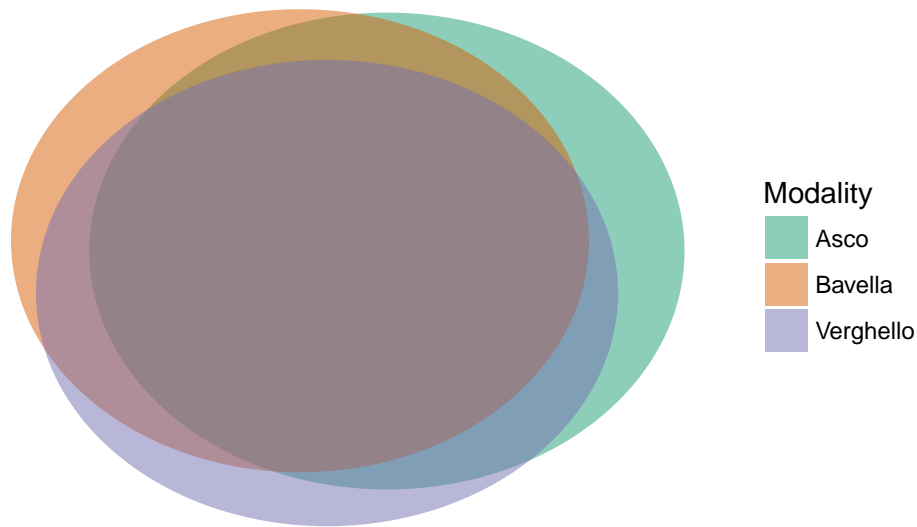


Figure 6.1: Venn diagramm of the distribution of OTUs among Sites

6 Effect of site, age and elevation on fungal endophytic beta-diversity

6.1 Venn diagramm

```
venn_phyloseq(data.f3, "Sites", printValues = F)
```

```
venn_phyloseq(data.f3, "Age", printValues = F)
```

```
venn_phyloseq(data.f3, "Elevation", printValues = F)
```

6.2 Venn diagramm for OTUs present in at least 3 samples

```
data.f3_3samp <- subset_taxa(data.f3, rowSums(data.f3@otu_table>0)>2)
venn_phyloseq(data.f3_3samp, "Sites", printValues = F)
```

```
venn_phyloseq(data.f3_3samp, "Age", printValues = F)
```

```
venn_phyloseq(data.f3_3samp, "Elevation", printValues = F)
```

6.3 Ordination

Ordination of the OTUs table using NMDS (Non-metric MultiDimensionstional Scaling).

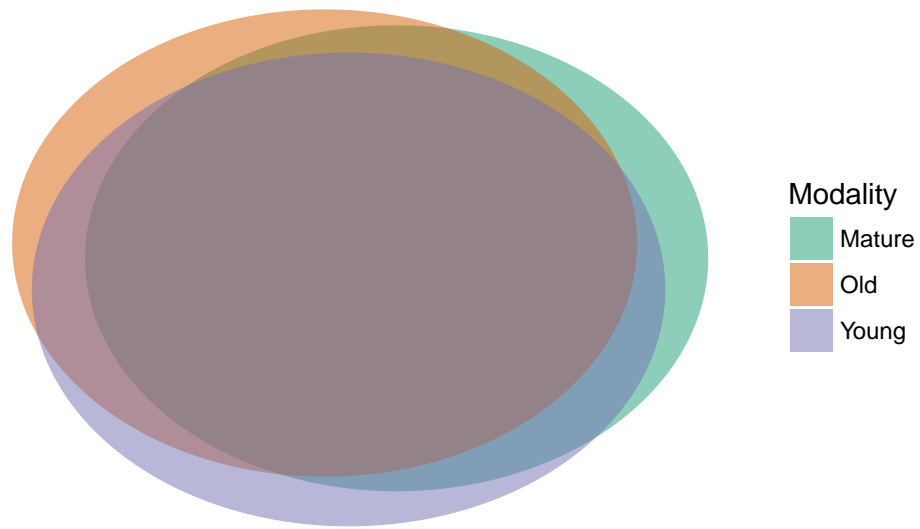


Figure 6.2: Venn diagramm of the distribution of OTUs among host age

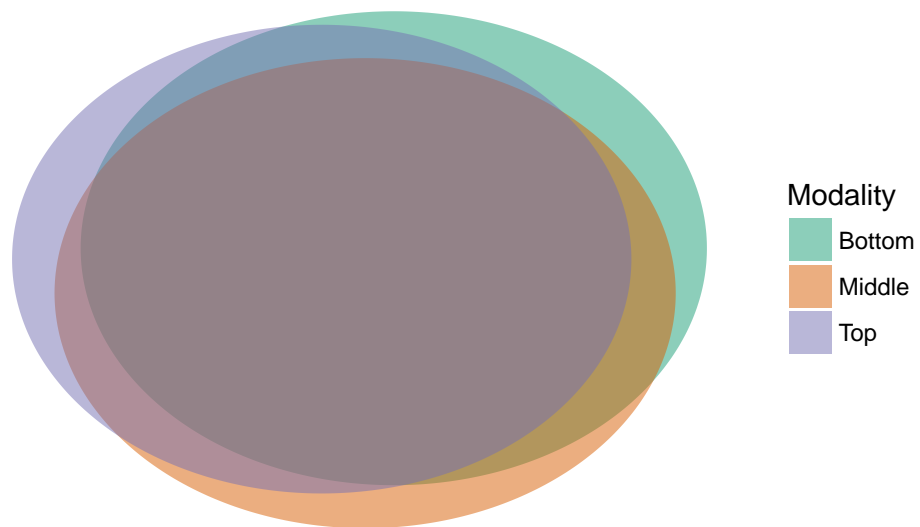


Figure 6.3: Venn diagramm of the distribution of OTUs among elevation of samples

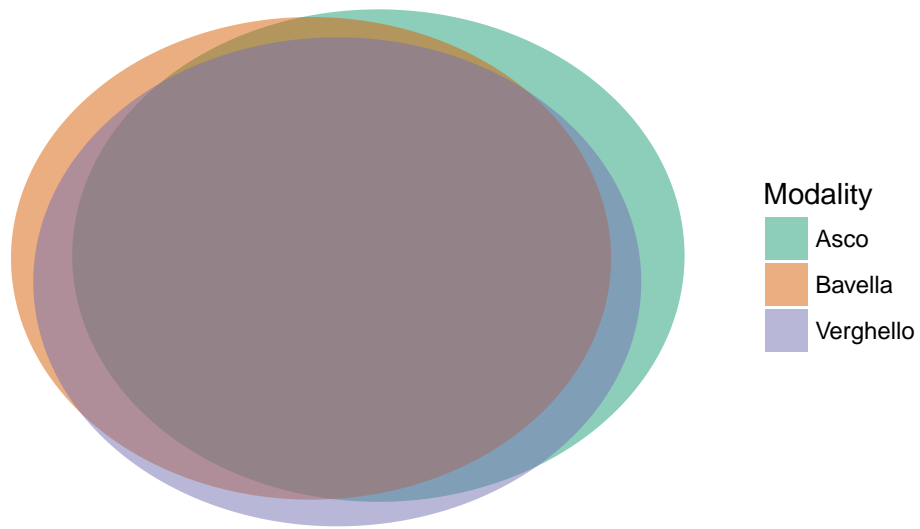


Figure 6.4: Venn diagramm of the distribution of OTUs among Sites

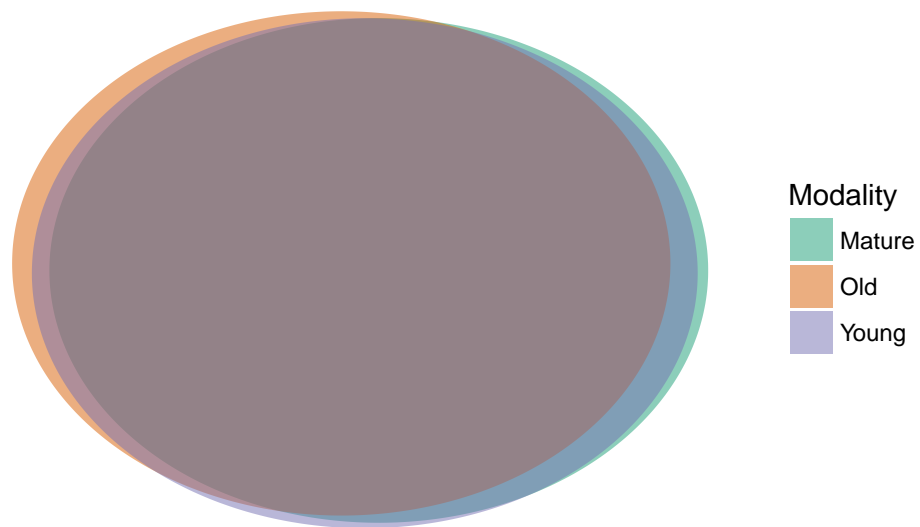


Figure 6.5: Venn diagramm of the distribution of OTUs among host age

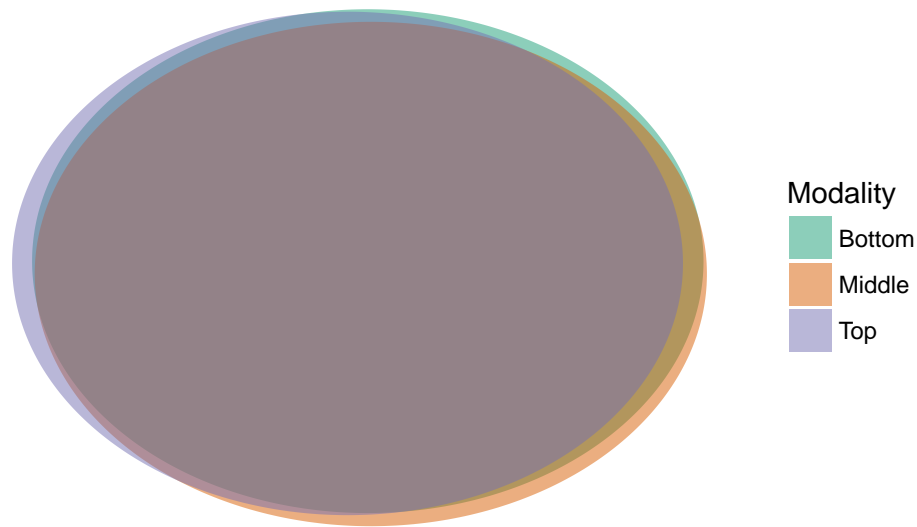


Figure 6.6: Venn diagramm of the distribution of OTUs among elevation of samples whitin the tree

```
my.ord.nmnds <- ordinate(data.f3, method = "NMDS")
my.ord.nmnds$stress
```

```
stressplot(my.ord.nmnds)
```

```
p <- plot_ordination(data.f3, my.ord.nmnds, color = "Sites", shape = "Age")
p + geom_point(size = 0) +
  geom_text(aes(x = p$data$NMDS1, y = p$data$NMDS2,
    label = as.character(as.vector(data.f3$sam_data[, "CODE"]$CODE))))
```

```
my.ord.nmnds_gower <- ordinate(data.f3, distance = "gower", method = "NMDS")
```

```
## Square root transformation
## Wisconsin double standardization
## Run 0 stress 0.2076191
## Run 1 stress 0.231293
## Run 2 stress 0.2076587
## ... Procrustes: rmse 0.02394137 max resid 0.1392384
## Run 3 stress 0.2068915
## ... New best solution
## ... Procrustes: rmse 0.02281764 max resid 0.132568
## Run 4 stress 0.2155622
## Run 5 stress 0.2085022
## Run 6 stress 0.211227
## Run 7 stress 0.2095934
## Run 8 stress 0.2090743
## Run 9 stress 0.206886
## ... New best solution
## ... Procrustes: rmse 0.001462729 max resid 0.008040082
```



```

## ... Similar to previous best
## Run 10 stress 0.216301
## Run 11 stress 0.2119772
## Run 12 stress 0.2071656
## ... Procrustes: rmse 0.007344136  max resid 0.04991657
## Run 13 stress 0.2121752
## Run 14 stress 0.2086672
## Run 15 stress 0.2164855
## Run 16 stress 0.2132907
## Run 17 stress 0.2128912
## Run 18 stress 0.2119254
## Run 19 stress 0.2084085
## Run 20 stress 0.2121553
## *** Solution reached

my.ord.PCoA <- ordinate(data.f3, method = "PCoA")
my.ord.PCoA_gower <- ordinate(data.f3, distance = "gower", method = "PCoA")
my.ord.DCA <- ordinate(data.f3, method = "DCA")
my.ord.DCA_gower <- ordinate(data.f3, distance = "gower", method = "DCA")

p_NMDS_BRAY <- plot_ordination(data.f3, my.ord.nmds, color = "Sites",
                              shape = "Age", label = "CODE") + geom_point(size = 5)
p_NMDS_GOWER <- plot_ordination(data.f3, my.ord.nmds_gower, color = "Sites",
                              shape = "Age", label = "CODE") + geom_point(size = 5)
p_PCoA_BRAY <- plot_ordination(data.f3, my.ord.PCoA, color = "Sites",
                              shape = "Age", label = "CODE") + geom_point(size = 5)
p_PCoA_GOWER <- plot_ordination(data.f3, my.ord.PCoA_gower, color = "Sites",
                              shape = "Age", label = "CODE") + geom_point(size = 5)
p_DCA_BRAY <- plot_ordination(data.f3, my.ord.DCA, color = "Sites",
                              shape = "Age", label = "CODE") + geom_point(size = 5)
p_DCA_GOWER <- plot_ordination(data.f3, my.ord.DCA_gower, color = "Sites",
                              shape = "Age", label = "CODE") + geom_point(size = 5)

```

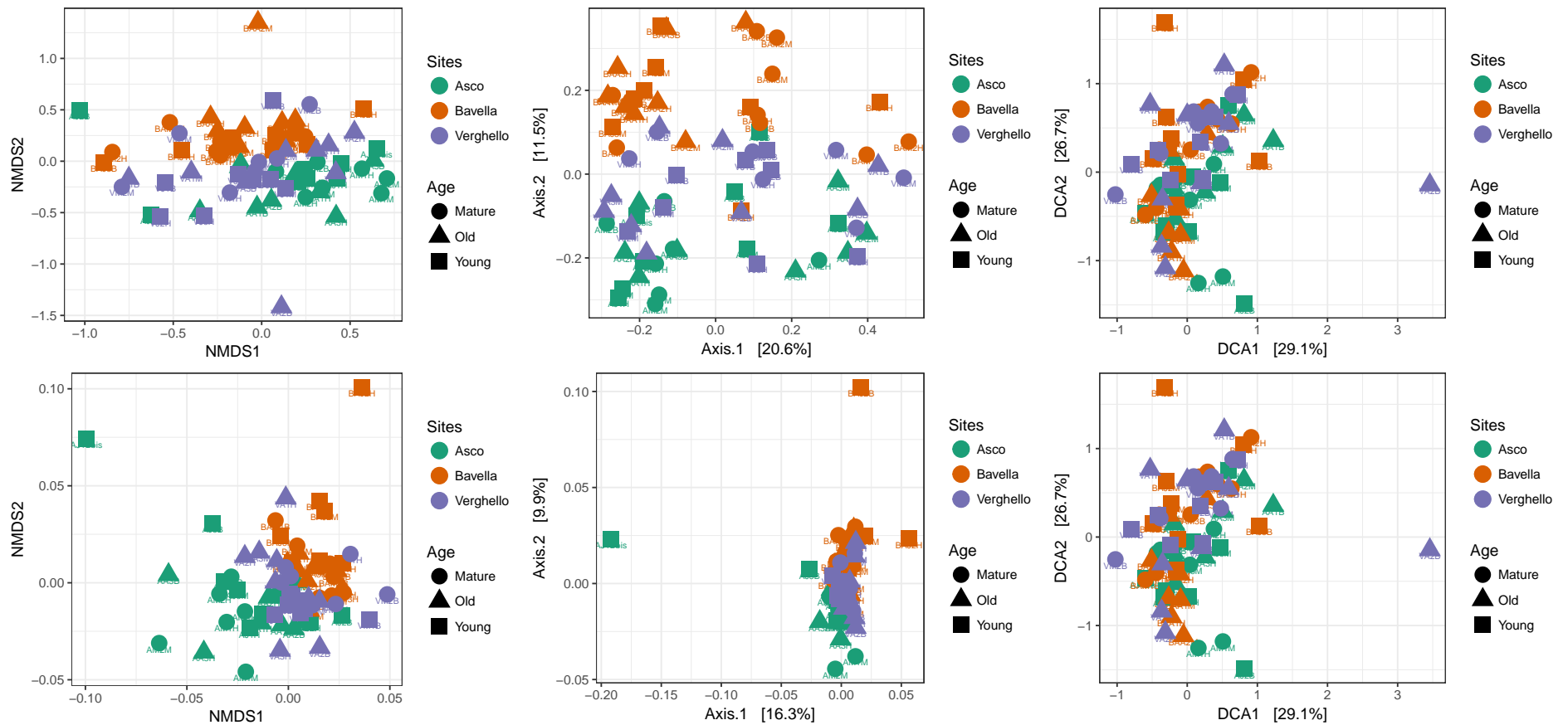


Figure 6.9: Comparison of different distances (bray (up) and gower (bottom)) and ordination methods (NMDS (left), PCoA (center) and DCA (right)).

```
multiplot(p_NMDS_BRAY, p_NMDS_GOWER, p_PCoA_BRAY, p_PCoA_GOWER, p_DCA_BRAY, p_DCA_GOWER,
          cols = 3)
```

6.4 Permanova on sites, host ages and elevation

```
sam_data <- as.data.frame(unclass(data.f3@sam_data))
sam_data$IndividualTree <- paste(sam_data$Sites, sam_data$Age, sam_data$Tree)
res.ado <- adonis(t(data.f3@otu_table) ~ Sites * Age * Elevation, sam_data,
  permutation = 9999)
```

If we only keep the 593 OTUs present in more than 30 sample, the Permanova results is the following:

```
res.ado_sampMin30 <- adonis(t(data.f3@otu_table[rowSums(data.f3@otu_table>0)>=30,]) ~
  Sites * Age * Elevation, sam_data, permutation = 9999)
```

```
data.f3_without_C_minus <- subset_taxa(data.f3, taxa_names(data.f3)!="OTU_1")
res.ado_without_C_minus <- adonis(t(data.f3_without_C_minus@otu_table) ~
  Sites * Age * Elevation, sam_data, permutation = 9999)
```

```
xtable(res.ado$aov.tab, caption = "Result of the permanova on abundances
(number of sequence).")
```

| | Df | SumsOfSqs | MeanSqs | F.Model | R2 | Pr(>F) |
|---------------------|----|-----------|---------|---------|------|--------|
| Sites | 2 | 2.05 | 1.02 | 4.52 | 0.11 | 0.0001 |
| Age | 2 | 0.71 | 0.35 | 1.56 | 0.04 | 0.0334 |
| Elevation | 2 | 0.60 | 0.30 | 1.33 | 0.03 | 0.1078 |
| Sites:Age | 4 | 1.67 | 0.42 | 1.84 | 0.09 | 0.0009 |
| Sites:Elevation | 4 | 0.90 | 0.22 | 0.99 | 0.05 | 0.4675 |
| Age:Elevation | 4 | 1.15 | 0.29 | 1.27 | 0.06 | 0.0902 |
| Sites:Age:Elevation | 8 | 1.96 | 0.25 | 1.09 | 0.10 | 0.2559 |
| Residuals | 45 | 10.18 | 0.23 | | 0.53 | |
| Total | 71 | 19.21 | | | 1.00 | |

Table 7: Result of the permanova on abundances (number of sequence).

```
xtable(res.ado_sampMin30$aov.tab, caption = "Result of the permanova on abundances
(number of sequence) using only OTUs present in more than 30 samples")
```

```
res.ado_bin <- adonis(t(as.binaryOtuTable(data.f3@otu_table) ~ Sites * Age *
  Elevation, sam_data, permutation = 9999)
```

```
xtable(res.ado_bin$aov.tab, caption = "Result of the permanova on OTUs
(each OTU is representing by one sequence)).")
```

6.5 Permanova on sites, host ages and individual trees

| | Df | SumsOfSqs | MeanSqs | F.Model | R2 | Pr(>F) |
|---------------------|----|-----------|---------|---------|------|--------|
| Sites | 2 | 2.01 | 1.00 | 4.69 | 0.11 | 0.0001 |
| Age | 2 | 0.68 | 0.34 | 1.59 | 0.04 | 0.0383 |
| Elevation | 2 | 0.59 | 0.29 | 1.37 | 0.03 | 0.1100 |
| Sites:Age | 4 | 1.63 | 0.41 | 1.90 | 0.09 | 0.0010 |
| Sites:Elevation | 4 | 0.86 | 0.21 | 1.00 | 0.05 | 0.4627 |
| Age:Elevation | 4 | 1.12 | 0.28 | 1.30 | 0.06 | 0.0884 |
| Sites:Age:Elevation | 8 | 1.87 | 0.23 | 1.09 | 0.10 | 0.2615 |
| Residuals | 45 | 9.63 | 0.21 | | 0.52 | |
| Total | 71 | 18.38 | | | 1.00 | |

Table 8: Result of the permanova on abundances (number of sequence) using only OTUs present in more than 30 samples

| | Df | SumsOfSqs | MeanSqs | F.Model | R2 | Pr(>F) |
|---------------------|----|-----------|---------|---------|------|--------|
| Sites | 2 | 1.05 | 0.52 | 3.72 | 0.09 | 0.0001 |
| Age | 2 | 0.45 | 0.22 | 1.58 | 0.04 | 0.0082 |
| Elevation | 2 | 0.29 | 0.15 | 1.04 | 0.03 | 0.3633 |
| Sites:Age | 4 | 0.84 | 0.21 | 1.49 | 0.08 | 0.0025 |
| Sites:Elevation | 4 | 0.48 | 0.12 | 0.84 | 0.04 | 0.8859 |
| Age:Elevation | 4 | 0.58 | 0.14 | 1.02 | 0.05 | 0.4056 |
| Sites:Age:Elevation | 8 | 1.06 | 0.13 | 0.94 | 0.10 | 0.7053 |
| Residuals | 45 | 6.35 | 0.14 | | 0.57 | |
| Total | 71 | 11.09 | | | 1.00 | |

Table 9: Result of the permanova on OTUs (each OTU is representing by one sequence)).

```
res.ado_Tree <- adonis(t(data.f3@otu_table) ~ Sites*Age + Sites:Age:IndividualTree ,
                      sam_data, permutation = 9999)
res.ado_Tree_sampMin30 <- adonis(t(data.f3@otu_table[rowSums(data.f3@otu_table>0)>=30,]) ~
                                Sites*Age + Sites:Age:IndividualTree , sam_data,
                                permutation = 9999)
res.ado_Tree_bin <- adonis(t(as.binaryOtuTable(data.f3@otu_table) ~
                              Sites*Age + Sites:Age:IndividualTree , sam_data,
                              permutation = 9999)
```

```
xtable(res.ado_Tree$aov.tab, caption = "Result of the permanova on abundances
(number of sequence).")
```

```
xtable(res.ado_Tree_sampMin30$aov.tab, caption = "Result of the permanova on abundances
(number of sequence) using only OTUs present in more than 30 samples")
```

```
xtable(res.ado_Tree_bin$aov.tab, caption = "Result of the permanova on OTUs
(each OTU is representing by one sequence)).")
```

6.6 Differences in abundances and OTUs number by Order.

| | Df | SumsOfSqs | MeanSqs | F.Model | R2 | Pr(>F) |
|--------------------------|----|-----------|---------|---------|------|--------|
| Sites | 2 | 2.05 | 1.02 | 4.91 | 0.11 | 0.0001 |
| Age | 2 | 0.71 | 0.35 | 1.69 | 0.04 | 0.0204 |
| Sites:Age | 4 | 1.68 | 0.42 | 2.02 | 0.09 | 0.0003 |
| Sites:Age:IndividualTree | 18 | 5.40 | 0.30 | 1.44 | 0.28 | 0.0001 |
| Residuals | 45 | 9.38 | 0.21 | | 0.49 | |
| Total | 71 | 19.21 | | | 1.00 | |

Table 10: Result of the permanova on abundances (number of sequence).

| | Df | SumsOfSqs | MeanSqs | F.Model | R2 | Pr(>F) |
|--------------------------|----|-----------|---------|---------|------|--------|
| Sites | 2 | 2.01 | 1.00 | 5.10 | 0.11 | 0.0001 |
| Age | 2 | 0.68 | 0.34 | 1.73 | 0.04 | 0.0225 |
| Sites:Age | 4 | 1.64 | 0.41 | 2.09 | 0.09 | 0.0004 |
| Sites:Age:IndividualTree | 18 | 5.20 | 0.29 | 1.47 | 0.28 | 0.0002 |
| Residuals | 45 | 8.85 | 0.20 | | 0.48 | |
| Total | 71 | 18.38 | | | 1.00 | |

Table 11: Result of the permanova on abundances (number of sequence) using only OTUs present in more than 30 samples

| | Df | SumsOfSqs | MeanSqs | F.Model | R2 | Pr(>F) |
|--------------------------|----|-----------|---------|---------|------|--------|
| Sites | 2 | 1.05 | 0.52 | 4.18 | 0.09 | 0.0001 |
| Age | 2 | 0.45 | 0.22 | 1.78 | 0.04 | 0.0026 |
| Sites:Age | 4 | 0.85 | 0.21 | 1.70 | 0.08 | 0.0001 |
| Sites:Age:IndividualTree | 18 | 3.10 | 0.17 | 1.37 | 0.28 | 0.0001 |
| Residuals | 45 | 5.65 | 0.13 | | 0.51 | |
| Total | 71 | 11.09 | | | 1.00 | |

Table 12: Result of the permanova on OTUs (each OTU is representing by one sequence)).

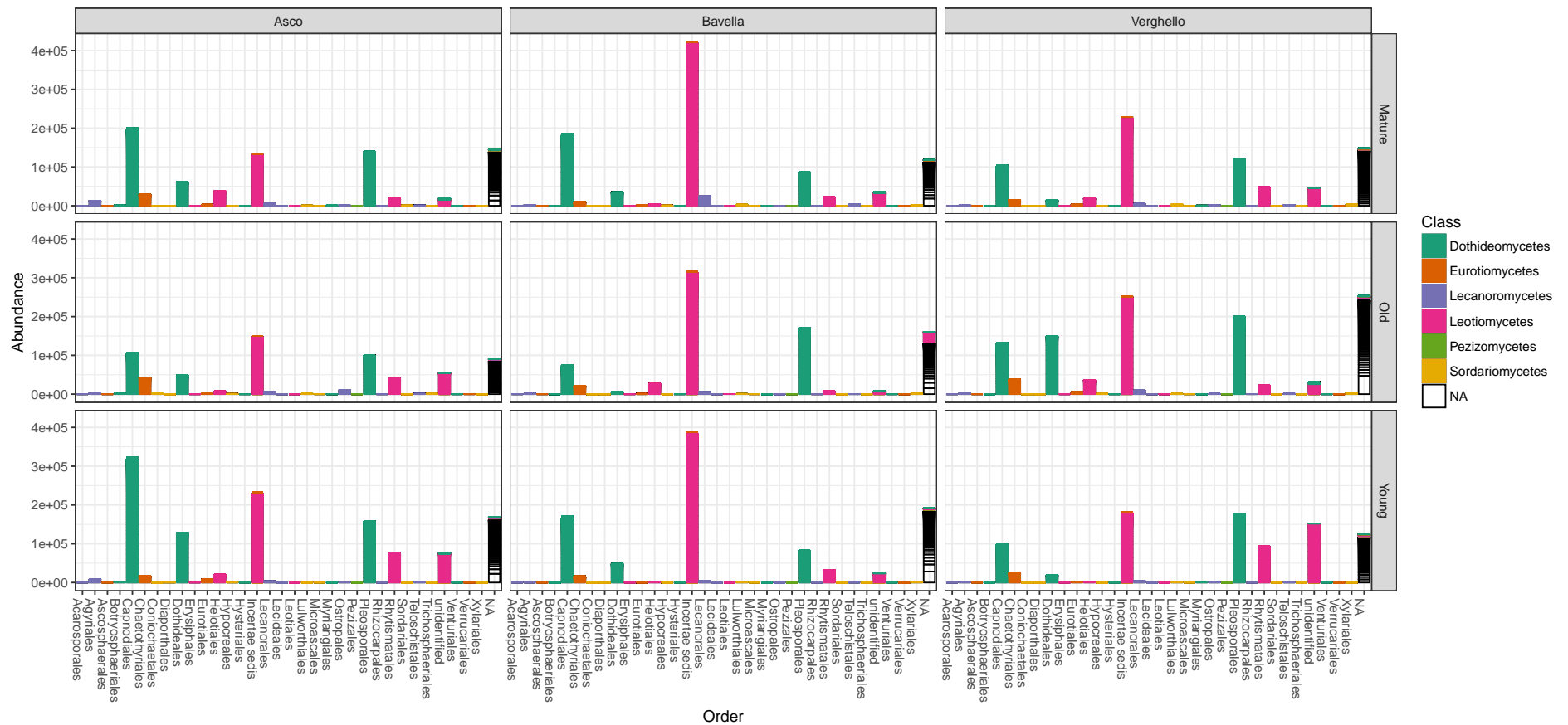


Figure 6.10: Taxonomic distribution of sequences in the different site * age combinaison.

```
data.f3_taxo_known <- prune_taxa(taxa_names(data.f3@tax_table)
  [!(data.f3@tax_table[, 4] %in% c("unidentified", "Incertae sedis"))], data.f3)
p <- plot_bar(data.f3_taxo_known, "Order", fill = "Class", facet_grid = Age ~ Sites)
p + geom_bar(aes(color = Class, fill = Class), stat = "identity", position = "stack")
```

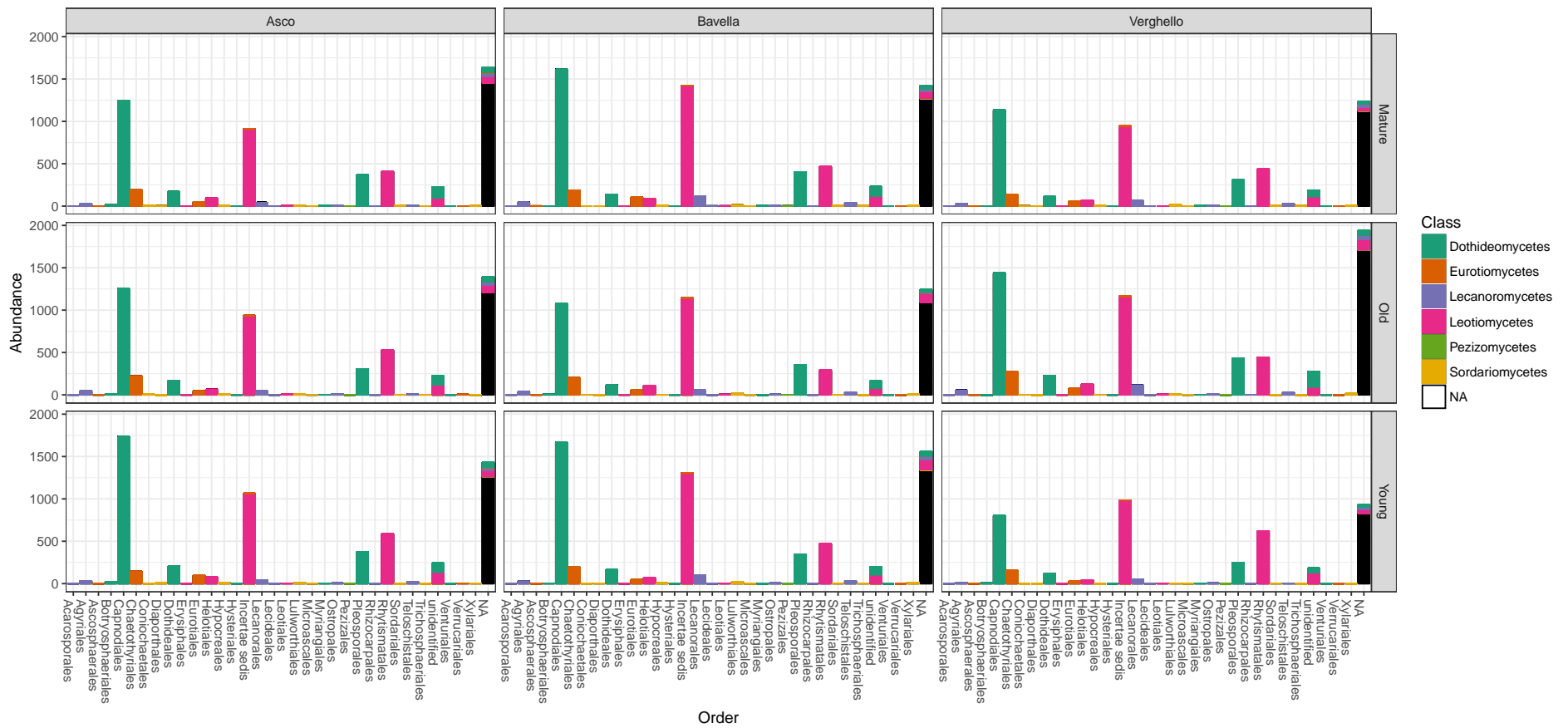


Figure 6.11: Taxonomic distribution of OTUs in the different site * age combinaison.

```
p <- plot_bar(as.binaryOtuTable(data.f3_taxo_known), "Order", fill = "Class",
             facet_grid = Age ~ Sites)
p + geom_bar(aes(color = Class, fill = Class), stat = "identity", position = "stack")
```



Figure 6.12: OTUs significantly different in terms of abundances between Verghello (positive values) and Asco (negative values)

6.7 Differences in abundances for each OTUs

6.7.1 Pairwise comparison of the OTUs composition by sites

```
library("DESeq2")
packageVersion("DESeq2")

## [1] '1.16.1'

data.f3_deseq2 <- phyloseq_to_deseq2(data.f3, ~ Sites)
data.f3_deseq2 <- DESeq(data.f3_deseq2, test = "Wald", fitType = "parametric")
res.f3_deseq2 <- results(data.f3_deseq2)
```

```
res_VA <- plot_deseq2_phyloseq(data.f3_deseq2, tax_table = data.f3@tax_table,
                               contrast = c("Sites", "Verghello", "Asco"),
                               taxa = "Species", color_tax = "Class")
res_VA
```

```
res_VB <- plot_deseq2_phyloseq(data.f3_deseq2, tax_table = data.f3@tax_table,
                               contrast = c("Sites", "Verghello", "Bavella"),
                               taxa = "Species", color_tax = "Class")
res_VB
```

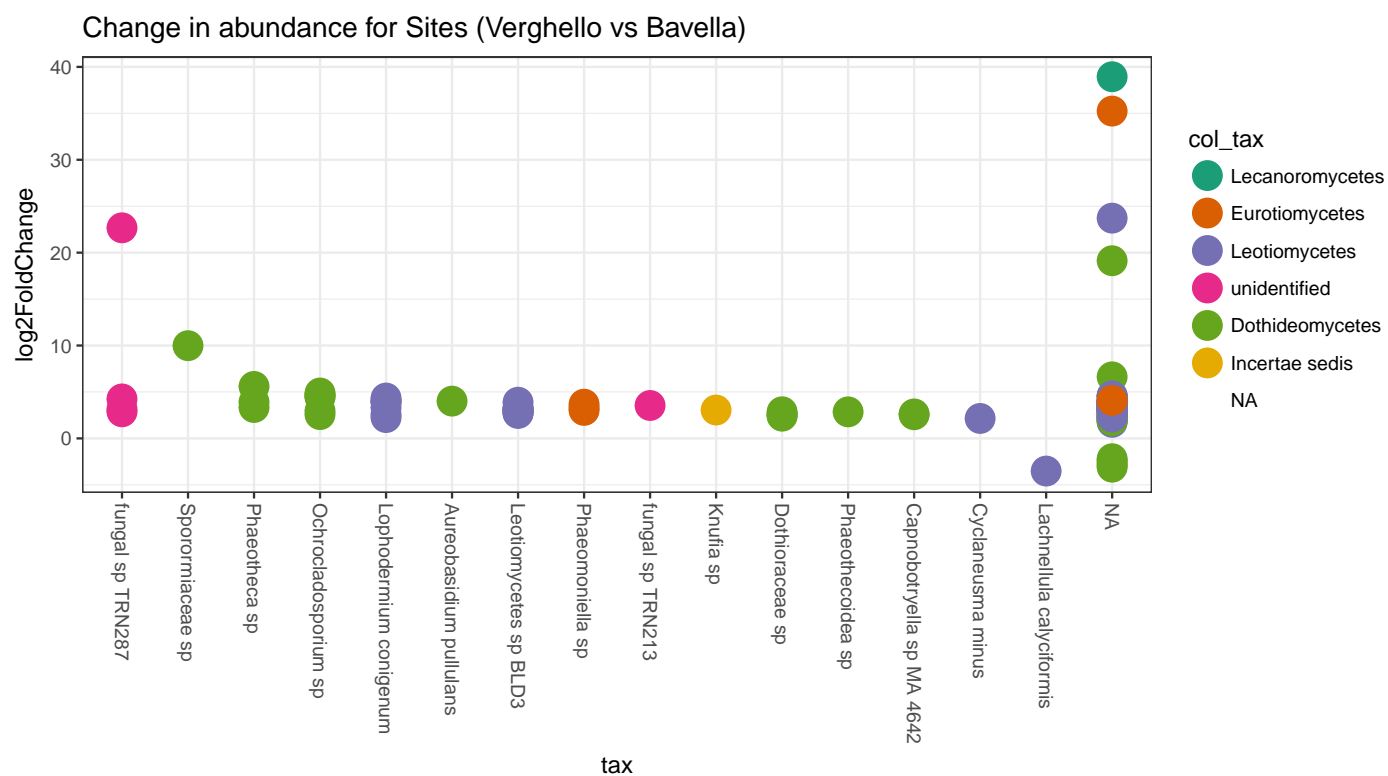



Figure 6.13: OTUs significantly different in terms of abundances between Verghello (positive values) and Bavella (negative values)

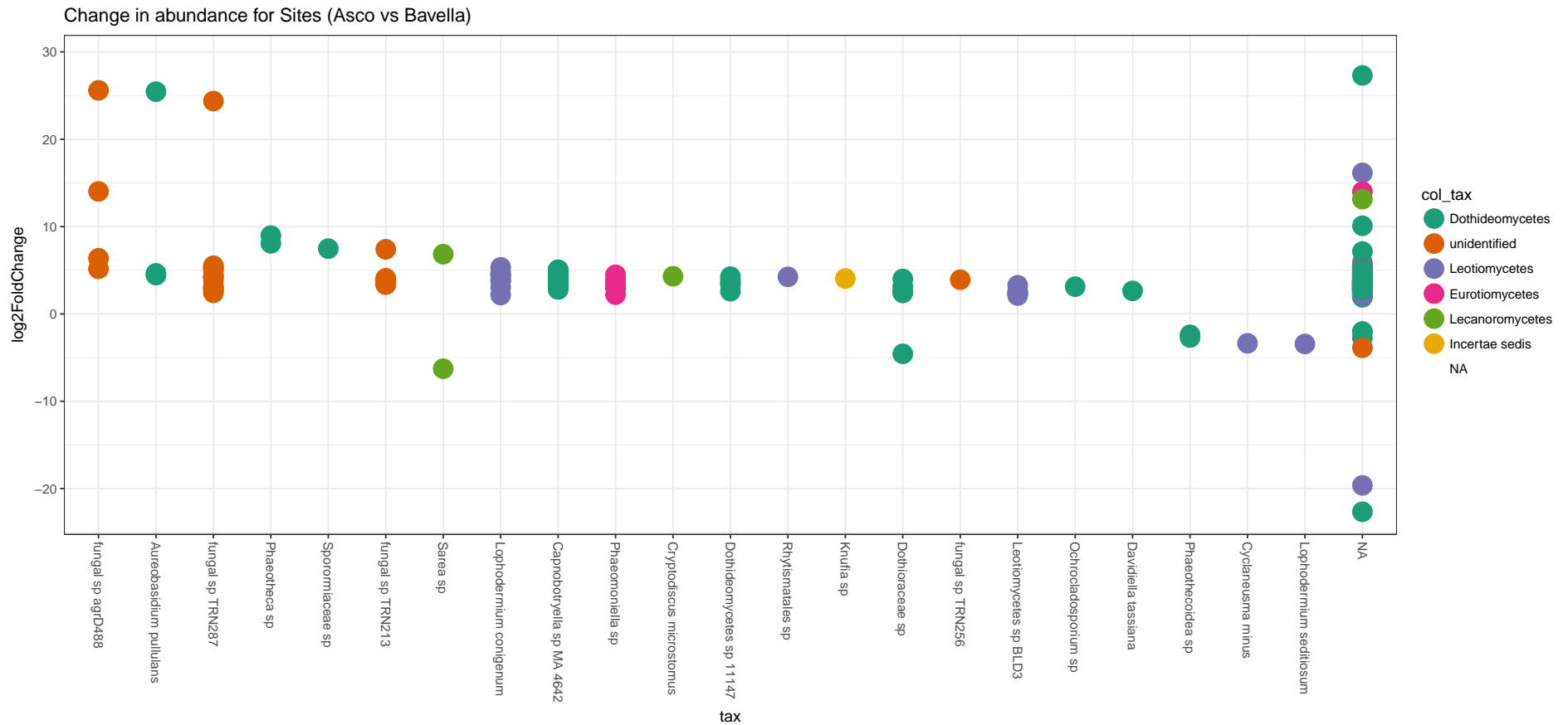


Figure 6.14: OTUs significantly different in terms of abundances between Asco (positive values) and Bavella (negative values)

```
res_AB <- plot_deseq2_phyloseq(data.f3_deseq2, tax_table = data.f3@tax_table,
                               contrast = c("Sites", "Asco", "Bavella"),
                               taxa = "Species", color_tax = "Class")
res_AB
```

6.7.2 Pairwise comparison of Order composition by sites

```
res_VA_o <- plot_deseq2_phyloseq(data.f3, contrast = c("Sites", "Verghello", "Asco"),
                                taxDepth = "Order", color_tax = "Class")
res_VB_o <- plot_deseq2_phyloseq(data.f3, contrast = c("Sites", "Verghello", "Bavella"),
                                taxDepth = "Order", color_tax = "Class")
res_AB_o <- plot_deseq2_phyloseq(data.f3, contrast = c("Sites", "Asco", "Bavella"),
                                taxDepth = "Order", color_tax = "Class")
```

6.8 Distribution of OTUs abundance in the fungal phylogeny

```
library("cluster")
library("phytools")

## Loading required package: maps
##
## Attaching package: 'maps'
## The following object is masked from 'package:plyr':
##
##   ozone
## The following object is masked from 'package:cluster':
##
##   votes.repub

data.f3_interm <- data.f3
data.f3_interm@otu_table <- otu_table(apply(data.f3@otu_table, 2, function(x) tapply(x, as.factor(data.f3@tax_table[, "Order"]), FUN = function(y) sum(y > 0))), MARGIN = 2)
data.f3_interm@tax_table <- tax_table(apply(data.f3@tax_table, 2, function(x) tapply(x, as.factor(data.f3@tax_table[, "Order"]), FUN = function(y) sum(y > 0))), MARGIN = 2)
data.f3_interm@refseq <- NULL

data.f3_interm <- subset_taxa(data.f3_interm, !grepl("uncultured", data.f3_interm@tax_table[, "Species"]))
data.f3_interm <- subset_taxa(data.f3_interm, !grepl("sp$", data.f3_interm@tax_table[, "Species"]))
data.f3_interm <- subset_taxa(data.f3_interm, !grepl("unidentified", data.f3_interm@tax_table[, "Family"]))
data.f3_interm <- subset_taxa(data.f3_interm, !grepl("unidentified", data.f3_interm@tax_table[, "Order"]))
data.f3_interm <- subset_taxa(data.f3_interm, !grepl("unidentified", data.f3_interm@tax_table[, "Class"]))
data.f3_interm <- subset_taxa(data.f3_interm, !grepl("Myxotrichaceae", data.f3_interm@tax_table[, "Species"]))
data.f3_interm <- subset_taxa(data.f3_interm, rowSums(data.f3_interm@otu_table) > 100)

tree_tax_interm <- as.data.frame(unclass(data.f3_interm@tax_table))
tree_tax_interm$OTUs <- rownames(tree_tax_interm)

tree_tax_interm <- as.data.frame(replace(as.matrix(tree_tax_interm), which(is.na(tree_tax_interm)), NA))

data.f3_interm@tax_table <- tax_table(as.matrix(tree_tax_interm))

tree_tax_interm$pathString <- paste("Fungi",
                                   tree_tax_interm$Phylum,
                                   tree_tax_interm$Class,
                                   tree_tax_interm$Order,
```

```

        tree_tax_interm$Family,
        tree_tax_interm$Genus,
        tree_tax_interm$OTUs,
        sep = "/" )

write(ToNewick(as.Node(tree_tax_interm, na.rm = TRUE)), file="tree.txt")
tree <- phytools::read.newick(file="tree.txt")
tree <- ape::collapse.singles(tree)

data.f3_interm@phy_tree <- tree
taxa_names(data.f3_interm@phy_tree) <- gsub("_", " ", taxa_names(data.f3_interm@phy_tree))
taxa_names(data.f3_interm@otu_table) <- gsub("_", " ", taxa_names(data.f3_interm@otu_table))
taxa_names(data.f3_interm@tax_table) <- gsub("_", " ", taxa_names(data.f3_interm@tax_table))
taxa_names(data.f3_interm@phy_tree) <- gsub(" ", "", taxa_names(data.f3_interm@phy_tree))
taxa_names(data.f3_interm@otu_table) <- gsub(" ", "", taxa_names(data.f3_interm@otu_table))
taxa_names(data.f3_interm@tax_table) <- gsub(" ", "", taxa_names(data.f3_interm@tax_table))

ptree <- plot_tree(data.f3_interm, color = "Class", shape = "Sites", ladderize = "left", justify = "left")

cond <- gsub(" ", "", rownames(data.f3_interm@otu_table)[rowSums(data.f3_interm@otu_table) >= 1])
df_cond <- as.data.frame(ptree$data)[ptree$data$OTU %in% cond,]
df_cond$Species <- data.f3_interm@tax_table[taxa_names(data.f3_interm) %in% cond, "OTUs"]

cond_Deseq <- levels(df_cond$Species)
df_cond_Deseq <- as.data.frame(ptree$data)[ptree$data$OTU %in% cond_Deseq,]
df_cond_Deseq$Species <- data.f3_interm@tax_table[gsub("_", " ", taxa_names(data.f3_interm)) %in% cond_Deseq, "OTUs"]

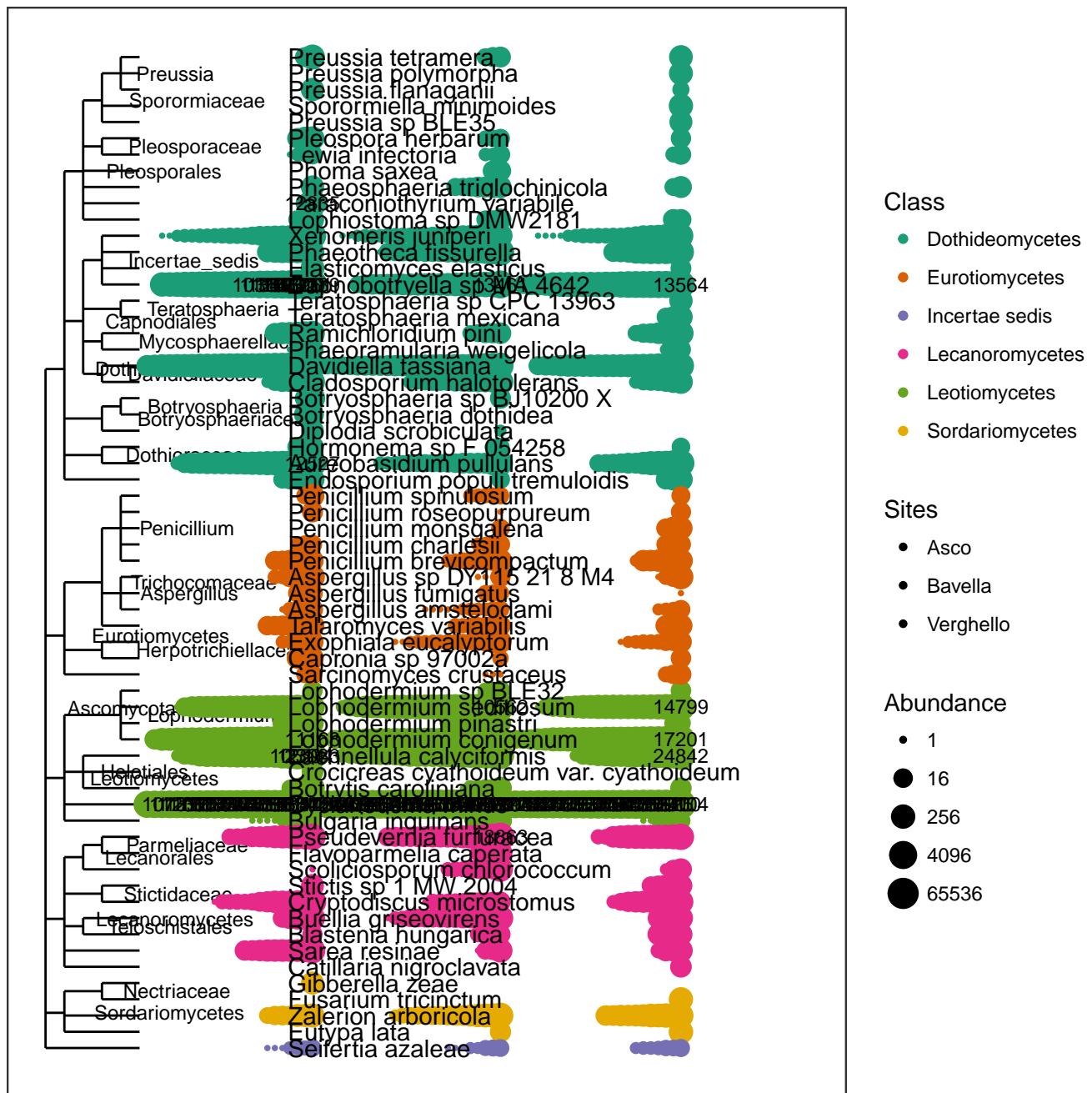
ptree + geom_text(data = df_cond, aes(x = 215, y = y, label = OTU), hjust = "left") + scale_shape_manual(values =
sum(!is.na(match(gsub("_", " ", data.f3@tax_table[, "Species"]), gsub("_", " ", tree$tip.label))))

## [1] 613

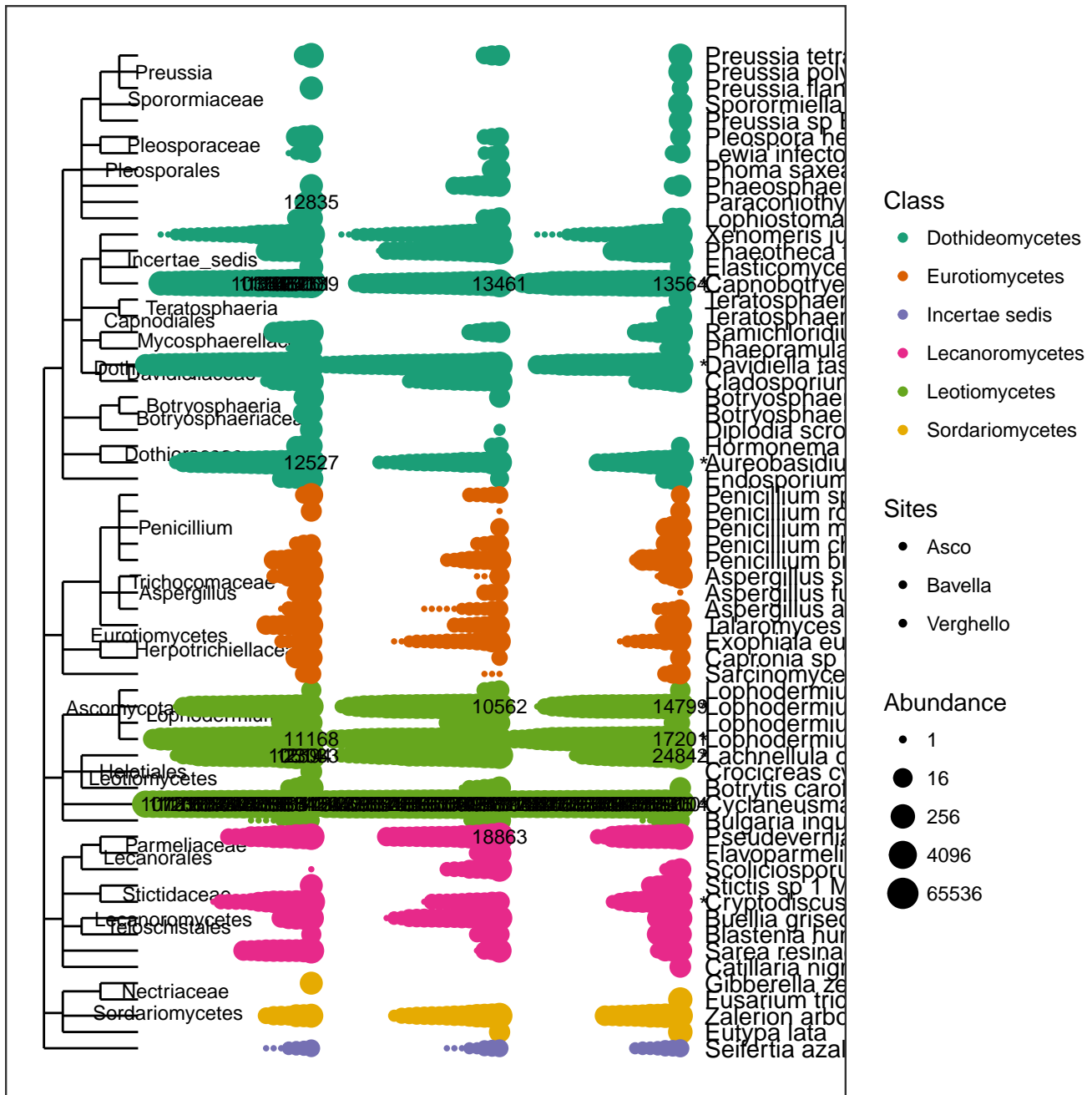
sum(rowSums(data.f3@otu_table)[gsub("_", " ", data.f3@tax_table[, "Species"]) %in% gsub("_", " ", tree$tip.label)])

## [1] 38.44194

```



```
ptree + geom_text(data = df_cond, aes(x = 585, y = y, label = OTU), hjust = "left") + scale_shape_ma
ggsave("phylo_map.pdf", width = 20, height = 15)
```



7 Summary

7.1 Filtering summary

The raw data are made of 8.398038×10^6 sequences representing 4373 OTUs allocated to 80 samples.

After filtering, the dataset includes 8.373567×10^6 sequences representing 3382 OTUs allocated to 72 samples.

7.2 Alpha diversity

Host age and elevation within tree do not impact any aspect of fungal local diversity. Despite similar OTUs richness, Asco is a site more diverse than Verghello and Bavella.

7.3 Beta diversity

Site ($R^2 = 0.106$), age ($R^2 = 0.037$) and interaction age*site ($R^2 = 0.087$) statistically structured the fungal endophytic beta-diversity.

| | Comparison | OTU_names | Species | Class | log2FoldChange (negative = more on second le |
|-----|----------------------|---------------------------------------|---------------------------|-----------------|--|
| 1 | Verghello vs Asco | SH200057.07FU_AY843076_reps | | | -4.0407542541417 |
| 2 | Verghello vs Asco | New.CleanUp.ReferenceOTU1770 | | | -24.6790763314907 |
| 3 | Verghello vs Asco | SH127907.07FU_DQ780388_refs | Davidiella tassiana | Dothideomycetes | -2.38306389114883 |
| 4 | Verghello vs Asco | New.CleanUp.ReferenceOTU5157 | | Dothideomycetes | -3.73720412281358 |
| 5 | Verghello vs Asco | New.CleanUp.ReferenceOTU2469 | | Dothideomycetes | -3.94502965159243 |
| 6 | Verghello vs Asco | New.CleanUp.ReferenceOTU3996 | Capnobotryella sp MA 4642 | Dothideomycetes | -2.3439677843879 |
| 7 | Verghello vs Asco | New.CleanUp.ReferenceOTU5258 | | | -3.49945135093844 |
| 8 | Verghello vs Asco | New.CleanUp.ReferenceOTU2018 | | | -2.49803884657837 |
| 9 | Verghello vs Asco | New.CleanUp.ReferenceOTU1461 | | Dothideomycetes | -3.06836744519646 |
| 10 | Verghello vs Asco | New.ReferenceOTU14 | | | -2.16563566468442 |
| 11 | Verghello vs Asco | New.CleanUp.ReferenceOTU1023 | | | -2.24952094123825 |
| 12 | Verghello vs Asco | New.ReferenceOTU69 | | | 3.44297543431105 |
| 13 | Verghello vs Asco | New.CleanUp.ReferenceOTU2768 | Dothideomycetes sp 11147 | Dothideomycetes | -3.56819089741535 |
| 14 | Verghello vs Asco | New.CleanUp.ReferenceOTU4452 | Phaeomoniella sp | Eurotiomycetes | -3.17267563744643 |
| 15 | Verghello vs Asco | New.ReferenceOTU128 | | | 2.891473978103 |
| 16 | Verghello vs Asco | New.CleanUp.ReferenceOTU2140 | | Dothideomycetes | -2.87782099181598 |
| 17 | Verghello vs Asco | New.ReferenceOTU83 | Capnobotryella sp MA 4642 | Dothideomycetes | -2.6675117357848 |
| 18 | Verghello vs Asco | New.CleanUp.ReferenceOTU5403 | | | -2.71599552048226 |
| 19 | Verghello vs Asco | New.CleanUp.ReferenceOTU6286 | | Leotiomycetes | -3.37227698939496 |
| 20 | Verghello vs Asco | New.CleanUp.ReferenceOTU405 | | | 3.02222564519143 |
| 21 | Verghello vs Asco | New.CleanUp.ReferenceOTU2257 | | Leotiomycetes | -1.94181850226078 |
| 22 | Verghello vs Asco | New.ReferenceOTU91 | fungal sp TRN287 | unidentified | -3.35310980259584 |
| 23 | Verghello vs Asco | New.CleanUp.ReferenceOTU1582 | | Dothideomycetes | -3.8894236232724 |
| 24 | Verghello vs Asco | New.CleanUp.ReferenceOTU3034 | | Dothideomycetes | -3.7691205317524 |
| 25 | Verghello vs Asco | New.CleanUp.ReferenceOTU3877 | | | -2.91496849506285 |
| 26 | Verghello vs Asco | New.CleanUp.ReferenceOTU4128 | | Dothideomycetes | -2.68118762168957 |
| 27 | Verghello vs Asco | New.CleanUp.ReferenceOTU2204 | | | 3.48918515064991 |
| 28 | Verghello vs Asco | New.CleanUp.ReferenceOTU1124 | | Dothideomycetes | -3.0201750258822 |
| 29 | Verghello vs Asco | New.ReferenceOTU77 | | Dothideomycetes | -3.90157972475388 |
| 30 | Verghello vs Asco | New.CleanUp.ReferenceOTU4251 | | unidentified | -3.74237103535745 |
| 31 | Verghello vs Asco | New.CleanUp.ReferenceOTU4270 | fungal sp TRN287 | | 3.65642945371858 |
| 32 | Verghello vs Asco | New.CleanUp.ReferenceOTU1866 | | | -4.29516742120644 |
| 33 | Verghello vs Asco | New.CleanUp.ReferenceOTU2594 | | Dothideomycetes | -2.92019526244776 |
| 34 | Verghello vs Asco | New.CleanUp.ReferenceOTU4335 | Phaeothecoidea sp | Dothideomycetes | 2.23132133531876 |
| 35 | Verghello vs Asco | New.CleanUp.ReferenceOTU1211 | Phaeotheca sp | Dothideomycetes | -4.15907125495757 |
| 36 | Verghello vs Asco | New.CleanUp.ReferenceOTU3089 | | Dothideomycetes | -2.4625965536206 |
| 37 | Verghello vs Asco | New.CleanUp.ReferenceOTU2729 | | Dothideomycetes | -8.17280664276634 |
| 38 | Verghello vs Asco | New.CleanUp.ReferenceOTU1325 | | | 3.40278786005025 |
| 39 | Verghello vs Asco | New.CleanUp.ReferenceOTU1779 | Capnobotryella sp MA 4642 | Dothideomycetes | -2.71716108510325 |
| 40 | Verghello vs Asco | New.CleanUp.ReferenceOTU617 | | | -2.5522071382824 |
| 41 | Verghello vs Asco | New.CleanUp.ReferenceOTU3655 | | | -2.92594421012468 |
| 42 | Verghello vs Asco | SH027337.07FU_EU707898_reps_singleton | | Dothideomycetes | 2.51258937103459 |
| 43 | Verghello vs Asco | New.CleanUp.ReferenceOTU2256 | | Dothideomycetes | -3.40893529950565 |
| 44 | Verghello vs Asco | New.ReferenceOTU24 | | Dothideomycetes | -2.95229916712289 |
| 45 | Verghello vs Asco | New.ReferenceOTU21 | | Dothideomycetes | -3.17471556733971 |
| 46 | Verghello vs Asco | New.CleanUp.ReferenceOTU5825 | Capnobotryella sp MA 4642 | Dothideomycetes | -2.59883342572787 |
| 47 | Verghello vs Asco | New.CleanUp.ReferenceOTU3208 | | | 4.06811644065732 |
| 48 | Verghello vs Asco | New.CleanUp.ReferenceOTU5870 | | | -3.60797778187428 |
| 49 | Verghello vs Asco | SH215125.07FU_KJ406972_reps | Rhytismatales sp | Leotiomycetes | -4.83057303993148 |
| 50 | Verghello vs Asco | SH020264.07FU_KF800472_reps_singleton | Aureobasidium pullulans | Dothideomycetes | -24.7698542518091 |
| 51 | Verghello vs Asco | New.CleanUp.ReferenceOTU3804 | Lachnellula calyciformis | Leotiomycetes | -4.47134827172345 |
| 52 | Verghello vs Asco | SH018491.07FU_JN053173_reps_singleton | fungal sp agrD488 | unidentified | -9.15113175973644 |
| 53 | Verghello vs Asco | New.CleanUp.ReferenceOTU5038 | | Dothideomycetes | -3.77537799709519 |
| 54 | Verghello vs Asco | New.CleanUp.ReferenceOTU186 | Capnobotryella sp MA 4642 | Dothideomycetes | -3.12530974729654 |
| 55 | Verghello vs Asco | New.CleanUp.ReferenceOTU733 | fungal sp agrD488 | unidentified | -5.17756965843196 |
| 56 | Verghello vs Asco | New.CleanUp.ReferenceOTU496 | | | -5.24203193721767 |
| 57 | Verghello vs Asco | New.CleanUp.ReferenceOTU485 | fungal sp TRN287 | unidentified | -3.50658784085513 |
| 58 | Verghello vs Asco | New.CleanUp.ReferenceOTU4078 | Capnobotryella sp MA 4642 | Dothideomycetes | -3.28942465133106 |
| 59 | Verghello vs Asco | New.CleanUp.ReferenceOTU551 | | Dothideomycetes | -2.90877746553527 |
| 60 | Verghello vs Asco | New.ReferenceOTU2 | | | -28.775503383438 |
| 61 | Verghello vs Asco | New.CleanUp.ReferenceOTU1645 | Lophodermium seditiosum | Leotiomycetes | 2.54496105588887 |
| 62 | Verghello vs Asco | New.CleanUp.ReferenceOTU2382 | Phaeotheca sp | Dothideomycetes | 2.85177899653228 |
| 63 | Verghello vs Asco | New.CleanUp.ReferenceOTU1566 | | | 3.45767555951727 |
| 64 | Verghello vs Asco | New.CleanUp.ReferenceOTU49 | | | 3.46893361479709 |
| 65 | Verghello vs Asco | New.CleanUp.ReferenceOTU1312 | | Dothideomycetes | -3.75954143874704 |
| 66 | Verghello vs Asco | New.CleanUp.ReferenceOTU2784 | | Dothideomycetes | -2.91450576680127 |
| 67 | Verghello vs Asco | New.CleanUp.ReferenceOTU5448 | | Dothideomycetes | -2.82804917380157 |
| 68 | Verghello vs Asco | New.CleanUp.ReferenceOTU1778 | | Dothideomycetes | -6.90844720707245 |
| 69 | Verghello vs Asco | New.CleanUp.ReferenceOTU5003 | Sarea sp | Lecanoromycetes | 6.4277295951405 |
| 70 | Verghello vs Asco | SH006640.07FU_GU910926_reps_singleton | Sporormiaceae sp | Dothideomycetes | -5.97998200105863 |
| 71 | Verghello vs Asco | New.CleanUp.ReferenceOTU4791 | fungal sp agrD488 | unidentified | -24.7666624761591 |
| 72 | Verghello vs Asco | New.CleanUp.ReferenceOTU2190 | | Dothideomycetes | -3.22755115301385 |
| 73 | Verghello vs Asco | New.CleanUp.ReferenceOTU2400 | | Dothideomycetes | -4.21110850370276 |
| 74 | Verghello vs Asco | SH184176.07FU_GU909656_reps | Sporormiaceae sp | Dothideomycetes | -6.00364485230586 |
| 75 | Verghello vs Asco | SH205426.07FU_AY843155_reps | | | -24.9579666000812 |
| 76 | Verghello vs Asco | New.CleanUp.ReferenceOTU4861 | | Leotiomycetes | 22.2388704716995 |
| 77 | Verghello vs Asco | New.CleanUp.ReferenceOTU3295 | | | 6.65123975257568 |
| 78 | Verghello vs Asco | SH202297.07FU_KF675366_reps | | Dothideomycetes | 23.3317918150654 |
| 79 | Verghello vs Asco | New.CleanUp.ReferenceOTU5274 | | Eurotiomycetes | 21.2386330777242 |
| 80 | Verghello vs Asco | New.CleanUp.ReferenceOTU5764 | | Lecanoromycetes | 25.8033702549455 |
| 81 | Verghello vs Bavella | SH197740.07FU_EF419976_reps | Phaeomoniella sp | Eurotiomycetes | 3.01191870374097 |
| 82 | Verghello vs Bavella | SH214165.07FU_AY843074_reps | | Dothideomycetes | 1.88265028828519 |
| 83 | Verghello vs Bavella | New.CleanUp.ReferenceOTU1083 | Lophodermium conigenum | Leotiomycetes | 3.95004407392942 |
| 84 | Verghello vs Bavella | New.ReferenceOTU4 | fungal sp TRN287 | unidentified | 2.86251081858689 |
| 85 | Verghello vs Bavella | SH017706.07FU_AJ971406_reps_singleton | Capnobotryella sp MA 4642 | Dothideomycetes | 2.54623521500715 |
| 86 | Verghello vs Bavella | New.CleanUp.ReferenceOTU4468 | | Dothideomycetes | 2.7737238869659 |
| 87 | Verghello vs Bavella | SH006502.07FU_FN868467_reps_singleton | Leotiomycetes sp BLD3 | Leotiomycetes | 3.89839831920109 |
| 88 | Verghello vs Bavella | New.ReferenceOTU19 | | Dothideomycetes | 2.55312805953063 |
| 89 | Verghello vs Bavella | New.ReferenceOTU16 | | | 3.78324511694868 |
| 90 | Verghello vs Bavella | New.CleanUp.ReferenceOTU2078 | | Leotiomycetes | 3.97186051632267 |
| 91 | Verghello vs Bavella | New.CleanUp.ReferenceOTU1128 | | Leotiomycetes | 3.88411531596102 |
| 92 | Verghello vs Bavella | New.CleanUp.ReferenceOTU5051 | | | 5.47137216131685 |
| 93 | Verghello vs Bavella | SH206392.07FU_AM921728_reps | Dothioraceae sp | Dothideomycetes | 2.49584211265739 |
| 94 | Verghello vs Bavella | New.CleanUp.ReferenceOTU69 | Phaeotheca sp | Dothideomycetes | 5.59688808234295 |
| 95 | Verghello vs Bavella | New.CleanUp.ReferenceOTU1913 | | | 3.12787189465673 |
| 96 | Verghello vs Bavella | New.CleanUp.ReferenceOTU6463 | | Dothideomycetes | 2.20132711628261 |
| 97 | Verghello vs Bavella | New.ReferenceOTU51 | Lophodermium conigenum | Leotiomycetes | 4.03896922221753 |
| 98 | Verghello vs Bavella | New.CleanUp.ReferenceOTU5284 | | Dothideomycetes | 2.19621641754827 |
| 99 | Verghello vs Bavella | New.CleanUp.ReferenceOTU4893 | fungal sp TRN287 | unidentified | 3.1463010425868 |
| 100 | Verghello vs Bavella | SH211751.07FU_KJ406871_reps | | | 4.07499554637067 |
| 101 | Verghello vs Bavella | New.CleanUp.ReferenceOTU765 | Ochrocladosporium sp | Dothideomycetes | 4.49907182970484 |
| 102 | Verghello vs Bavella | New.CleanUp.ReferenceOTU2528 | | Leotiomycetes | 3.88850861419167 |
| 103 | Verghello vs Bavella | New.CleanUp.ReferenceOTU4409 | | Leotiomycetes | 4.5910924430294 |
| 104 | Verghello vs Bavella | New.CleanUp.ReferenceOTU1703 | | | 6.53824338573641 |
| 105 | Verghello vs Bavella | New.CleanUp.ReferenceOTU3609 | fungal sp TRN287 | unidentified | 22.6742092560571 |
| 106 | Verghello vs Bavella | New.CleanUp.ReferenceOTU2145 | | Leotiomycetes | 2.33068716549373 |
| 107 | Verghello vs Bavella | New.CleanUp.ReferenceOTU406 | fungal sp TRN213 | unidentified | 2.54509520458045 |

| | Comparison | Order | Class | log2FoldChange (negative = more on second levels) |
|---|----------------------|-------------------|-----------------|---|
| 1 | Verghello vs Asco | Xylariales | Sordariomycetes | 5.04736158531515 |
| 2 | Verghello vs Bavella | Incertae sedis | Leotiomycetes | -1.3470109106199 |
| 3 | Verghello vs Bavella | unidentified | unidentified | 1.59296035953185 |
| 4 | Asco vs Bavella | Botryosphaeriales | Dothideomycetes | 7.46092665033677 |
| 5 | Asco vs Bavella | Eurotiales | Eurotiomycetes | 1.80585521286681 |
| 6 | Asco vs Bavella | Incertae sedis | Leotiomycetes | -1.68998134956396 |
| 7 | Asco vs Bavella | unidentified | unidentified | 1.46261385771009 |
| 8 | Asco vs Bavella | Xylariales | Sordariomycetes | -4.7735371857996 |

Table 14: Order showing differential abundances in the different sites.

| | Nb.of.OTUs | Nb.of.samples | Nb.of.sequences |
|--|------------|---------------|-----------------|
| No filter | 4373 | 80 | 8398038.00 |
| Nb of sequences by sample ≥ 20000 | 4359 | 72 | 8375892.00 |
| Nb of sample by OTUs ≥ 1 | 4359 | 72 | 8375892.00 |
| Nb of sequences by OTUs ≥ 5 | 3382 | 72 | 8373567.00 |

Table 15: Number of OTUs, samples and sequences after filtering

List of Figures

| | | |
|------|--|----|
| 2.1 | Number of sequences by sample | 9 |
| 2.2 | Number of OTU present in a given number of samples | 10 |
| 2.3 | Number of sequences by OTU (log10 transformed) | 11 |
| 3.1 | Number of OTUs by sample in fonction of the number of sequences by sample (log10 axe) | 12 |
| 3.2 | Distribution of reference sequences length | 13 |
| 3.3 | Number of sequences by OTUs (log10 axe) in fonction of the number of samples where OTUs were found | 13 |
| 3.4 | Distribution of the number of sequences in the Ascomycota taxonomy | 14 |
| 3.5 | Number of sequences of the 30 most abundant OTUs (number of sequences) | 16 |
| 4.1 | Distribution of OTUs into functional Guild | 18 |
| 4.2 | Distribution of sequences (log10 transformed) into functional Guild | 19 |
| 5.1 | Rarefaction curves for each site | 20 |
| 5.2 | Rarefaction curves for each sample using sequences number on x-axes | 21 |
| 5.3 | Diversity of each sites | 22 |
| 5.4 | Rarefaction curves for each host age | 23 |
| 5.5 | Diversity in fonction of tree age | 24 |
| 5.6 | Rarefaction curves for each elevation | 25 |
| 5.7 | Diversity in fonction of elevation | 26 |
| 5.8 | Results of the Tuckey HSD testing for differences in mean Hill numbers among pairs of modalities | 27 |
| 6.1 | Venn diagramm of the distribution of OTUs among Sites | 28 |
| 6.2 | Venn diagramm of the distribution of OTUs among host age | 29 |
| 6.3 | Venn diagramm of the distribution of OTUs among elevation of samples | 29 |
| 6.4 | Venn diagramm of the distribution of OTUs among Sites | 30 |
| 6.5 | Venn diagramm of the distribution of OTUs among host age | 30 |
| 6.6 | Venn diagramm of the distribution of OTUs among elevation of samples whitin the tree | 31 |
| 6.7 | Stress plot of the NMDS | 32 |
| 6.8 | NMDS of OTU table | 32 |
| 6.9 | Comparison of different distances (bray (up) and gower (bottom)) and ordination methods (NMDS (left), PCoA (center) and DCA (right)) | 34 |
| 6.10 | Taxonomic distribution of sequences in the different site * age combinaison | 38 |
| 6.11 | Taxonomic distribution of OTUs in the different site * age combinaison | 39 |
| 6.12 | OTUs significantly different in terms of abundances between Verghello (positive values) and Asco (negative values) | 40 |
| 6.13 | OTUs significantly different in terms of abundances between Verghello (positive values) and Bavella (negative values) | 41 |
| 6.14 | OTUs significantly different in terms of abundances between Asco (positive values) and Bavella (negative values) | 42 |

List of Tables

| | | |
|---|---|----|
| 1 | Number of OTUs, samples and sequences after filtering | 11 |
| 2 | Taxonomie of the 30 more abundant OTUs (number of sequences) | 15 |
| 3 | Taxonomie of the 30 more frequent OTUs (number of samples) | 17 |
| 4 | Summary of the linear model of species richness (Hill number with $q = 0$) | 23 |

| | | |
|----|--|----|
| 5 | Summary of the linear model of the exponential of Shannon's entropy index (Hill number with $q = 1$) | 27 |
| 6 | Summary of the linear model of inverse of Simpson's concentration index (Hill number with $q = 2$) | 27 |
| 7 | Result of the permanova on abundances (number of sequence). | 35 |
| 8 | Result of the permanova on abundances (number of sequence) using only OTUs present in more than 30 samples | 36 |
| 9 | Result of the permanova on OTUs (each OTU is representing by one sequence)). | 36 |
| 10 | Result of the permanova on abundances (number of sequence). | 37 |
| 11 | Result of the permanova on abundances (number of sequence) using only OTUs present in more than 30 samples | 37 |
| 12 | Result of the permanova on OTUs (each OTU is representing by one sequence)). | 37 |
| 13 | OTUs showing differential abundances in the different sites. | 48 |
| 14 | Order showing differential abundances in the different sites. | 49 |
| 15 | Number of OTUs, samples and sequences after filtering | 49 |