# Appendix S5: results after UPARSE clustering allowing unique sequences (Usearch function sortbysize with argument -minsize 1). 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

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 UPARSE clustering allowing unique sequences (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('ramnathv/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 15:49:27. 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 <- "Uparse_min1"

#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)
```

```
# Sort taxonomy
sort_taxRDP_brut <- unlist(strsplit(unlist(strsplit(rownames(dataBiom), split = ";"))
                                    [seq(1, length(rownames(dataBiom))*2, by = 2)],
                                    split =  "_"))[seq(2, length(rownames(dataBiom))*2,
                                                       by = 2)]
taxRDP_brut <- taxRDP_brut[1:dim(taxRDP_brut)[1] %in% sort_taxRDP_brut,]

# Format taxonomy for phyloseq
taxRDP <- taxRDP_brut[match(taxa_names(dataBiom),
                            paste(taxRDP_brut[, 1], taxRDP_brut[, 2], "", sep = ";")),
                      c(5, 7, 9, 11, 13, 15, 17)]
taxRDP <- tax_table(as.matrix(taxRDP))
taxa_names(taxRDP) <- taxa_names(dataBiom)
colnames(taxRDP) <- c("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(paste(funguild$OTU_ID,";", sep = ""), gsub(";size=", "_",
                                                                 rownames(taxRDP2)))

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


taxRDP2 <- tax_table(as.matrix(taxRDP2))
taxa_names(taxRDP2) <- taxa_names(dataBiom)
colnames(taxRDP2) <- c("Domain", "Phylum", "Class", "Order", "Family", "Genus", "Species",
                       "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...
```

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

### 2.2.7 Caracteristics of the phyloseq data

```
data_all

## phyloseq-class experiment-level object
## otu_table()   OTU Table:         [ 1667 taxa and 80 samples ]
## sample_data() Sample Data:       [ 80 samples by 6 sample variables ]
## tax_table()   Taxonomy Table:    [ 1667 taxa by 12 taxonomic ranks ]
## refseq()      DNAStringSet:      [ 1667 reference sequences ]
```

The data are made of $8.335341 \times 10^6$ sequences representing 1667 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 $2 \times 10^4$ sequences, we keep 72 on the 80 samples (90%).

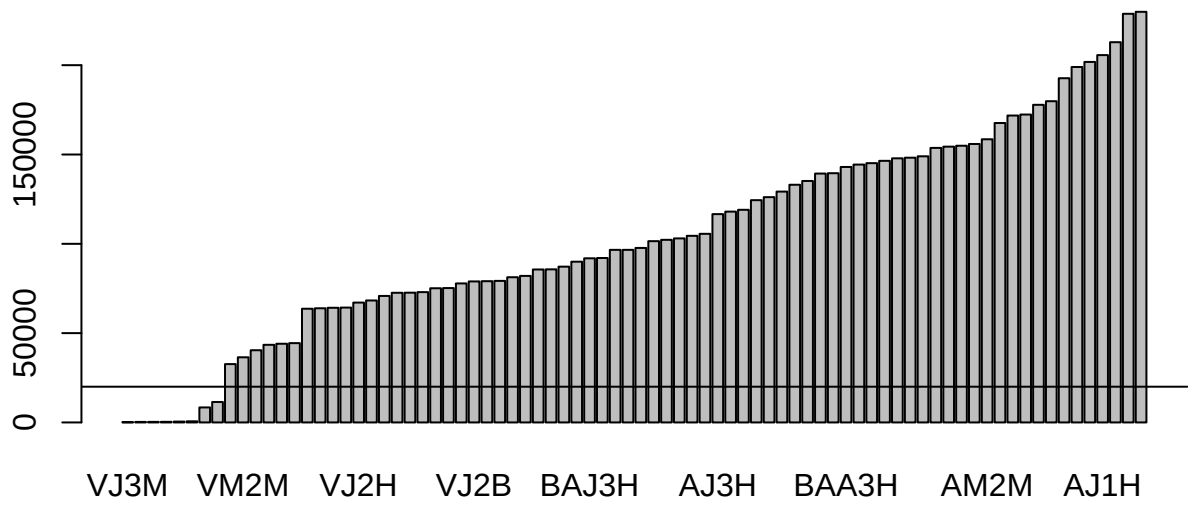



**Figure 2.1:** Number of sequences by sample. Horizontal line indicates the filtering parameter.

```
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).

```
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   18.25   35.50   36.00   53.75   72.00
```

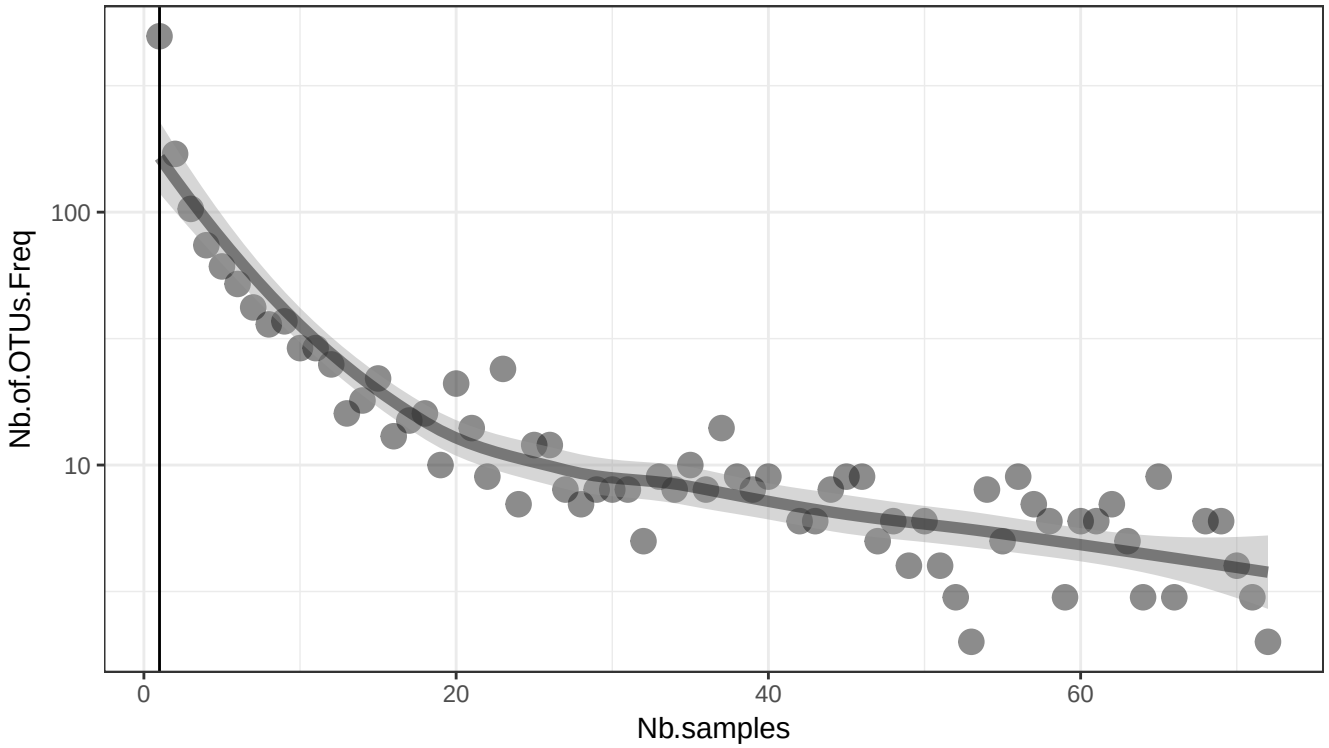



**Figure 2.2:** Number of OTU present in a given number of samples. Vertical bar illustrates the filtering parameter.

```
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 1650 on the 1650 OTUs (100%).

## 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.
##       1.0       7.0      46.0    5038.3     337.8 1943171.0
```

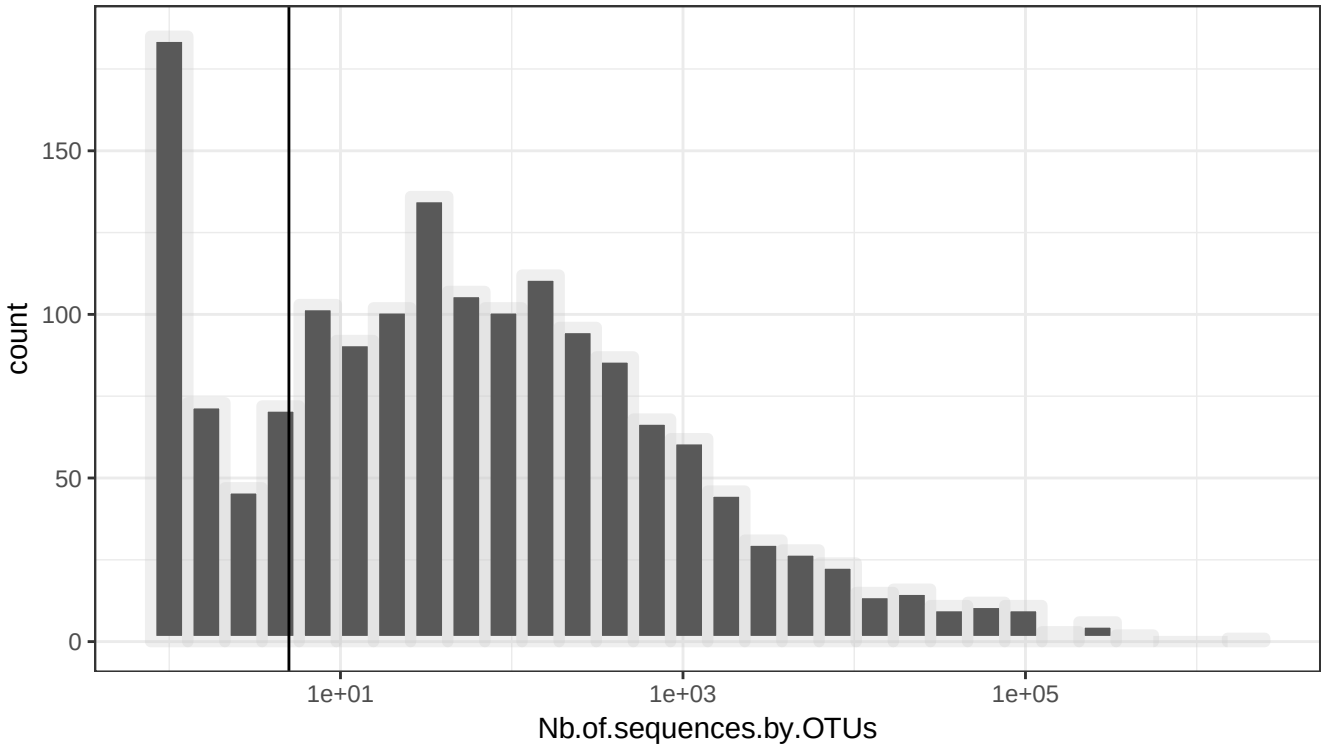



**Figure 2.3:** Number of sequences by OTU (log10 transformed). Horizontal bar illustrates the filtering parameter.

```
N_seq_otu_min

## [1] 5
```

If we discard OTUs with less than 1 sequences, we keep 1302 on the 1667 OTUs (78.1%).

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

## 2.6 Summary of filtration workflow

The filtered data are made of $8.312594 \times 10^6$ sequences representing 1302 OTUs allocate to 72 samples.

| | Nb.of.OTUs | Nb.of.samples | Nb.of.sequences |
|---|---|---|---|
| No filter | 1667 | 80 | 8335341.00 |
| Nb of sequences by sample >= 20000 | 1650 | 72 | 8313238.00 |
| Nb of sample by OTUs >= 1 | 1650 | 72 | 8313238.00 |
| Nb of sequences by OTUs >= 5 | 1302 | 72 | 8312594.00 |

**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

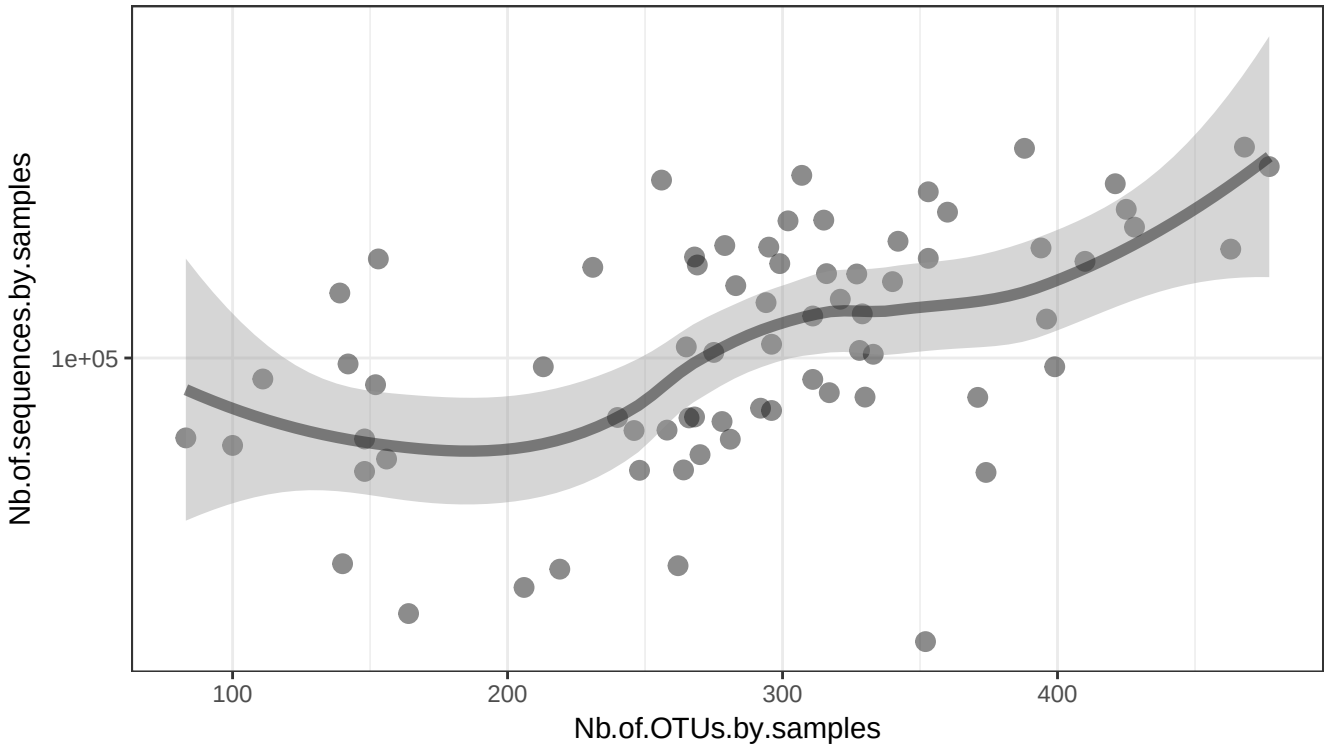



**Figure 3.1:** Number of OTUs by sample in fonction of the number of sequences by sample (log10 axe). The tendency is represented by the line obtained from loess (Local Polynomial Regression Fitting).

```
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') +
  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])
```

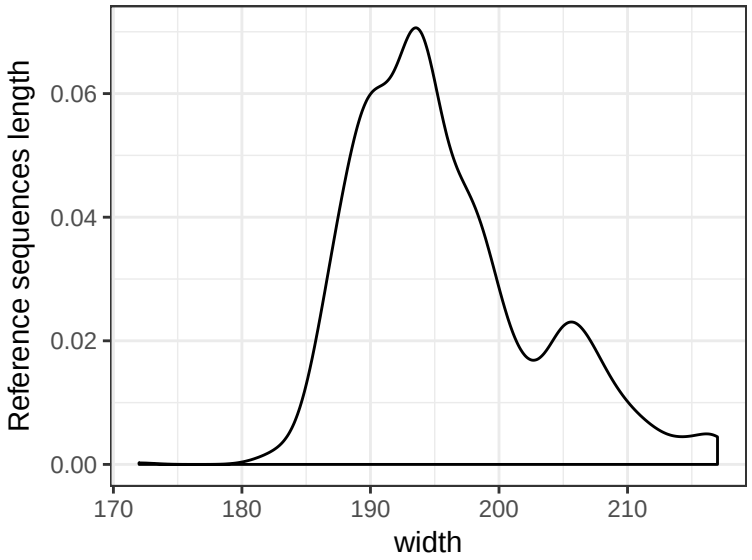



**Figure 3.2:** Distribution of reference sequences length.

```
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"))
```

## 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)
```

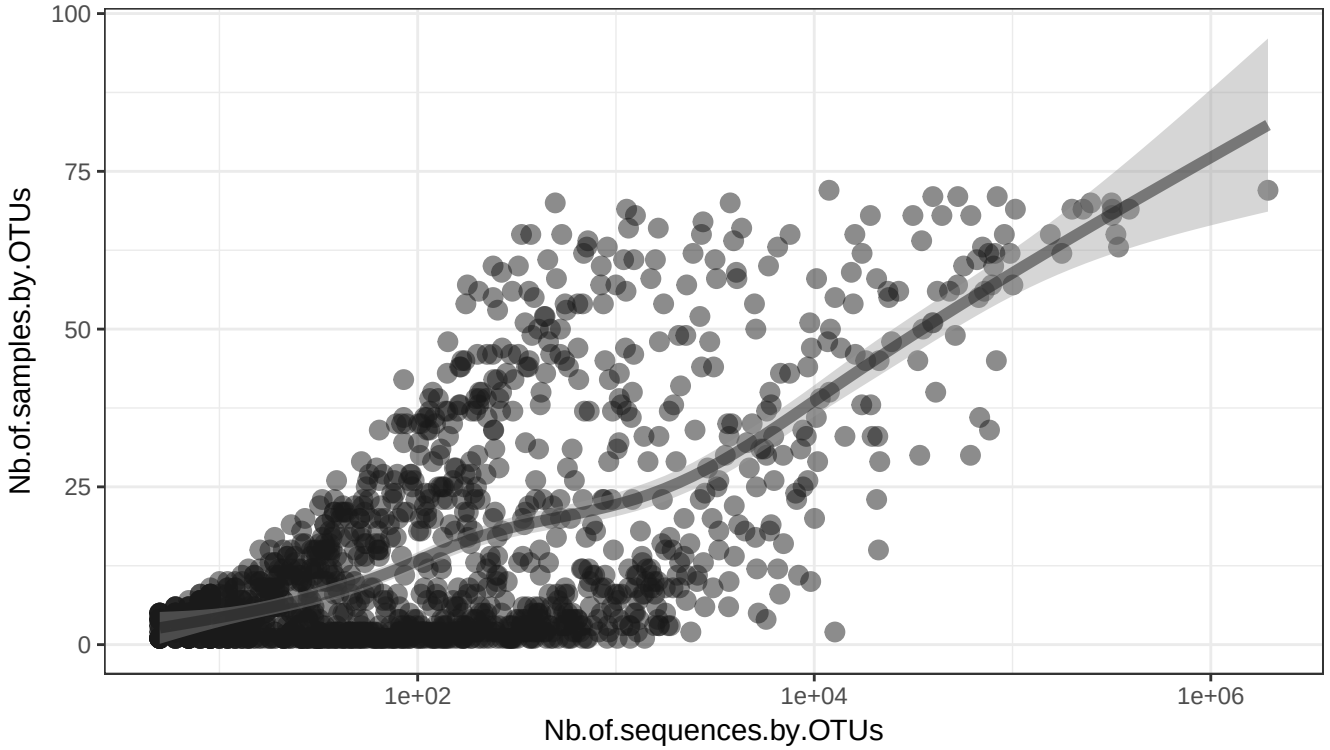



**Figure 3.3:** Number of sequences by OTUs (log10 axe) in fonction of the number of samples where OTUs were found. The tendency is represented by the line obtain from gam (Generalized additive models with integrated smoothness estimation).

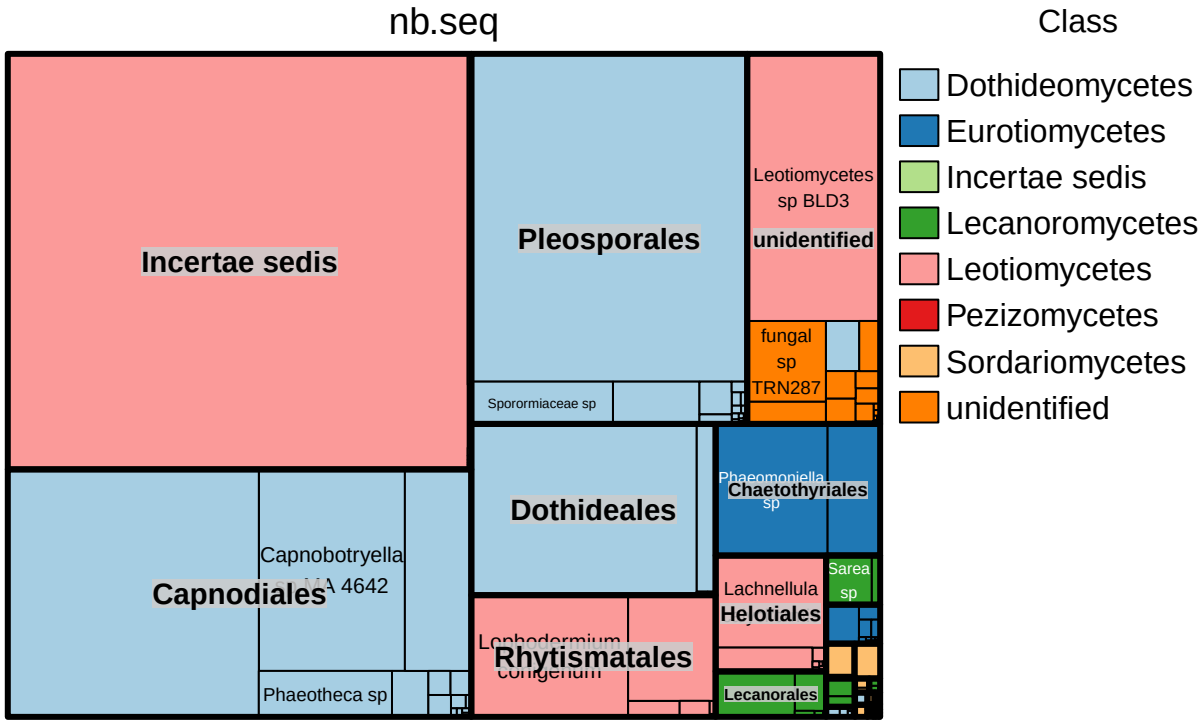



**Figure 3.4:** Distribution of the number of sequences in the Ascomycota taxonomy. Colors represent Class, bold lines delimit Order and thick line delimit species.

```
print(xtable(df_the30mostfrequents[, c(1:8, 12)], auto = T,
             caption = "Taxonomie of the 30 more
             abundant OTUs (number of sequences)"),
      size = "\\tiny", include.rownames = FALSE)
```

| Phylum | Class | Order | Family | Genus | Species | Trophic_Mode | Guild | Nb.sequences |
|---|---|---|---|---|---|---|---|---|
| Ascomycota | Leotiomycetes | Incertae sedis | Incertae sedis | Cyclaneusma | Cyclaneusma minus | - | - | 1943171 |
| Ascomycota | Dothideomycetes | Pleosporales | | | | - | - | 387885 |
| Ascomycota | Dothideomycetes | Capnodiales | Incertae sedis | Capnobotryella | Capnobotryella sp MA 4642 | Saprotroph | Undefined Saprotroph | 342517 |
| Ascomycota | Dothideomycetes | Pleosporales | Incertae sedis | | | - | - | 332263 |
| Ascomycota | Leotiomycetes | unidentified | unidentified | unidentified | Leotiomycetes sp BLD3 | - | - | 318495 |
| Ascomycota | Dothideomycetes | Dothideales | Dothioraceae | unidentified | Dothioraceae sp | - | - | 317246 |
| Ascomycota | Dothideomycetes | Capnodiales | | | | - | - | 315723 |
| | | | | | | - | - | 247896 |
| Ascomycota | Leotiomycetes | Rhytismatales | Rhytismataceae | Lophodermium | Lophodermium conigenum | Pathotroph | Plant Pathogen | 226872 |
| Ascomycota | Dothideomycetes | Pleosporales | | | | - | - | 199465 |
| | | | | | | - | - | 177383 |
| Ascomycota | Dothideomycetes | Capnodiales | Mycosphaerellaceae | Phaeothecoidea | Phaeothecoidea sp | Saprotroph | Undefined Saprotroph | 155089 |
| Ascomycota | Leotiomycetes | unidentified | unidentified | unidentified | Leotiomycetes sp BLD3 | - | - | 103401 |
| | | | | | | - | - | 100459 |
| Ascomycota | Dothideomycetes | Dothideales | Dothioraceae | unidentified | Dothioraceae sp | - | - | 97232 |
| Ascomycota | Dothideomycetes | Pleosporales | | | | - | - | 90937 |
| Ascomycota | Leotiomycetes | Incertae sedis | Incertae sedis | Cyclaneusma | Cyclaneusma minus | - | - | 83720 |
| | | | | | | - | - | 82876 |
| Ascomycota | Eurotiomycetes | Chaetothyriales | Herpotrichiellaceae | Phaeomoniella | Phaeomoniella sp | Saprotroph | Undefined Saprotroph | 81451 |
| Ascomycota | Leotiomycetes | Rhytismatales | Rhytismataceae | Lophodermium | | Pathotroph | Plant Pathogen | 80357 |
| | | | | | | - | - | 78492 |
| Ascomycota | | | | | | - | - | 76573 |
| Ascomycota | Leotiomycetes | Incertae sedis | Incertae sedis | Cyclaneusma | Cyclaneusma minus | - | - | 75793 |
| unidentified | unidentified | unidentified | unidentified | unidentified | fungal sp TRN287 | - | - | 72163 |
| Ascomycota | Dothideomycetes | Capnodiales | | | | - | - | 70557 |
| Ascomycota | | | | | | - | - | 68326 |
| Ascomycota | Leotiomycetes | Helotiales | Hyaloscyphaceae | Lachnellula | Lachnellula calyciformis | Saprotroph | Undefined Saprotroph | 67576 |
| Ascomycota | Eurotiomycetes | Chaetothyriales | Herpotrichiellaceae | Phaeomoniella | Phaeomoniella sp | Saprotroph | Undefined Saprotroph | 66072 |
| Ascomycota | Dothideomycetes | Capnodiales | Davidiellaceae | Cladosporium | | - | - | 61649 |

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

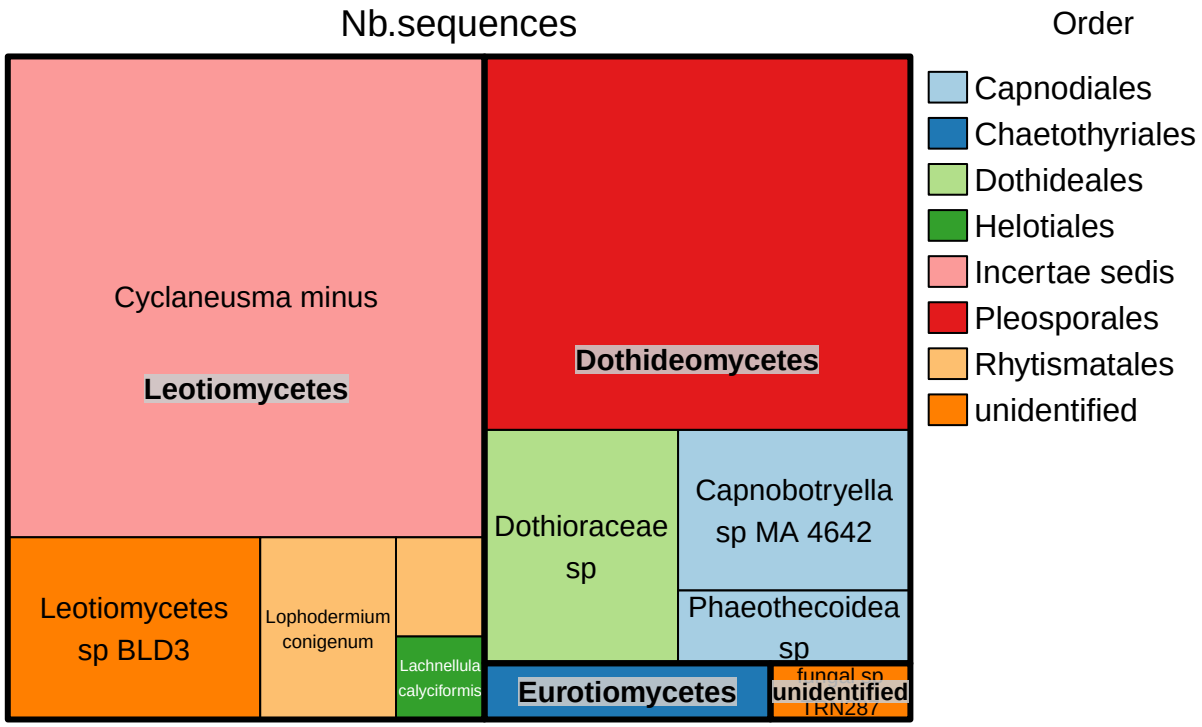



**Figure 3.5:** Number of sequences of the 30 most abundant OTUs (number of sequences). Colors indicate Order, bold lines delimit Class and thick lines delimit species.

```
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)
```

| Phylum | Class | Order | Family | Genus | Species | Trophic_Mode | Guild | Nb.samples |
|---|---|---|---|---|---|---|---|---|
| Ascomycota | Leotiomycetes | Incertae sedis | Incertae sedis | Cyclaneusma | Cyclaneusma minus | - | - | 72 |
| Ascomycota | Leotiomycetes | Incertae sedis | Incertae sedis | Cyclaneusma | Cyclaneusma minus | - | - | 72 |
| Ascomycota | Leotiomycetes | Incertae sedis | Incertae sedis | Cyclaneusma | Cyclaneusma minus | - | - | 71 |
| Ascomycota | Leotiomycetes | Incertae sedis | Incertae sedis | Cyclaneusma | Cyclaneusma minus | - | - | 71 |
| Ascomycota | Leotiomycetes | Incertae sedis | Incertae sedis | Cyclaneusma | Cyclaneusma minus | - | - | 71 |
| | | | | | | - | - | 70 |
| Ascomycota | Dothideomycetes | Capnodiales | | | | - | - | 70 |
| Ascomycota | Leotiomycetes | Incertae sedis | Incertae sedis | Cyclaneusma | Cyclaneusma minus | - | - | 70 |
| Ascomycota | Leotiomycetes | Incertae sedis | Incertae sedis | Cyclaneusma | Cyclaneusma minus | - | - | 70 |
| Ascomycota | Leotiomycetes | Rhytismatales | Rhytismataceae | Lophodermium | Lophodermium conigenum | Pathotroph | Plant Pathogen | 69 |
| Ascomycota | Dothideomycetes | Pleosporales | | | | - | - | 69 |
| Ascomycota | Leotiomycetes | unidentified | unidentified | unidentified | Leotiomycetes sp BLD3 | - | - | 69 |
| Ascomycota | Leotiomycetes | unidentified | unidentified | unidentified | Leotiomycetes sp BLD3 | - | - | 69 |
| Ascomycota | Dothideomycetes | Pleosporales | | | | - | - | 69 |
| Ascomycota | Leotiomycetes | Incertae sedis | Incertae sedis | Cyclaneusma | Cyclaneusma minus | - | - | 69 |
| Ascomycota | Dothideomycetes | Dothideales | Dothioraceae | unidentified | Dothioraceae sp | - | - | 68 |
| Ascomycota | Dothideomycetes | Pleosporales | Incertae sedis | Ochrocladosporium | Ochrocladosporium sp | Saprotroph | Undefined Saprotroph | 68 |
| Ascomycota | Dothideomycetes | Capnodiales | Davidiellaceae | Cladosporium | | - | - | 68 |
| Ascomycota | Leotiomycetes | Incertae sedis | Incertae sedis | Cyclaneusma | Cyclaneusma minus | - | - | 68 |
| Ascomycota | Leotiomycetes | Incertae sedis | Incertae sedis | Cyclaneusma | Cyclaneusma minus | - | - | 68 |
| Ascomycota | Leotiomycetes | Incertae sedis | Incertae sedis | Cyclaneusma | Cyclaneusma minus | - | - | 68 |
| Ascomycota | Leotiomycetes | unidentified | unidentified | unidentified | Leotiomycetes sp BLD3 | - | - | 67 |
| Ascomycota | Dothideomycetes | Capnodiales | | | | - | - | 66 |
| | | | | | | - | - | 66 |
| Ascomycota | Leotiomycetes | Incertae sedis | Incertae sedis | Cyclaneusma | Cyclaneusma minus | - | - | 66 |

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

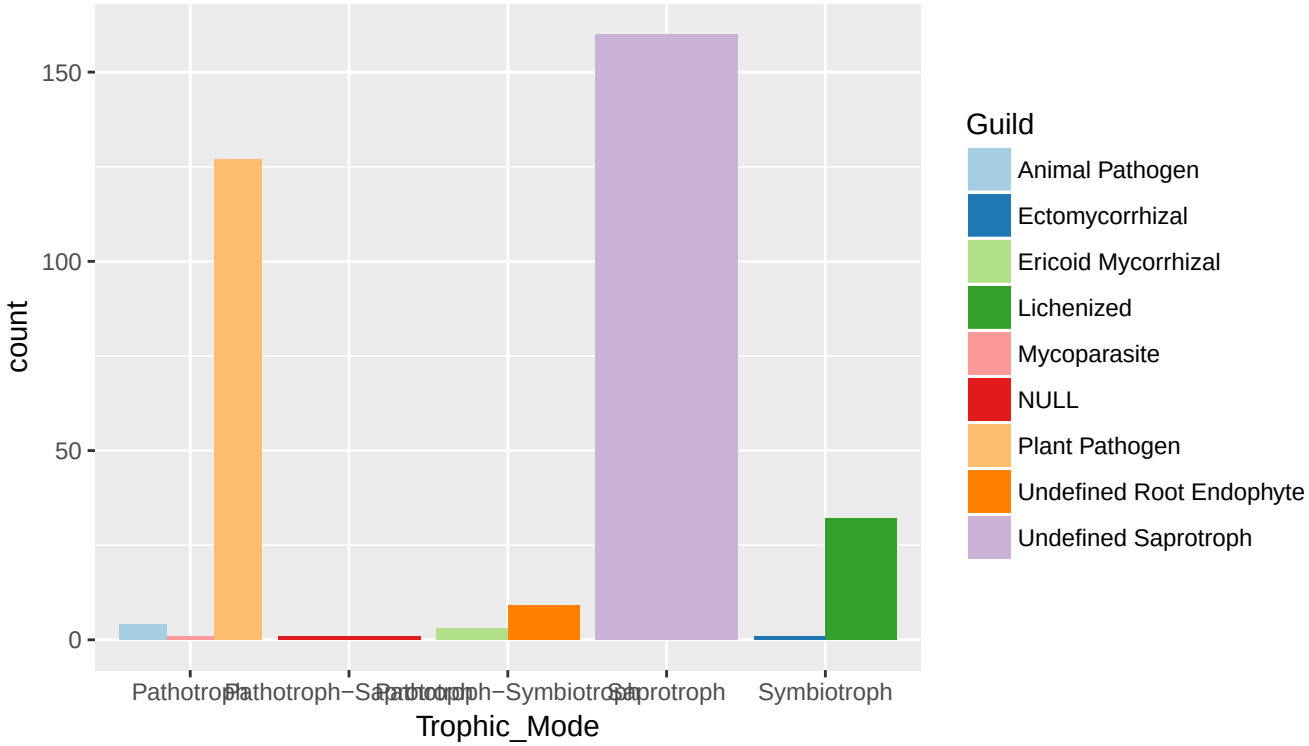



**Figure 4.1:** Distribution of OTUs into functional Guild.

# 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] 82.18797

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()
```

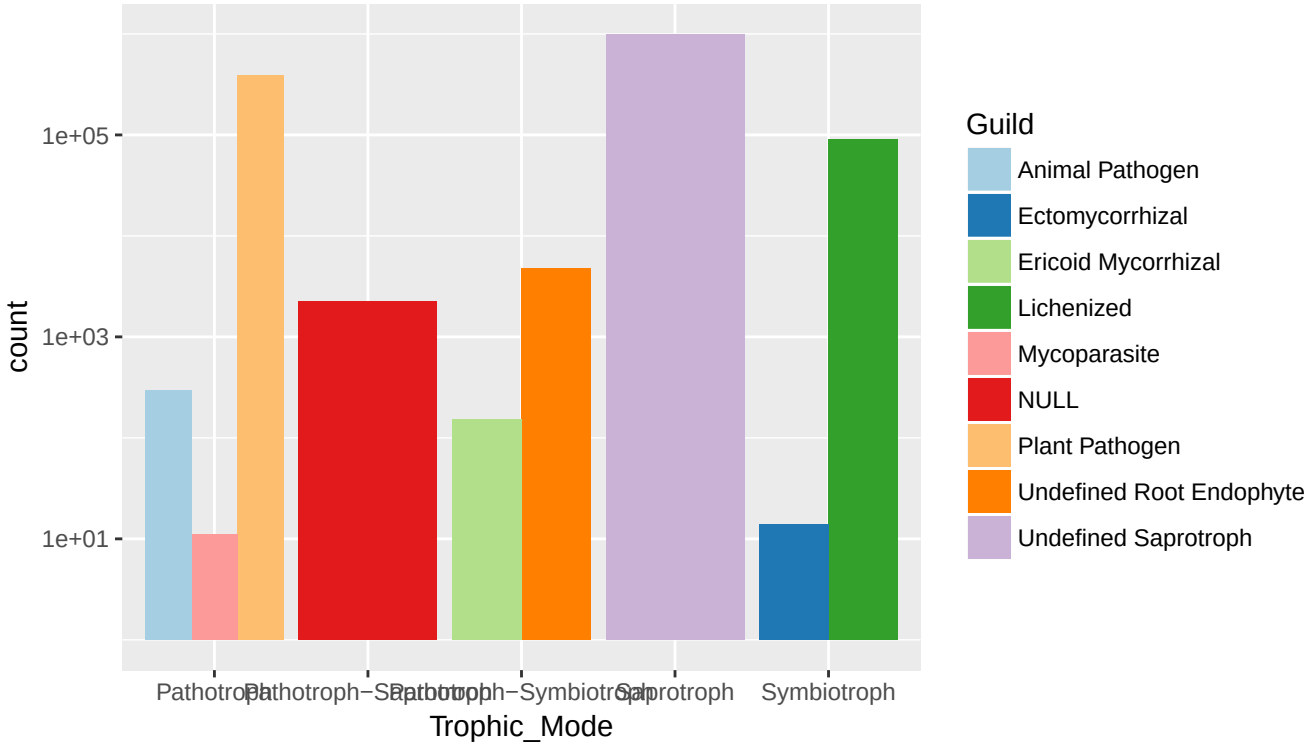



**Figure 4.2:** Distribution of sequences (log10 transformed) into functional Guild.

# 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)
```

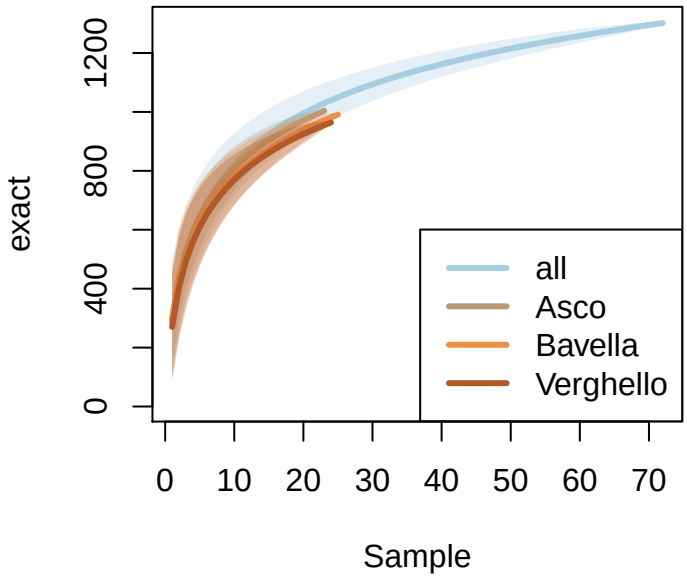



**Figure 5.1:** Rarefaction curves for each site. Note that if singletons were removed, these curves are biaised.

## 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)
```

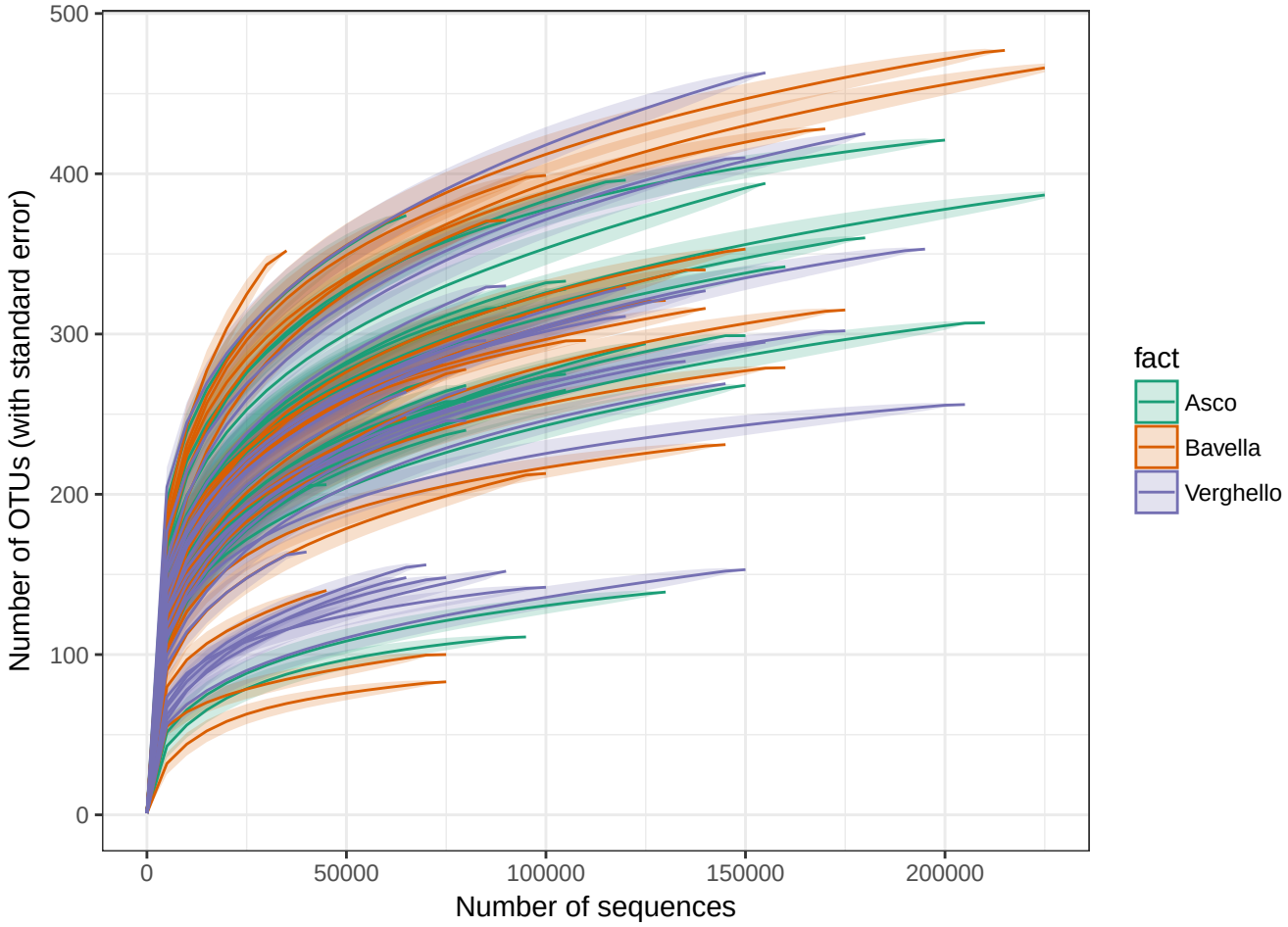



**Figure 5.2:** Rarefaction curves for each sample using sequences number on x-axes. Note that if singletons were removed, these curves are biaised.

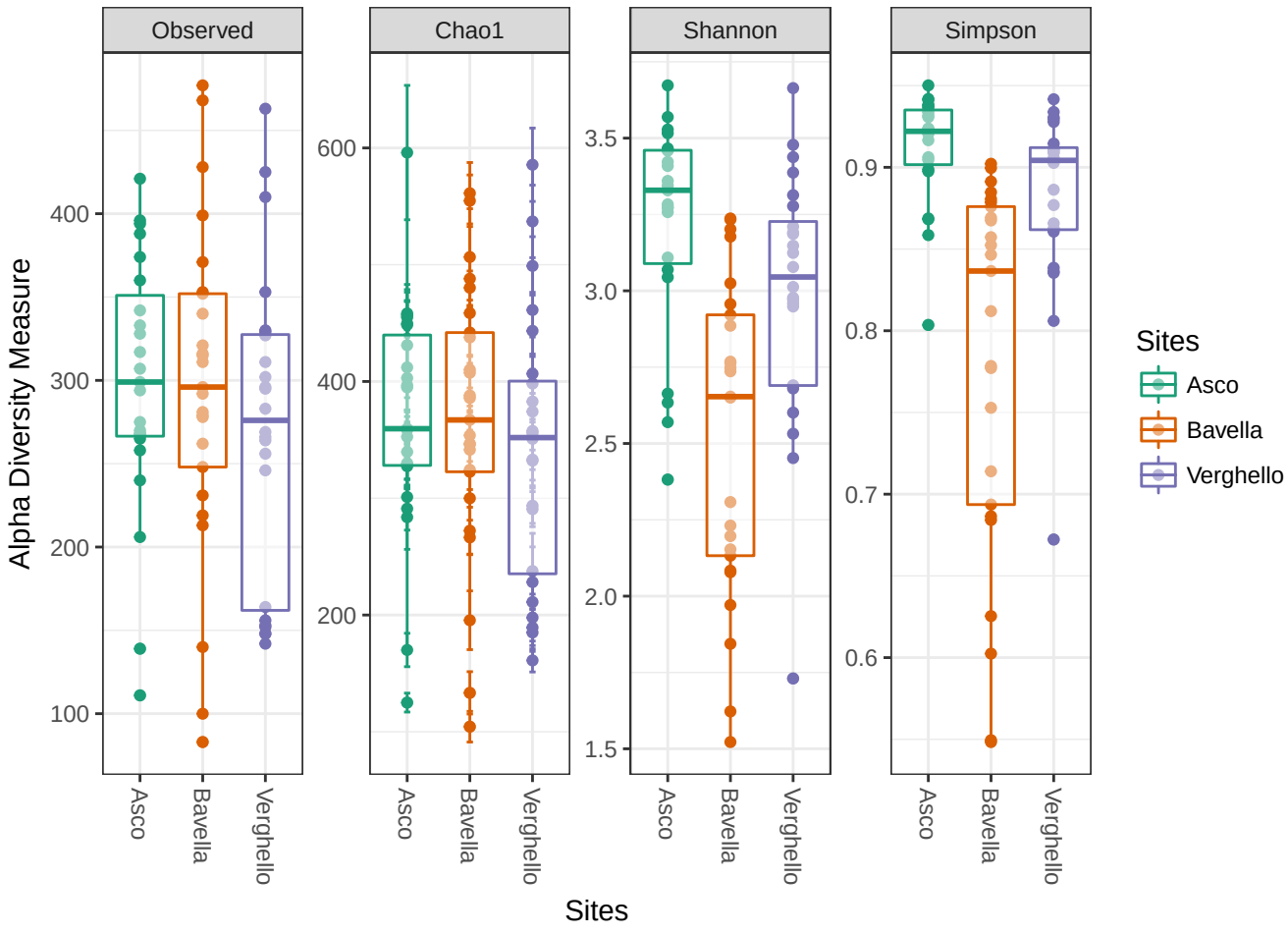



**Figure 5.3:** Diversity of each sites

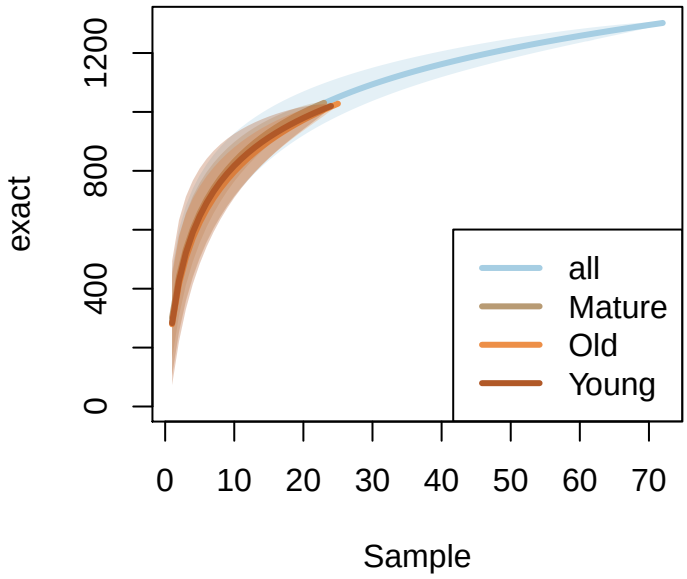



**Figure 5.4:** Rarefaction curves for each host age. Note that if singletons were removed, these curves are biaised.

```
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) | 62.8502239 | 49.5872070 | 1.2674685 | 0.2095802 |
| sqrt(readNumbers) | 0.7288918 | 0.1267135 | 5.7522816 | 0.0000003 |
| data.f3@sam_data$SitesBavella | 9.9500272 | 22.0573697 | 0.4510976 | 0.6534440 |
| data.f3@sam_data$SitesVerghello | -24.5733747 | 22.0808586 | -1.1128813 | 0.2699220 |
| data.f3@sam_data$AgeOld | -13.7853535 | 21.9625265 | -0.6276761 | 0.5324494 |
| data.f3@sam_data$AgeYoung | -37.0937028 | 22.4841290 | -1.6497727 | 0.1038895 |
| data.f3@sam_data$ElevationMiddle | 20.1789158 | 22.2394735 | 0.9073468 | 0.3676272 |
| data.f3@sam_data$ElevationTop | -5.5313782 | 21.9385577 | -0.2521304 | 0.8017482 |

**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))
```

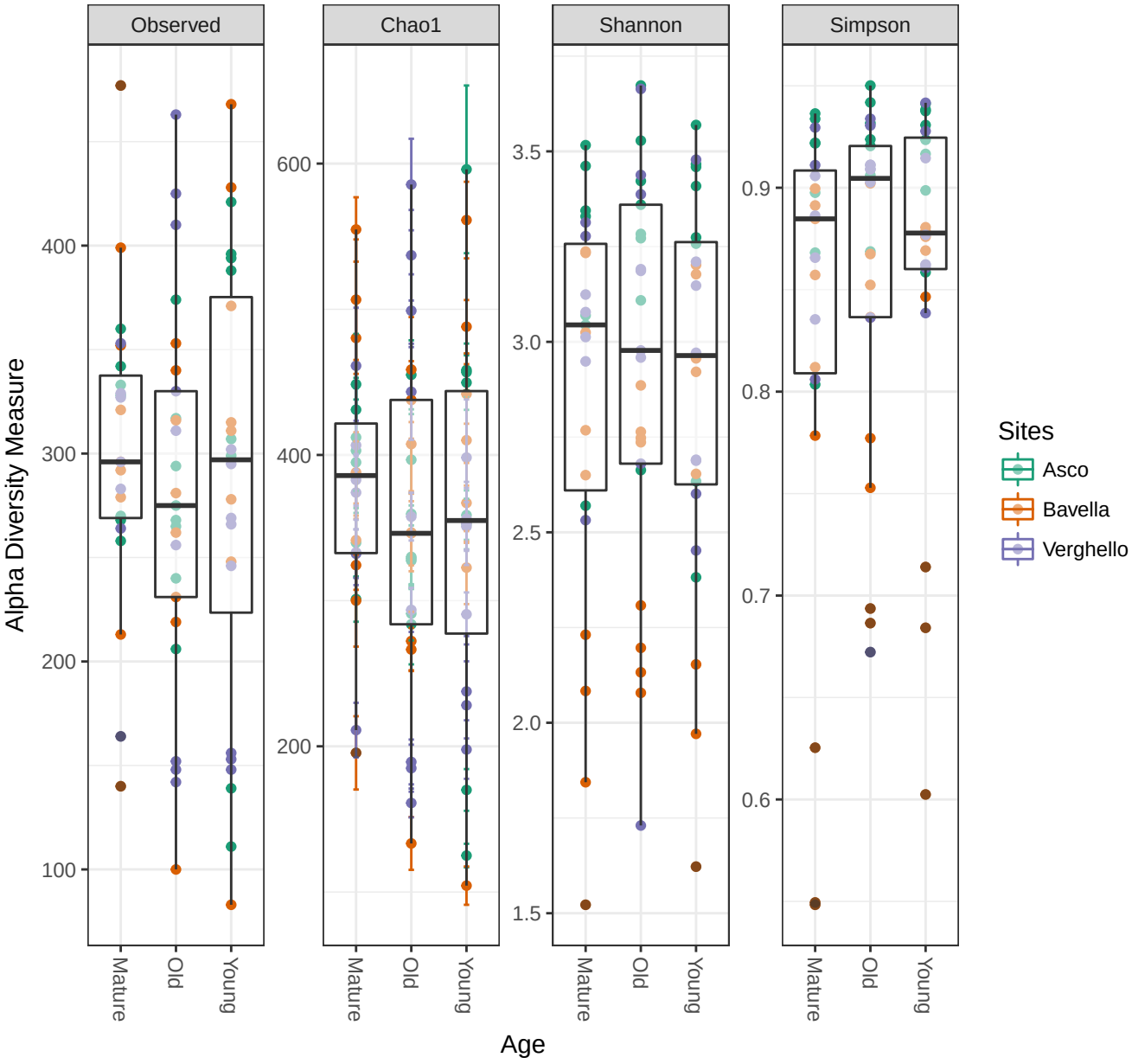



**Figure 5.5:** Diversity in function of tree age. Color represent sites.

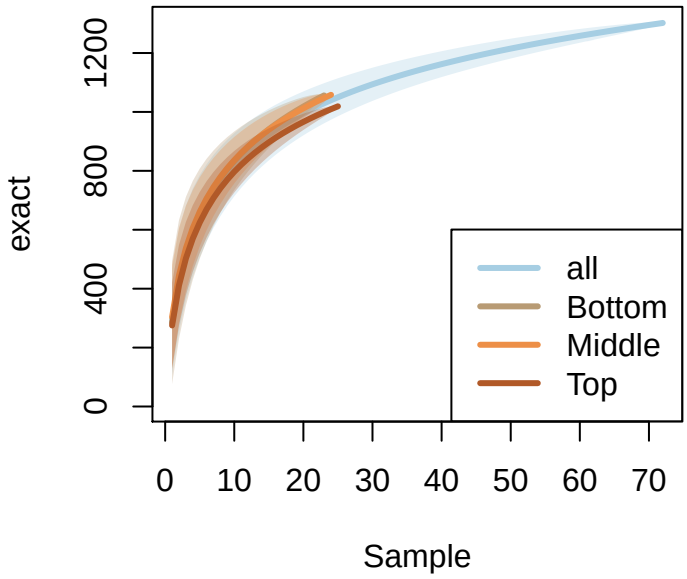



**Figure 5.6:** Rarefaction curves for each elevation. Notes that if singletons were removed, these curves are biaised.

```
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("")
```

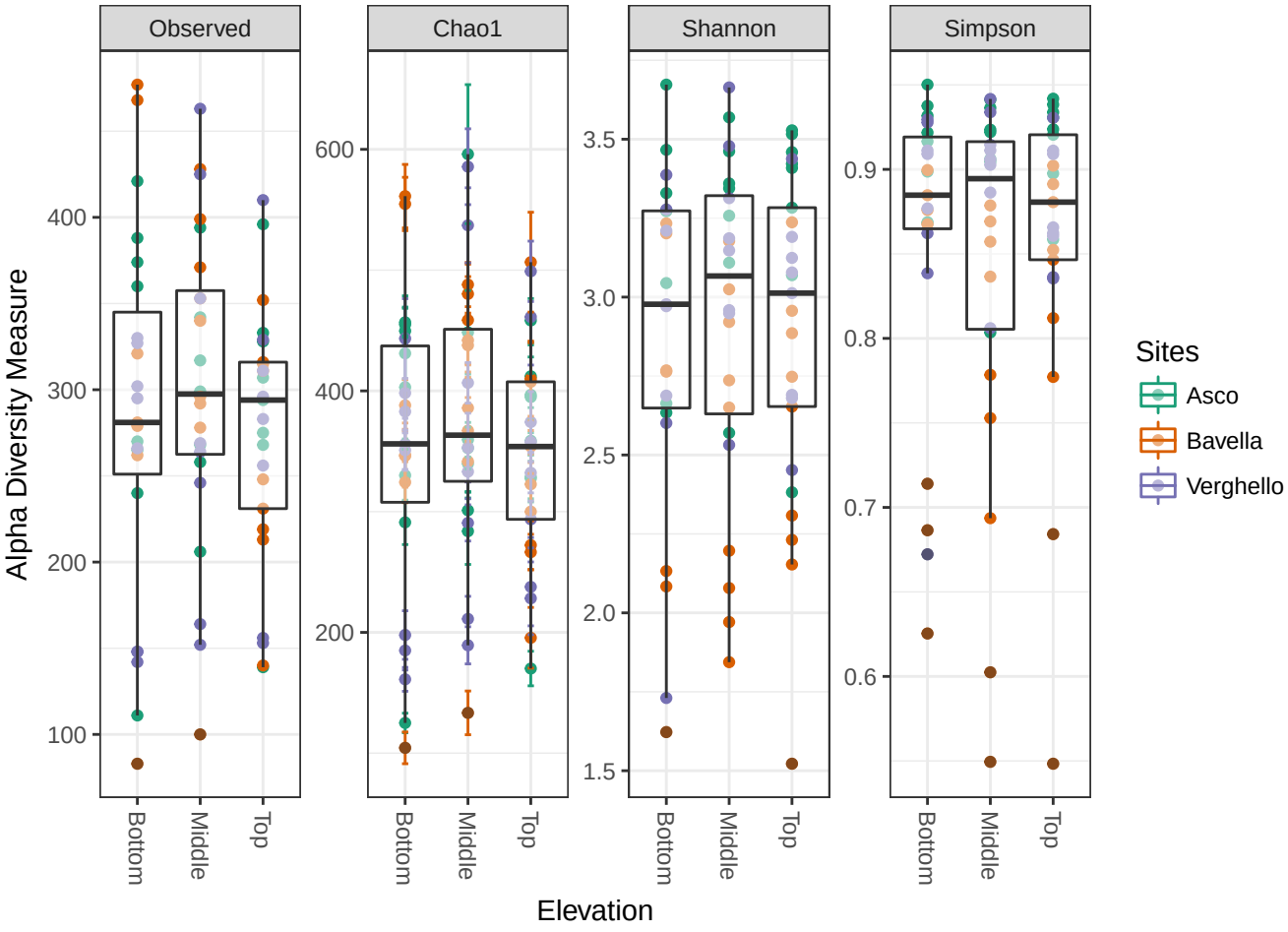



**Figure 5.7:** Diversity in function of elevation. Color represent sites.

| | Estimate | Std. Error | t value | Pr(>\|t\|) |
|---|---|---|---|---|
| (Intercept) | 15.9643217 | 4.7426774 | 3.3660990 | 0.0012932 |
| sqrt(readNumbers) | 0.0273900 | 0.0121193 | 2.2600329 | 0.0272300 |
| data.f3@sam_data$SitesBavella | -11.8649769 | 2.1096366 | -5.6241804 | 0.0000004 |
| data.f3@sam_data$SitesVerghello | -4.9746545 | 2.1118832 | -2.3555539 | 0.0215712 |
| data.f3@sam_data$AgeOld | 1.0906065 | 2.1005655 | 0.5191966 | 0.6054145 |
| data.f3@sam_data$AgeYoung | -0.9054816 | 2.1504532 | -0.4210655 | 0.6751180 |
| data.f3@sam_data$ElevationMiddle | 2.1257168 | 2.1270536 | 0.9993715 | 0.3213784 |
| data.f3@sam_data$ElevationTop | 1.0378518 | 2.0982731 | 0.4946219 | 0.6225607 |

**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) | 9.3970926 | 2.2046679 | 4.2623619 | 0.0000678 |
| sqrt(readNumbers) | 0.0071962 | 0.0056337 | 1.2773406 | 0.2060965 |
| data.f3@sam_data$SitesBavella | -6.6403487 | 0.9806799 | -6.7711687 | 0.0000000 |
| data.f3@sam_data$SitesVerghello | -2.8126265 | 0.9817242 | -2.8649865 | 0.0056353 |
| data.f3@sam_data$AgeOld | 1.1277239 | 0.9764631 | 1.1549068 | 0.2524225 |
| data.f3@sam_data$AgeYoung | 0.7541150 | 0.9996537 | 0.7543762 | 0.4533902 |
| data.f3@sam_data$ElevationMiddle | 0.2513083 | 0.9887763 | 0.2541610 | 0.8001860 |
| data.f3@sam_data$ElevationTop | 0.1848083 | 0.9753974 | 0.1894698 | 0.8503243 |

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

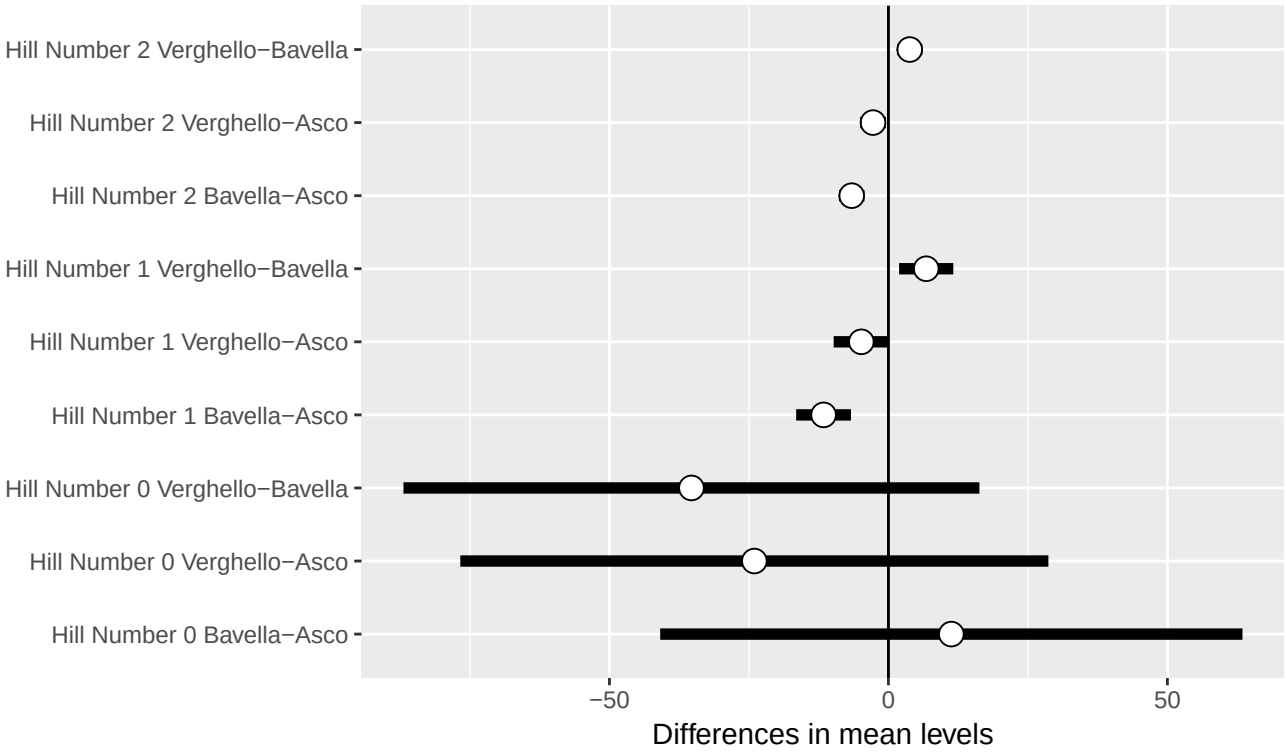



**Figure 5.8:** Results of the Tuckey HSD testing for differences in mean Hill numbers among pairs of modalities

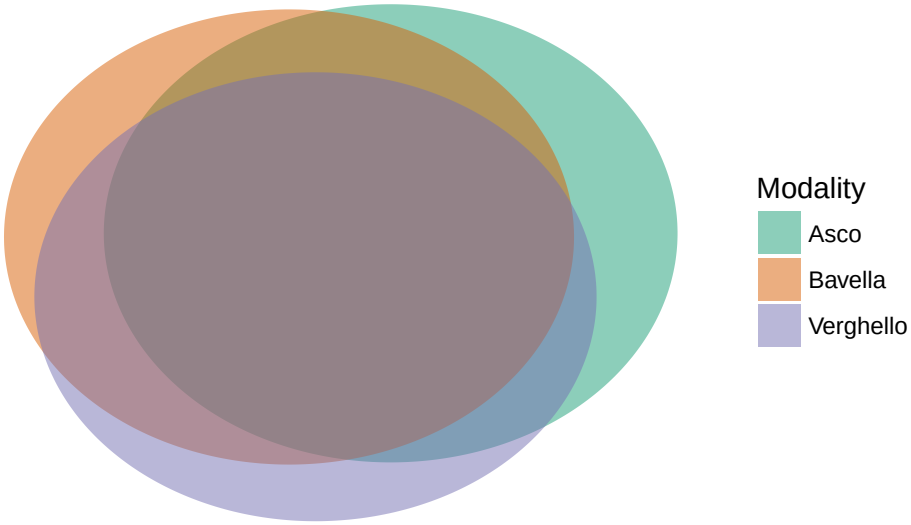



**Figure 6.1:** Venn diagramm of the distribution of OTUs among Sites

# 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 MultiDimenstional Scaling).

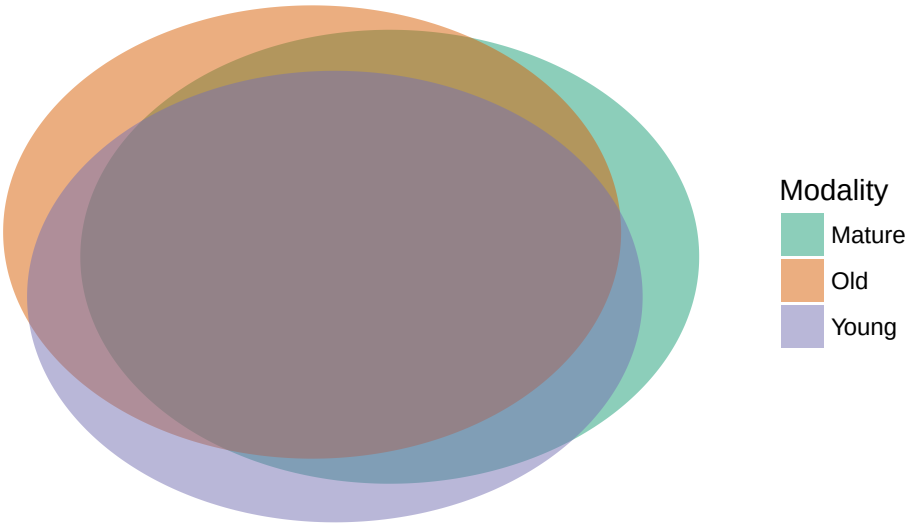



**Figure 6.2:** Venn diagramm of the distribution of OTUs among host age

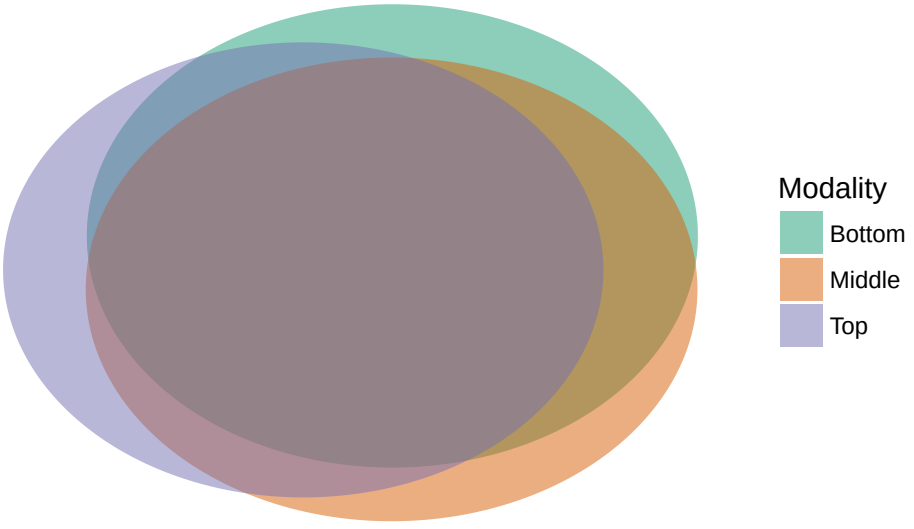



**Figure 6.3:** Venn diagramm of the distribution of OTUs among elevation of samples

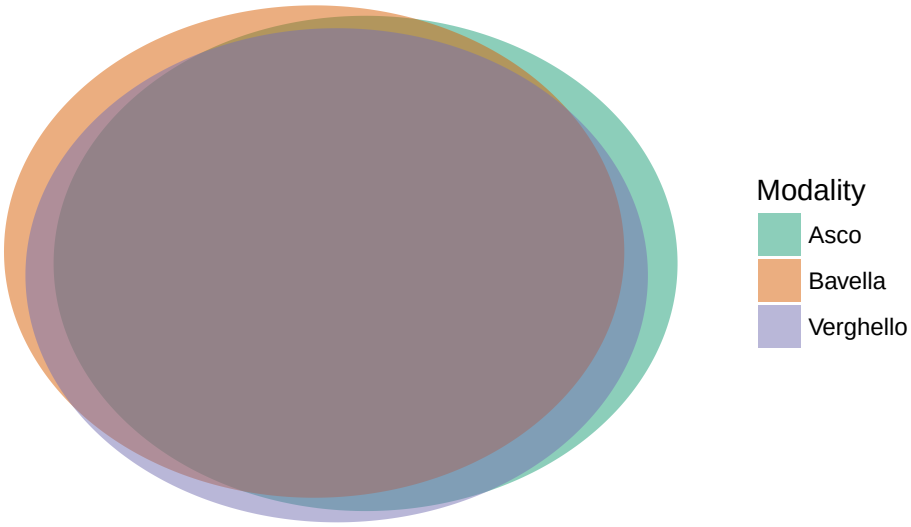



**Figure 6.4:** Venn diagramm of the distribution of OTUs among Sites

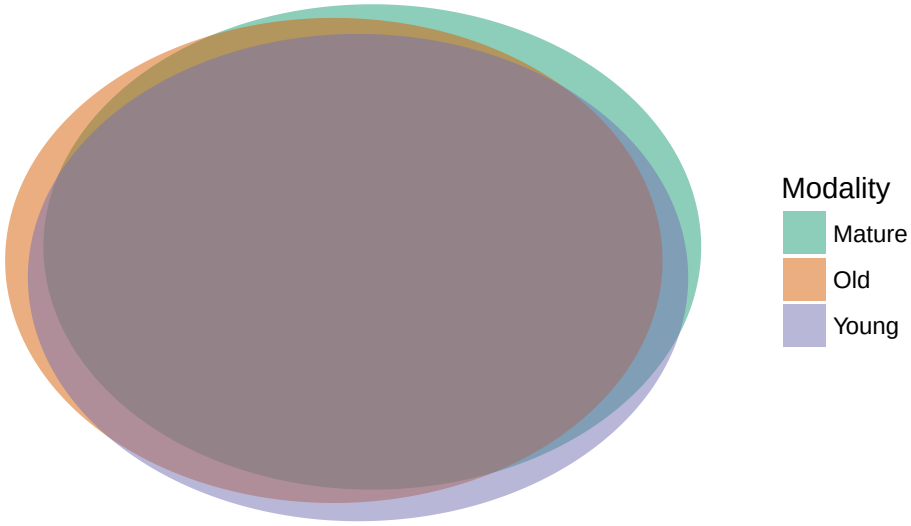



**Figure 6.5:** Venn diagramm of the distribution of OTUs among host age

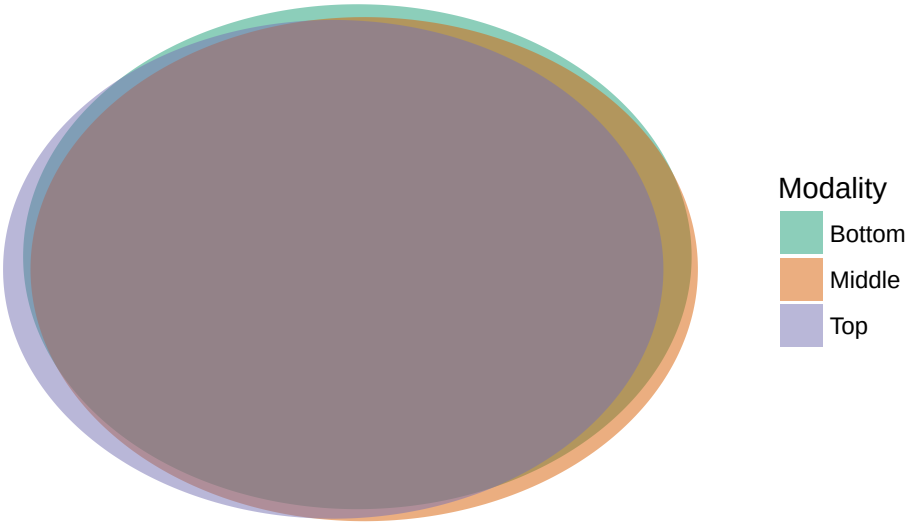



**Figure 6.6:** Venn diagramm of the distribution of OTUs among elevation of samples whitin the tree

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

```
stressplot(my.ord.nmds)
```

```
p <- plot_ordination(data.f3, my.ord.nmds, 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.nmds_gower <- ordinate(data.f3, distance = "gower",  method = "NMDS")

## Square root transformation
## Wisconsin double standardization
## Run 0 stress 0.2393401
## Run 1 stress 0.2438282
## Run 2 stress 0.4083443
## Run 3 stress 0.2425071
## Run 4 stress 0.4083585
## Run 5 stress 0.2406903
## Run 6 stress 0.2464409
## Run 7 stress 0.2394845
## ... Procrustes: rmse 0.01228488  max resid 0.07431725
## Run 8 stress 0.2394414
## ... Procrustes: rmse 0.01007789  max resid 0.06804078
## Run 9 stress 0.2452266
## Run 10 stress 0.239005
## ... New best solution
## ... Procrustes: rmse 0.0104553  max resid 0.07359731
```

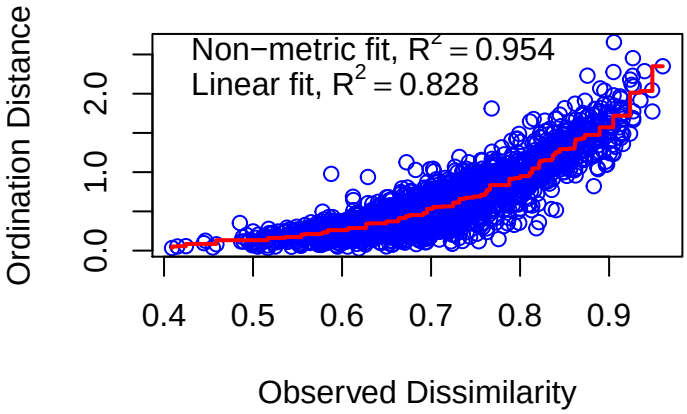



**Figure 6.7:** Stress plot of the NMDS

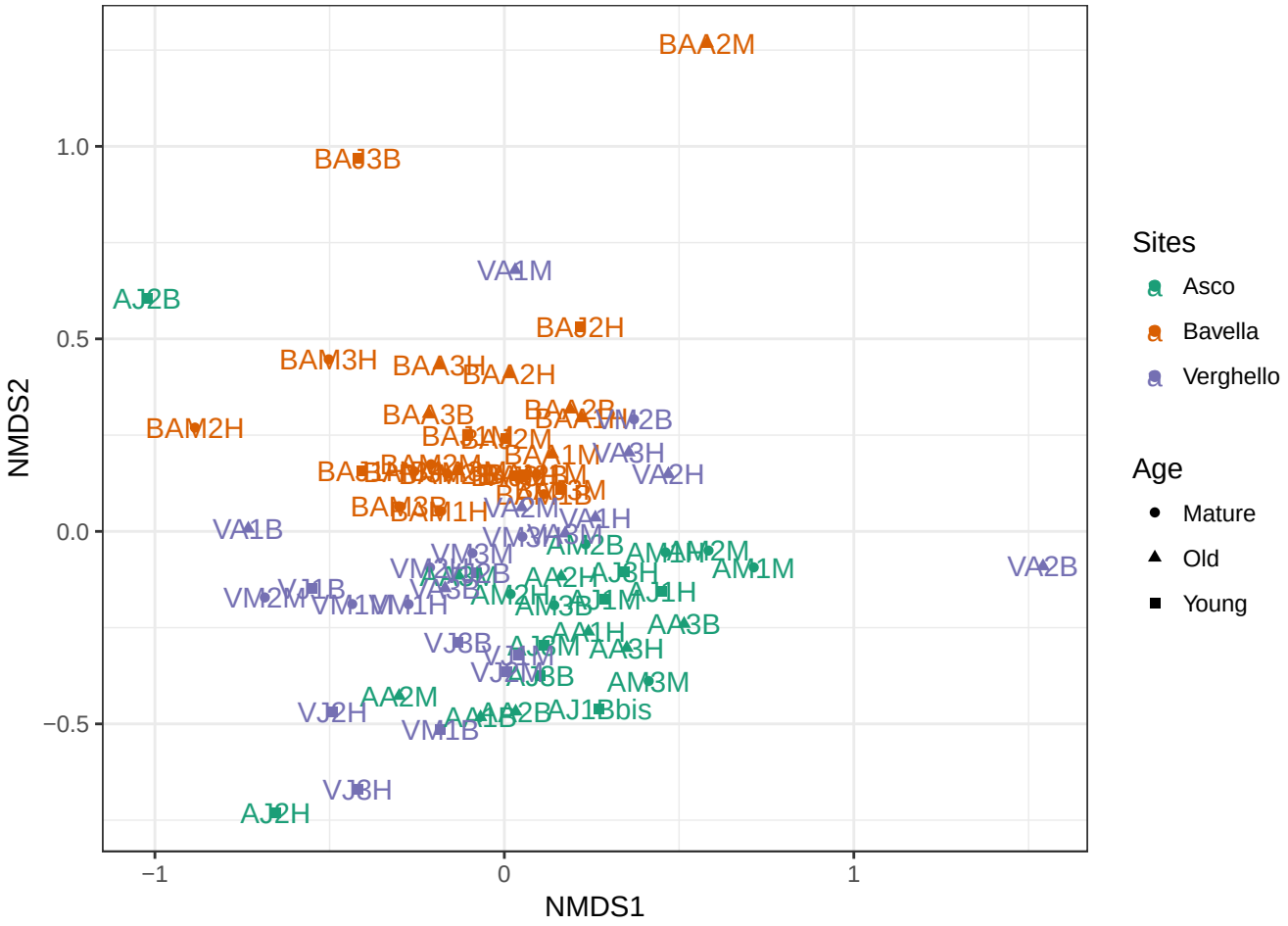



**Figure 6.8:** NMDS of OTU table. Colors represent sites and shape the age of tree.

```
## Run 11 stress 0.2399895
## Run 12 stress 0.2411228
## Run 13 stress 0.2424844
## Run 14 stress 0.2480004
## Run 15 stress 0.2418689
## Run 16 stress 0.2393749
## ... Procrustes: rmse 0.01189761  max resid 0.0778224
## Run 17 stress 0.2403961
## Run 18 stress 0.2448713
## Run 19 stress 0.2469104
## Run 20 stress 0.2480753
## *** No convergence -- monoMDS stopping criteria:
##      3: no. of iterations >= maxit
##     17: stress ratio > sratmax

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)
```

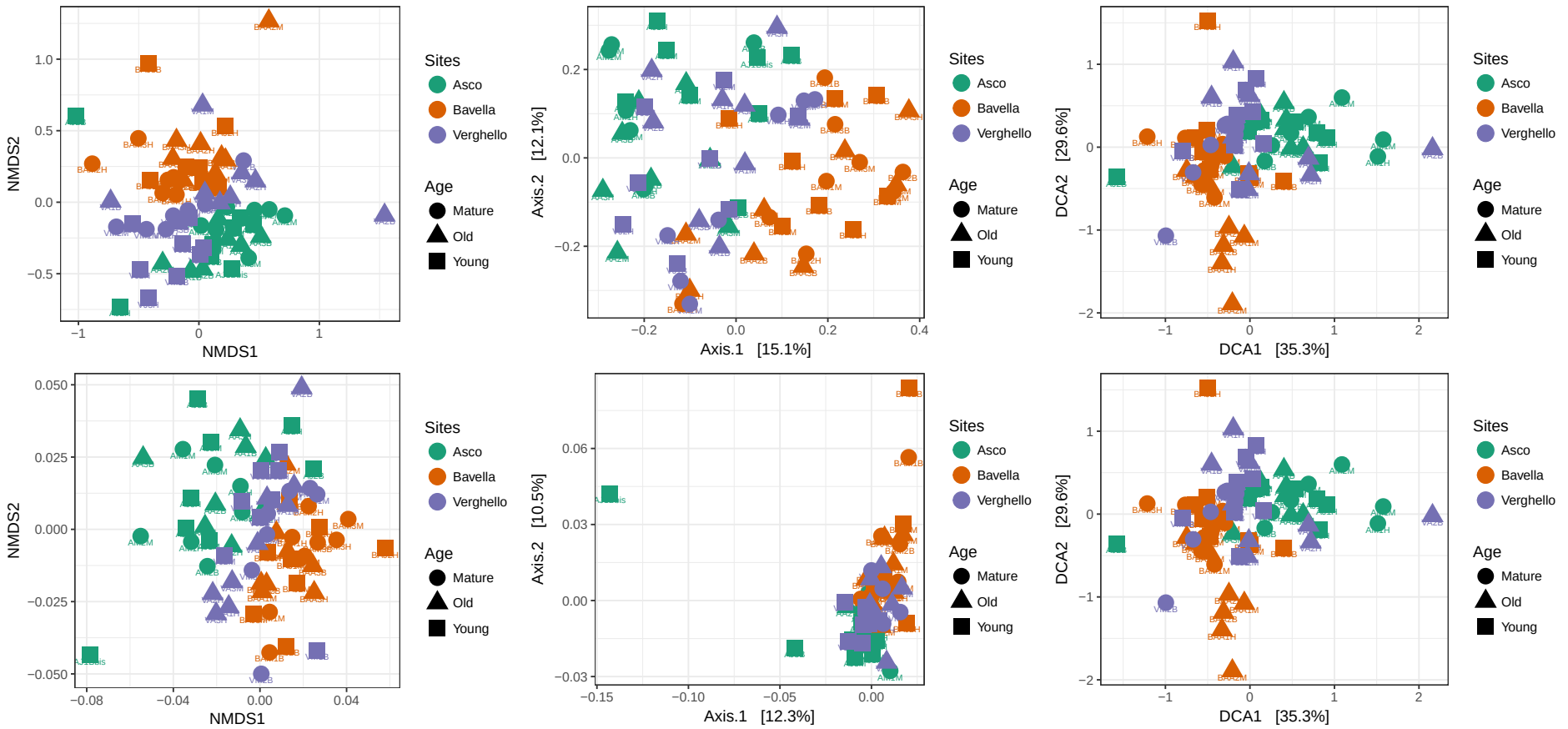



**Figure 6.9:** Comparison of different distances (bray (up) and gower (bottom)) and ordination methods (NMDS (left), PCoA (center) and DCA (right)).

```
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 264 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 | 2.03 | 1.02 | 5.41 | 0.13 | 0.0001 |
| Age | 2 | 0.63 | 0.32 | 1.68 | 0.04 | 0.0122 |
| Elevation | 2 | 0.48 | 0.24 | 1.29 | 0.03 | 0.1195 |
| Sites:Age | 4 | 1.42 | 0.35 | 1.88 | 0.09 | 0.0002 |
| Sites:Elevation | 4 | 0.68 | 0.17 | 0.90 | 0.04 | 0.6973 |
| Age:Elevation | 4 | 0.81 | 0.20 | 1.07 | 0.05 | 0.3267 |
| Sites:Age:Elevation | 8 | 1.43 | 0.18 | 0.95 | 0.09 | 0.6260 |
| Residuals | 45 | 8.45 | 0.19 | | 0.53 | |
| Total | 71 | 15.93 | | | 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 | 1.98 | 0.99 | 5.59 | 0.13 | 0.0001 |
| Age | 2 | 0.61 | 0.30 | 1.71 | 0.04 | 0.0111 |
| Elevation | 2 | 0.47 | 0.23 | 1.32 | 0.03 | 0.1197 |
| Sites:Age | 4 | 1.37 | 0.34 | 1.93 | 0.09 | 0.0002 |
| Sites:Elevation | 4 | 0.64 | 0.16 | 0.90 | 0.04 | 0.6962 |
| Age:Elevation | 4 | 0.77 | 0.19 | 1.08 | 0.05 | 0.3080 |
| Sites:Age:Elevation | 8 | 1.33 | 0.17 | 0.94 | 0.09 | 0.6591 |
| Residuals | 45 | 7.98 | 0.18 | | 0.53 | |
| Total | 71 | 15.14 | | | 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 | 0.88 | 0.44 | 3.66 | 0.09 | 0.0001 |
| Age | 2 | 0.44 | 0.22 | 1.83 | 0.05 | 0.0012 |
| Elevation | 2 | 0.25 | 0.13 | 1.04 | 0.03 | 0.3725 |
| Sites:Age | 4 | 0.69 | 0.17 | 1.43 | 0.07 | 0.0062 |
| Sites:Elevation | 4 | 0.39 | 0.10 | 0.82 | 0.04 | 0.9231 |
| Age:Elevation | 4 | 0.52 | 0.13 | 1.08 | 0.05 | 0.2686 |
| Sites:Age:Elevation | 8 | 0.88 | 0.11 | 0.92 | 0.09 | 0.7744 |
| Residuals | 45 | 5.42 | 0.12 | | 0.57 | |
| Total | 71 | 9.47 | | | 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 | 2.03 | 1.02 | 6.22 | 0.13 | 0.0001 |
| Age | 2 | 0.63 | 0.32 | 1.93 | 0.04 | 0.0024 |
| Sites:Age | 4 | 1.42 | 0.35 | 2.17 | 0.09 | 0.0001 |
| Sites:Age:IndividualTree | 18 | 4.49 | 0.25 | 1.53 | 0.28 | 0.0001 |
| Residuals | 45 | 7.35 | 0.16 | | 0.46 | |
| Total | 71 | 15.93 | | | 1.00 | |

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

| | Df | SumsOfSqs | MeanSqs | F.Model | R2 | Pr(>F) |
|---|---|---|---|---|---|---|
| Sites | 2 | 1.98 | 0.99 | 6.48 | 0.13 | 0.0001 |
| Age | 2 | 0.61 | 0.30 | 1.98 | 0.04 | 0.0035 |
| Sites:Age | 4 | 1.37 | 0.34 | 2.24 | 0.09 | 0.0001 |
| Sites:Age:IndividualTree | 18 | 4.30 | 0.24 | 1.56 | 0.28 | 0.0001 |
| Residuals | 45 | 6.88 | 0.15 | | 0.45 | |
| Total | 71 | 15.14 | | | 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 | 0.88 | 0.44 | 4.13 | 0.09 | 0.0001 |
| Age | 2 | 0.44 | 0.22 | 2.07 | 0.05 | 0.0002 |
| Sites:Age | 4 | 0.70 | 0.18 | 1.65 | 0.07 | 0.0005 |
| Sites:Age:IndividualTree | 18 | 2.64 | 0.15 | 1.38 | 0.28 | 0.0001 |
| Residuals | 45 | 4.80 | 0.11 | | 0.51 | |
| Total | 71 | 9.47 | | | 1.00 | |

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

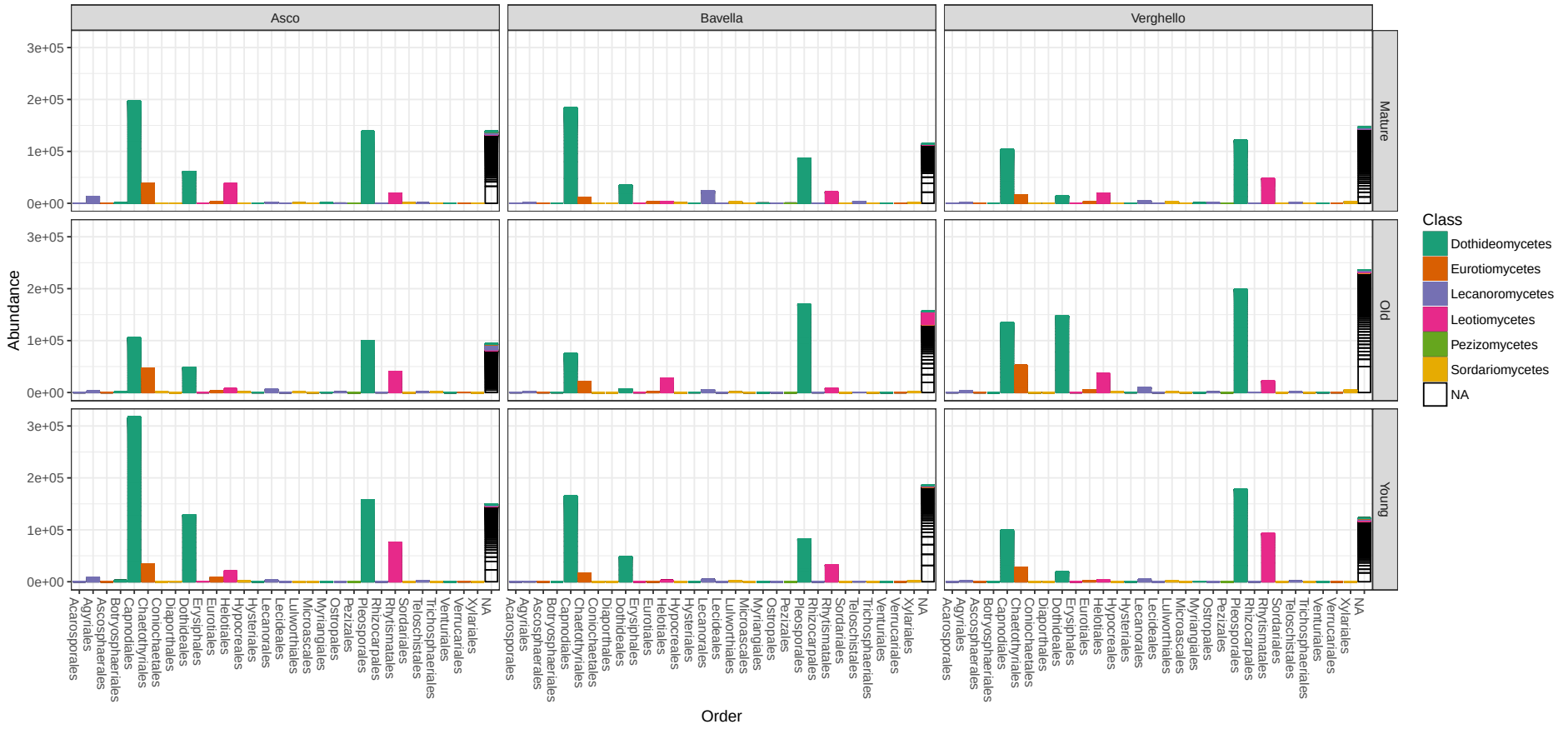



**Figure 6.10:** Taxonomic distribution of sequences in the different site * age combinaison.

```
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")
```

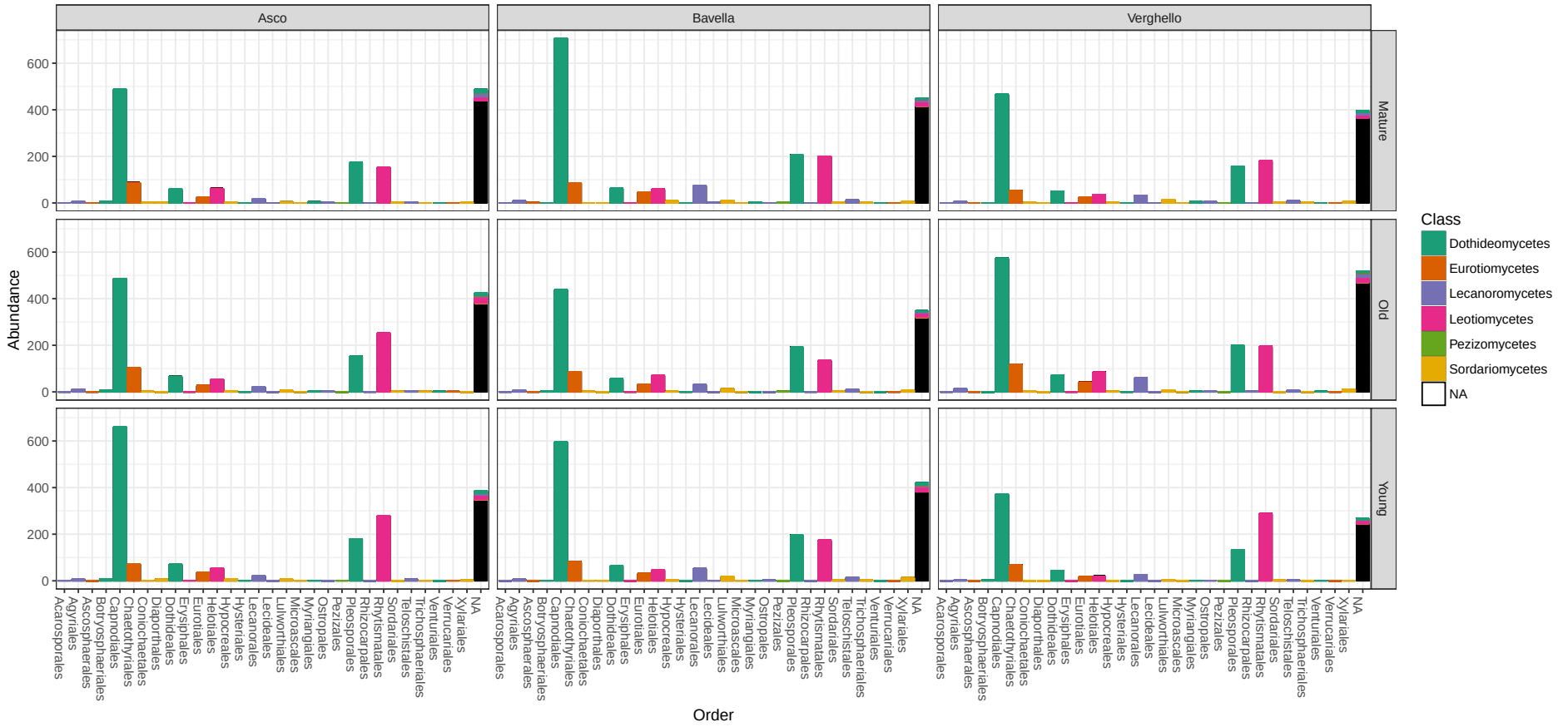



**Figure 6.11:** Taxonomic distribution of OTUs in the different site * age combinaison.

```
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
```

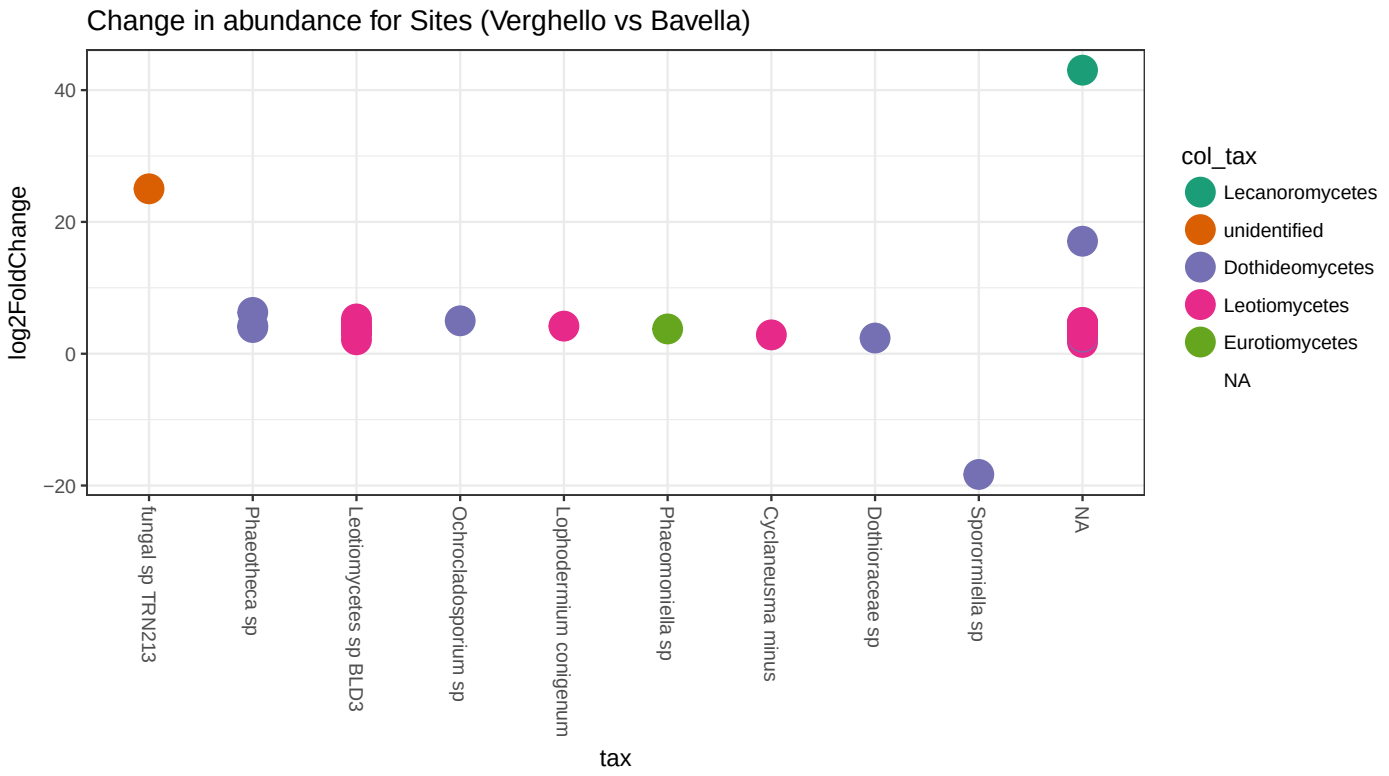



**Figure 6.13:** OTUs significantly different in terms of abundances between Verghello (positive values) and Bavella (negative values)

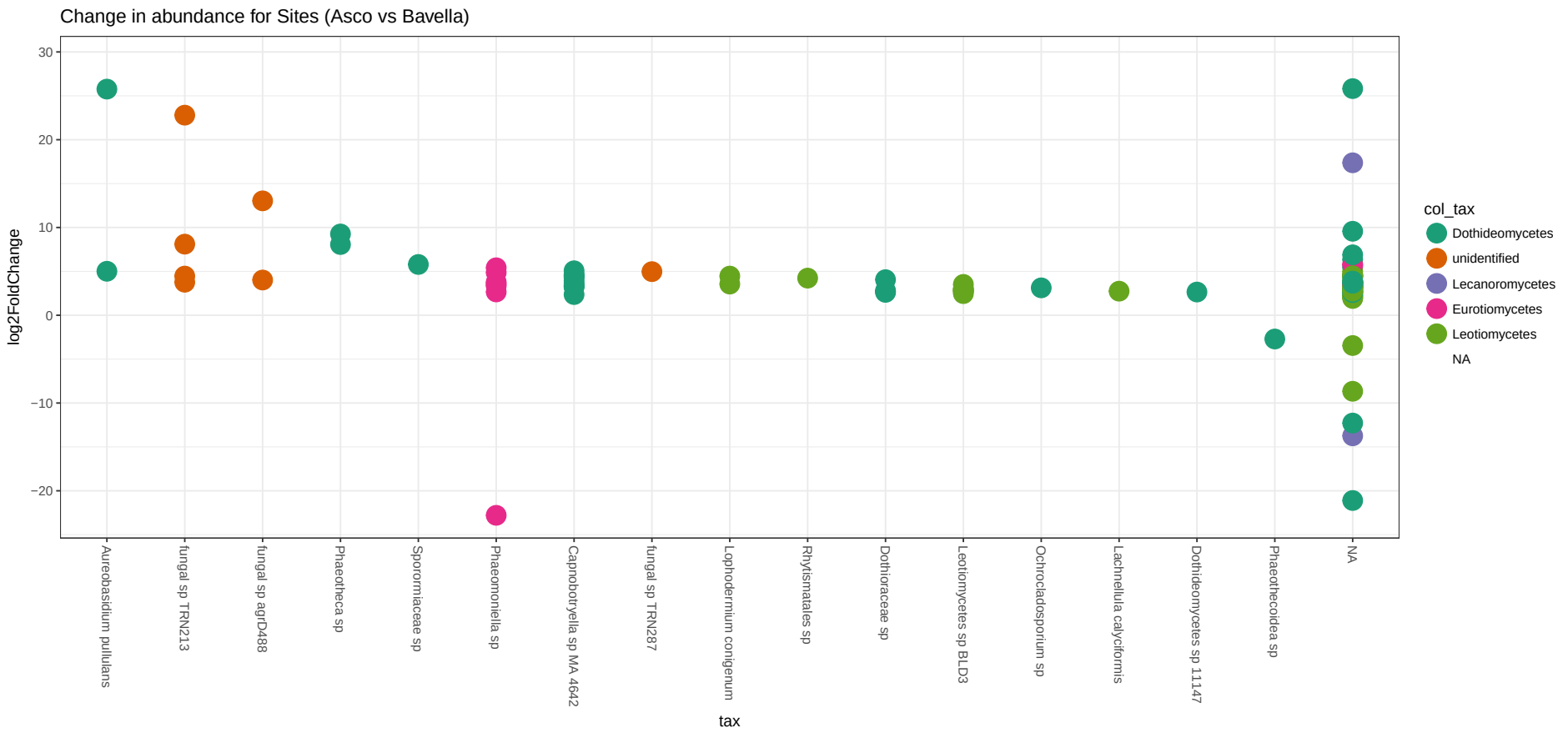



**Figure 6.14:** OTUs significantly different in terms of abundances between Asco (positive values) and Bavella (negative values)

```
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
```

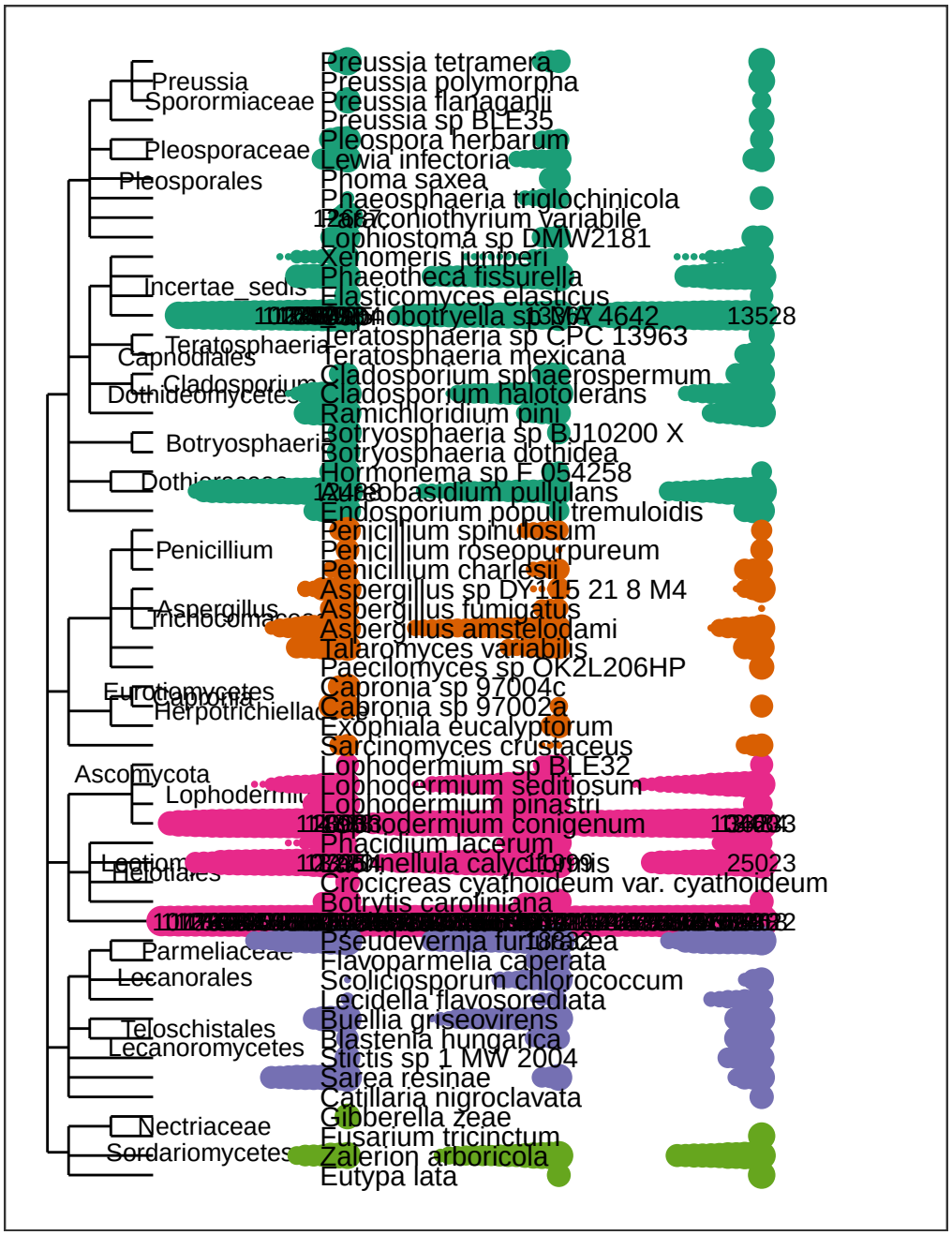
### 6.7.2 Pairwise comparison of Order composition by sites

```
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")
```

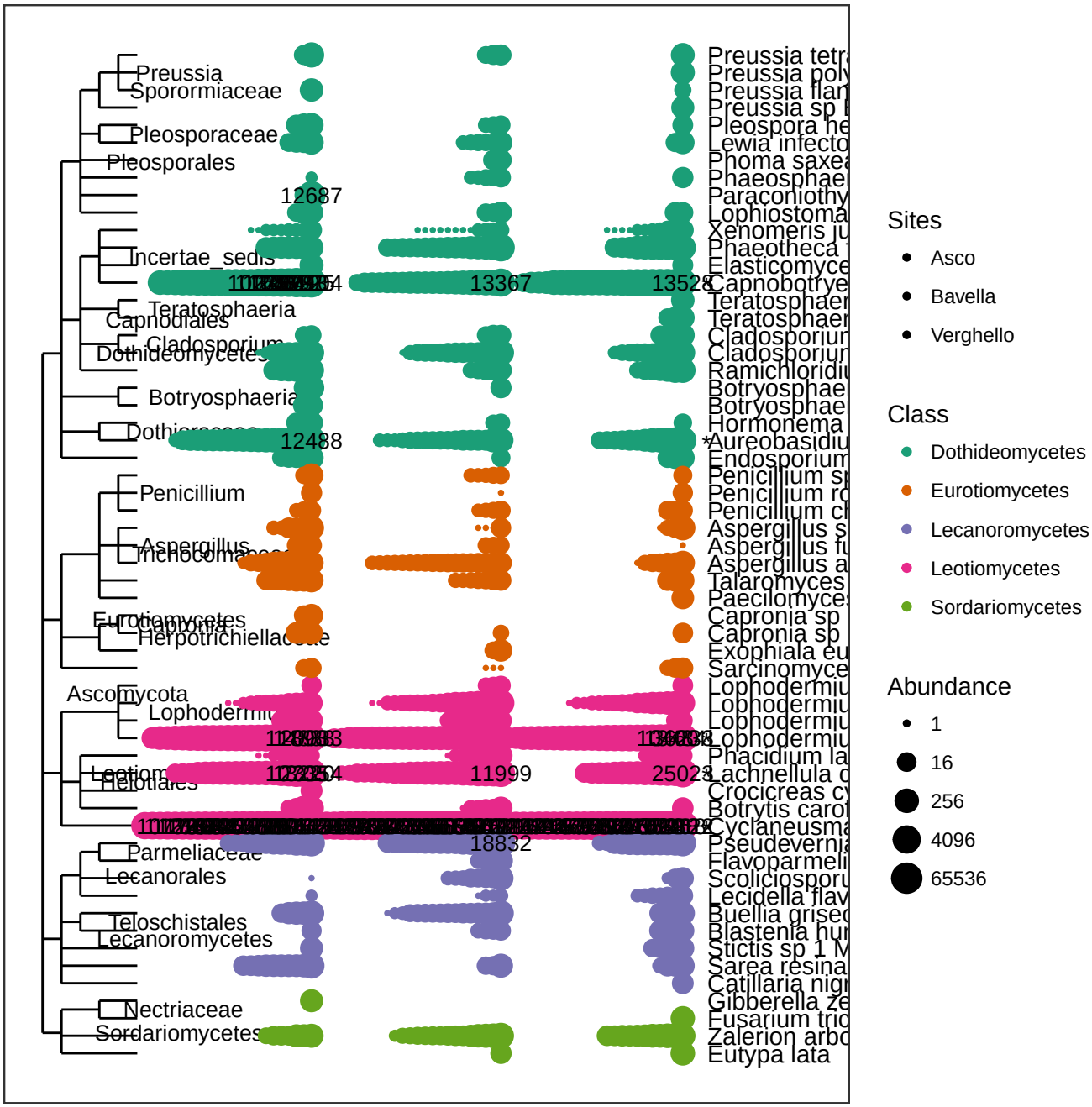
## 6.8 Distribution of OTUs abundance in the fungal phylogeny

```
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(da
data.f3_interm@tax_table <- tax_table(apply(data.f3@tax_table, 2, function(x) tapply(x, as.factor(da
data.f3_interm@refseq <- NULL

data.f3_interm <- subset_taxa(data.f3_interm, !grepl("uncultured", data.f3_interm@tax_table[,"Specie
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[,"Fami
data.f3_interm <- subset_taxa(data.f3_interm, !grepl("unidentified", data.f3_interm@tax_table[,"Orde
data.f3_interm <- subset_taxa(data.f3_interm, !grepl("unidentified", data.f3_interm@tax_table[,"Clas
data.f3_interm <- subset_taxa(data.f3_interm, !grepl("Myxotrichaceae", data.f3_interm@tax_table[,"Sp
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)),

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 = '

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

ptree + geom_text(data = df_cond, aes(x = 215, y = y, label = OTU), hjust = "left") + scale_shape_ma

sum(!is.na(match(gsub("_", " ", data.f3@tax_table[,"Species"]), gsub("_", " ", tree$tip.label))))

## [1] 269

sum(rowSums(data.f3@otu_table)[gsub("_", " ", data.f3@tax_table[,"Species"]) %in% gsub("_", " ", tre

## [1] 38.5203
```




```
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 $8.335341 \times 10^6$ sequences representing 1667 OTUs allocated to 80 samples.

After filtering, the dataset includes $8.312594 \times 10^6$ sequences representing 1302 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 (R2 = 0.128), age (R2 = 0.04) and interaction age*site (R2 = 0.089) statistically structured the fungal endophytic beta-diversity.

## 7.4 Special case of *Cyclaneusma minus*

*Cyclaneusma minus* account for 23.38% of total sequences.

```
cycla <- as.vector(data.f3@otu_table["OTU_1",]/ colSums(data.f3@otu_table) * 100)

par(mfrow=c(1,2))
boxplot(cycla~data.f3@sam_data$Age, ylab="% Cyclaneusma minus")
boxplot(cycla~data.f3@sam_data$Elevation, ylab="% Cyclaneusma minus")
boxplot(cycla~data.f3@sam_data$Sites, ylab="% Cyclaneusma minus")
par(mfrow=c(1,1))
```

```
tapply(cycla, data.f3@sam_data$Age, mean)

##   Mature      Old    Young
## 28.05235 22.51140 24.01578

tapply(cycla, data.f3@sam_data$Age, sd)

##   Mature      Old    Young
## 18.05840 13.30125 14.48586

tapply(cycla, data.f3@sam_data$Elevation, mean)

##   Bottom   Middle      Top
## 23.94633 27.48506 22.95842

tapply(cycla, data.f3@sam_data$Elevation, sd)

##   Bottom   Middle      Top
## 14.36857 16.84100 14.85478

tapply(cycla, data.f3@sam_data$Sites, mean)

##      Asco   Bavella Verghello
##  16.12246  36.73680  20.63047

tapply(cycla, data.f3@sam_data$Sites, sd)

##      Asco   Bavella Verghello
##  6.755436 17.731830 10.160594

wilcox.test(cycla[data.f3@sam_data$Sites=="Asco"], cycla[data.f3@sam_data$Sites=="Bavella"])

##
##  Wilcoxon rank sum test
##
## data:  cycla[data.f3@sam_data$Sites == "Asco"] and cycla[data.f3@sam_data$Sites == "Bavella"]
## W = 77, p-value = 3.837e-06
## alternative hypothesis: true location shift is not equal to 0
```

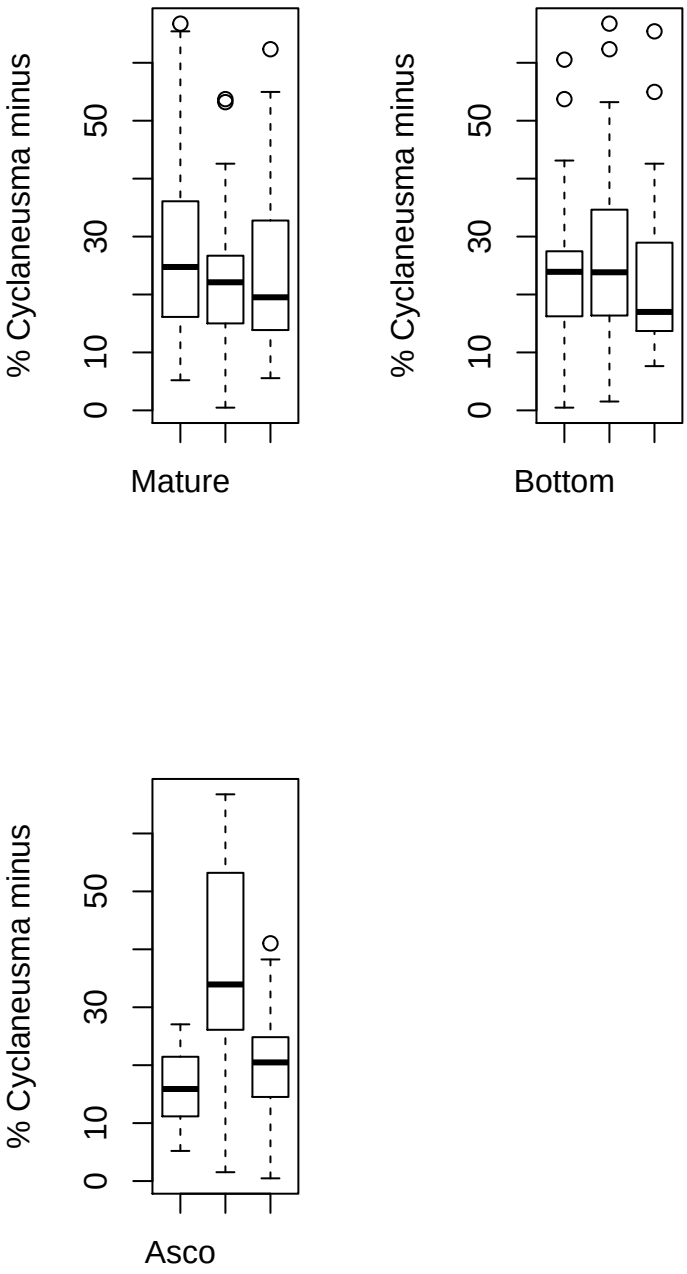



**Figure 7.1:** Number of sequences assigned to extitCyclaneusma minus across host age, elevation whitin tree and sites

```r
wilcox.test(cycla[data.f3@sam_data$Sites=="Verghello"], cycla[data.f3@sam_data$Sites=="Bavella"])

##
##  Wilcoxon rank sum test
##
## data:  cycla[data.f3@sam_data$Sites == "Verghello"] and cycla[data.f3@sam_data$Sites == "Bavella"
## W = 130, p-value = 0.000477
## alternative hypothesis: true location shift is not equal to 0

wilcox.test(cycla[data.f3@sam_data$Sites=="Asco"], cycla[data.f3@sam_data$Sites=="Verghello"])

##
##  Wilcoxon rank sum test
##
## data:  cycla[data.f3@sam_data$Sites == "Asco"] and cycla[data.f3@sam_data$Sites == "Verghello"]
## W = 204, p-value = 0.129
## alternative hypothesis: true location shift is not equal to 0
```

| | Comparison | OTU_names | Species | Class | log2FoldChange (negative = more on second levels) |
|---|---|---|---|---|---|
| 1 | Verghello vs Asco | OTU_885 | Phaeomoniella sp | Eurotiomycetes | -3.14180047934908 |
| 2 | Verghello vs Asco | OTU_263 | | | 3.64844007621285 |
| 3 | Verghello vs Asco | OTU_1192 | | Dothideomycetes | -4.06191788385273 |
| 4 | Verghello vs Asco | OTU_805 | | Dothideomycetes | -3.74527277873544 |
| 5 | Verghello vs Asco | OTU_688 | fungal sp TRN213 | unidentified | -3.05716134156103 |
| 6 | Verghello vs Asco | OTU_35 | | | 3.45787058219993 |
| 7 | Verghello vs Asco | OTU_861 | | | -2.84255276977124 |
| 8 | Verghello vs Asco | OTU_633 | Phaeotheca sp | Dothideomycetes | -3.79929856510545 |
| 9 | Verghello vs Asco | OTU_742 | | | 3.27422748242967 |
| 10 | Verghello vs Asco | OTU_570 | | Dothideomycetes | -2.95237737337865 |
| 11 | Verghello vs Asco | OTU_1203 | | Dothideomycetes | -3.37772312072497 |
| 12 | Verghello vs Asco | OTU_1546 | Capnobotryella sp MA 4642 | Dothideomycetes | -2.8270261217971 |
| 13 | Verghello vs Asco | OTU_399 | | Dothideomycetes | -3.24628121566645 |
| 14 | Verghello vs Asco | OTU_669 | | Dothideomycetes | -6.11297624195951 |
| 15 | Verghello vs Asco | OTU_1148 | Phaeomoniella sp | Eurotiomycetes | -2.46378616499525 |
| 16 | Verghello vs Asco | OTU_575 | | | 4.71085075476014 |
| 17 | Verghello vs Asco | OTU_1168 | | Dothideomycetes | -3.13089568886969 |
| 18 | Verghello vs Asco | OTU_809 | | Dothideomycetes | -2.80476330884076 |
| 19 | Verghello vs Asco | OTU_1285 | | Leotiomycetes | -2.0352037582083 |
| 20 | Verghello vs Asco | OTU_930 | | | 2.90203226388763 |
| 21 | Verghello vs Asco | OTU_25 | Phaeotheca sp | Dothideomycetes | 3.20057595263818 |
| 22 | Verghello vs Asco | OTU_18 | | Dothideomycetes | -5.82900067867463 |
| 23 | Verghello vs Asco | OTU_1050 | Cyclaneusma minus | Leotiomycetes | -2.1938663310028 |
| 24 | Verghello vs Asco | OTU_895 | | Dothideomycetes | -3.59194016172022 |
| 25 | Verghello vs Asco | OTU_1138 | | | 2.73811574455171 |
| 26 | Verghello vs Asco | OTU_72 | Rhytismatales sp | Leotiomycetes | -5.71309567878076 |
| 27 | Verghello vs Asco | OTU_33 | Aureobasidium pullulans | Dothideomycetes | -25.2920407676653 |
| 28 | Verghello vs Asco | OTU_30 | fungal sp agrD488 | unidentified | -8.90732449255831 |
| 29 | Verghello vs Asco | OTU_326 | Phaeotheca sp | Dothideomycetes | 2.90163549188653 |
| 30 | Verghello vs Asco | OTU_379 | | | -5.94480193488885 |
| 31 | Verghello vs Asco | OTU_826 | | Dothideomycetes | -3.70188271469607 |
| 32 | Verghello vs Asco | OTU_42 | | | -28.4093647742594 |
| 33 | Verghello vs Asco | OTU_21 | Sporormiaceae sp | Dothideomycetes | -6.12827048914855 |
| 34 | Verghello vs Asco | OTU_634 | | | -24.9062387805127 |
| 35 | Verghello vs Asco | OTU_1436 | | Dothideomycetes | -2.6205041473203 |
| 36 | Verghello vs Asco | OTU_54 | Sporormiella sp | Dothideomycetes | -25.378275998241 |
| 37 | Verghello vs Asco | OTU_1078 | Sporormiaceae sp | Dothideomycetes | -5.79837614995541 |
| 38 | Verghello vs Asco | OTU_188 | | | -24.7487110108513 |
| 39 | Verghello vs Asco | OTU_159 | Lophodermium conigenum | Leotiomycetes | -4.31908667235487 |
| 40 | Verghello vs Asco | OTU_1369 | | Dothideomycetes | -8.74996839038013 |
| 41 | Verghello vs Asco | OTU_149 | Phaeomoniella sp | Eurotiomycetes | 20.5726228422299 |
| 42 | Verghello vs Asco | OTU_111 | | Dothideomycetes | 23.052758808263 |
| 43 | Verghello vs Asco | OTU_152 | | Lecanoromycetes | 20.8234902163824 |
| 44 | Verghello vs Asco | OTU_419 | | Dothideomycetes | 22.303198860941 |
| 45 | Verghello vs Asco | OTU_74 | | | 24.365344777204 |
| 46 | Verghello vs Asco | OTU_105 | | Lecanoromycetes | 25.6569833302234 |
| 47 | Verghello vs Bavella | OTU_13 | Phaeomoniella sp | Eurotiomycetes | 3.759462667366 |
| 48 | Verghello vs Bavella | OTU_9 | Lophodermium conigenum | Leotiomycetes | 4.188191232945 |
| 49 | Verghello vs Bavella | OTU_1123 | | Dothideomycetes | 2.60757748632336 |
| 50 | Verghello vs Bavella | OTU_1102 | Leotiomycetes sp BLD3 | Leotiomycetes | 3.40611194797604 |
| 51 | Verghello vs Bavella | OTU_17 | | | 2.63932960489047 |
| 52 | Verghello vs Bavella | OTU_189 | Phaeotheca sp | Dothideomycetes | 6.26969179588288 |
| 53 | Verghello vs Bavella | OTU_11 | | | 5.77270198491659 |
| 54 | Verghello vs Bavella | OTU_6 | Dothioraceae sp | Dothideomycetes | 2.38082031291577 |
| 55 | Verghello vs Bavella | OTU_39 | Leotiomycetes sp BLD3 | Leotiomycetes | 2.11331395697968 |
| 56 | Verghello vs Bavella | OTU_46 | | Dothideomycetes | 2.68995694269571 |
| 57 | Verghello vs Bavella | OTU_20 | | Eurotiomycetes | 4.49523983865664 |
| 58 | Verghello vs Bavella | OTU_299 | Ochrocladosporium sp | Dothideomycetes | 4.99363124461955 |
| 59 | Verghello vs Bavella | OTU_263 | | | 6.1863215387924 |
| 60 | Verghello vs Bavella | OTU_756 | | Dothideomycetes | 4.70424392971001 |
| 61 | Verghello vs Bavella | OTU_227 | | Leotiomycetes | 2.23867944260312 |
| 62 | Verghello vs Bavella | OTU_935 | Leotiomycetes sp BLD3 | Leotiomycetes | 3.4448021030654 |
| 63 | Verghello vs Bavella | OTU_744 | Cyclaneusma minus | Leotiomycetes | 2.85886284279939 |
| 64 | Verghello vs Bavella | OTU_1657 | Leotiomycetes sp BLD3 | Leotiomycetes | 3.34719535552037 |
| 65 | Verghello vs Bavella | OTU_35 | | | 7.99414113690649 |
| 66 | Verghello vs Bavella | OTU_1527 | | Leotiomycetes | 3.77565606014944 |
| 67 | Verghello vs Bavella | OTU_503 | | Leotiomycetes | 4.0707538748451 |
| 68 | Verghello vs Bavella | OTU_861 | | | 3.99778432992676 |
| 69 | Verghello vs Bavella | OTU_633 | Phaeotheca sp | Dothideomycetes | 4.25864744298029 |
| 70 | Verghello vs Bavella | OTU_1120 | | Leotiomycetes | 3.37101321682374 |
| 71 | Verghello vs Bavella | OTU_1255 | | Leotiomycetes | 3.47507576148164 |
| 72 | Verghello vs Bavella | OTU_1121 | Leotiomycetes sp BLD3 | Leotiomycetes | 3.21363087956089 |
| 73 | Verghello vs Bavella | OTU_729 | | Leotiomycetes | 4.7515217647916 |
| 74 | Verghello vs Bavella | OTU_742 | | | 2.69179592736329 |
| 75 | Verghello vs Bavella | OTU_1388 | | Leotiomycetes | 1.7102831094592 |
| 76 | Verghello vs Bavella | OTU_1232 | | Leotiomycetes | 3.33173106215749 |
| 77 | Verghello vs Bavella | OTU_457 | fungal sp TRN213 | unidentified | 25.0251517044582 |
| 78 | Verghello vs Bavella | OTU_1565 | Leotiomycetes sp BLD3 | Leotiomycetes | 3.99810061312571 |
| 79 | Verghello vs Bavella | OTU_1286 | | Leotiomycetes | 3.44266738594465 |
| 80 | Verghello vs Bavella | OTU_595 | | Leotiomycetes | 2.44277599115821 |
| 81 | Verghello vs Bavella | OTU_679 | | Leotiomycetes | 4.53653641743523 |
| 82 | Verghello vs Bavella | OTU_734 | | Dothideomycetes | 2.5195236722982 |
| 83 | Verghello vs Bavella | OTU_1292 | | Dothideomycetes | 3.02284664254276 |
| 84 | Verghello vs Bavella | OTU_669 | | Dothideomycetes | 3.45549304491711 |
| 85 | Verghello vs Bavella | OTU_640 | | Leotiomycetes | 3.93876423355081 |
| 86 | Verghello vs Bavella | OTU_575 | | | 4.63323931848402 |
| 87 | Verghello vs Bavella | OTU_997 | | Eurotiomycetes | 3.76739950320754 |
| 88 | Verghello vs Bavella | OTU_1013 | | Dothideomycetes | 2.76149108480107 |
| 89 | Verghello vs Bavella | OTU_692 | Leotiomycetes sp BLD3 | Leotiomycetes | 3.15138780696417 |
| 90 | Verghello vs Bavella | OTU_930 | | | 2.70799737569365 |
| 91 | Verghello vs Bavella | OTU_714 | Leotiomycetes sp BLD3 | Leotiomycetes | 3.78172933086983 |
| 92 | Verghello vs Bavella | OTU_25 | Phaeotheca sp | Dothideomycetes | 3.97275762815806 |
| 93 | Verghello vs Bavella | OTU_1018 | | Leotiomycetes | 2.6965257401675 |
| 94 | Verghello vs Bavella | OTU_831 | | Leotiomycetes | 4.10403849059408 |
| 95 | Verghello vs Bavella | OTU_1064 | | Leotiomycetes | 3.46003270070723 |
| 96 | Verghello vs Bavella | OTU_267 | | | 4.1001095082275 |
| 97 | Verghello vs Bavella | OTU_1180 | | Leotiomycetes | 3.15223347765059 |
| 98 | Verghello vs Bavella | OTU_616 | | Leotiomycetes | 3.5551976152887 |
| 99 | Verghello vs Bavella | OTU_1138 | | | 3.61854236763842 |
| 100 | Verghello vs Bavella | OTU_1417 | | Leotiomycetes | 4.61348516916332 |
| 101 | Verghello vs Bavella | OTU_64 | | | 5.01883222005108 |
| 102 | Verghello vs Bavella | OTU_34 | | | 5.25930283096 |
| 103 | Verghello vs Bavella | OTU_1358 | Leotiomycetes sp BLD3 | Leotiomycetes | 5.27490409917799 |
| 104 | Verghello vs Bavella | OTU_862 | | | 3.15087972516567 |
| 105 | Verghello vs Bavella | OTU_1104 | | Leotiomycetes | 3.88500915389483 |
| 106 | Verghello vs Bavella | OTU_1495 | Leotiomycetes sp BLD3 | Leotiomycetes | 4.6302480280799 |
| 107 | Verghello vs Bavella | OTU_872 | Leotiomycetes sp BLD3 | Leotiomycetes | 3.56071879137086 |

| | Comparison | Order | Class | log2FoldChange (negative = more on second levels) |
|---|---|---|---|---|
| 1 | Verghello vs Asco | Xylariales | Sordariomycetes | 5.02800865384018 |
| 2 | Verghello vs Bavella | Incertae sedis | Leotiomycetes | -1.36534813356438 |
| 3 | Verghello vs Bavella | unidentified | unidentified | 1.5926930618187 |
| 4 | Asco vs Bavella | Botryosphaeriales | Dothideomycetes | 7.33720699516155 |
| 5 | Asco vs Bavella | Eurotiales | Eurotiomycetes | 1.82573680514222 |
| 6 | Asco vs Bavella | Incertae sedis | Leotiomycetes | -1.69236022105782 |
| 7 | Asco vs Bavella | unidentified | unidentified | 1.47100012610789 |
| 8 | Asco vs Bavella | Xylariales | Sordariomycetes | -4.76720083284185 |

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

| | Nb.of.OTUs | Nb.of.samples | Nb.of.sequences |
|---|---|---|---|
| No filter | 1667 | 80 | 8335341.00 |
| Nb of sequences by sample >= 20000 | 1650 | 72 | 8313238.00 |
| Nb of sample by OTUs >= 1 | 1650 | 72 | 8313238.00 |
| Nb of sequences by OTUs >= 5 | 1302 | 72 | 8312594.00 |

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

# List of Figures



# List of Tables