Appendix S9: results after Qiime Open ref clustering

Adrien Taudiere*

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

July 19, 2016

To set the filter parameter, see directly section 'Choice of filter parameters' 2.1.
Don't forgot to set working directory.

*adrien.taudiere@cefe.cnrs.fr

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 analyse here was computed using Qiime Open ref clustering (see article for more details).

1.1 R requirements

First we 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")
```

```
#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 3.3.1 (2016-06-21) on Windows the 2016-07-19 18:50:13. See below for more information.

```
sessionInfo()

## R version 3.3.1 (2016-06-21)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 8.1 x64 (build 9600)
##
## locale:
## [1] LC_COLLATE=French_France.1252  LC_CTYPE=French_France.1252
## [3] LC_MONETARY=French_France.1252 LC_NUMERIC=C
## [5] LC_TIME=French_France.1252
##
## attached base packages:
##  [1] parallel  stats4    grid      stats     graphics  grDevices utils
```

```
##  [8] datasets  methods   base
##
## other attached packages:
##  [1] vegan_2.4-0                lattice_0.20-33
##  [3] permute_0.9-0              gridExtra_2.2.1
##  [5] venneuler_1.1-0            rJava_0.9-8
##  [7] d3treeR_0.1                data.tree_0.3.5
##  [9] treemap_2.4-1              networkD3_0.2.11
## [11] multtest_2.28.0            adegenet_2.0.1
## [13] ade4_1.7-4                 ips_0.0-7
## [15] XML_3.98-1.4               colorspace_1.2-6
## [17] DECIPHER_2.0.2             RSQLite_1.0.0
## [19] DBI_0.4-1                  Biostrings_2.40.2
## [21] XVector_0.12.0             phangorn_2.0.4
## [23] ape_3.5                    edgeR_3.14.0
## [25] limma_3.28.12              mvabund_3.11.9
## [27] DESeq2_1.12.3              SummarizedExperiment_1.2.3
## [29] Biobase_2.32.0             GenomicRanges_1.24.2
## [31] GenomeInfoDb_1.8.2         IRanges_2.6.1
## [33] S4Vectors_0.10.1           BiocGenerics_0.18.0
## [35] schoRsch_1.2               xtable_1.8-2
## [37] circlize_0.3.7             VennDiagram_1.6.17
## [39] futile.logger_1.4.1        plyr_1.8.4
## [41] cluster_2.0.4              phyloseq_1.16.2
## [43] ggplot2_2.1.0              knitr_1.13
##
## loaded via a namespace (and not attached):
##  [1] seqinr_3.1-5         deldir_0.1-12        GlobalOptions_0.0.10
##  [4] rstudioapi_0.6       AnnotationDbi_1.34.3 codetools_0.2-14
##  [7] splines_3.3.1        geneplotter_1.50.0   Formula_1.2-1
## [10] jsonlite_0.9.22      gridBase_0.4-7       annotate_1.50.0
## [13] shiny_0.13.2         DiagrammeR_0.8.2     assertthat_0.1
## [16] Matrix_1.2-6         formatR_1.4          visNetwork_1.0.1
## [19] acepack_1.3-3.3      htmltools_0.3.5      tools_3.3.1
## [22] igraph_1.0.1         coda_0.18-1          gtable_0.2.0
## [25] reshape2_1.4.1       dplyr_0.5.0          gmodels_2.16.2
## [28] fastmatch_1.0-4      Rcpp_0.12.5          RJSONIO_1.3-0
## [31] spdep_0.6-5          gdata_2.17.0         nlme_3.1-128
## [34] iterators_1.0.8      stringr_1.0.0        mime_0.4
## [37] gtools_3.5.0         statmod_1.4.24       LearnBayes_2.15
## [40] zlibbioc_1.18.0      MASS_7.3-45          scales_0.4.0
## [43] biomformat_0.99.4    rhdf5_2.16.0         lambda.r_1.1.7
## [46] RColorBrewer_1.1-2   rpart_4.1-10         latticeExtra_0.6-28
## [49] stringi_1.1.1        highr_0.6            genefilter_1.54.2
## [52] gridSVG_1.5-0        foreach_1.4.3        boot_1.3-18
## [55] BiocParallel_1.6.2   shape_1.4.2          chron_2.3-47
## [58] evaluate_0.9         htmlwidgets_0.6      magrittr_1.5
## [61] R6_2.1.2             nnls_1.4             Hmisc_3.17-4
## [64] foreign_0.8-66       mgcv_1.8-12          survival_2.39-5
## [67] sp_1.2-3             nnet_7.3-12          tibble_1.0
## [70] futile.options_1.0.0 locfit_1.5-9.1       data.table_1.9.6
```

```
## [73] digest_0.6.9         httpuv_1.3.3         munsell_0.4.3
## [76] tweedie_2.2.1        quadprog_1.5-5
```

1.3 Some usefull functions

The function `as.binaryOtuTable` convert a phyloseq object into a phyloseq object with binary (*i.e.* 0/1) OTU table. It allow to suppress effect due to 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` allow to plot accumulation curves in fonction of a factor in samples data (`@sam_data` of phyloseq object).

`otu_circle` use 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, convert 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 (either using package DESeq2 or edgeR).

```
setwd("~/Documents/GitHub/FEF_paper/")
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(3, 5, 7, 9, 11, 13, 15)]
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=""))
funguild <- funguild[!is.na(match(funguild$OTU_ID, rownames(taxRDP2))),]

match_interm <- match(funguild$OTU_ID, 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...

names(repset) <- unlist(strsplit(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 Caracteristics 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 12 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)
```

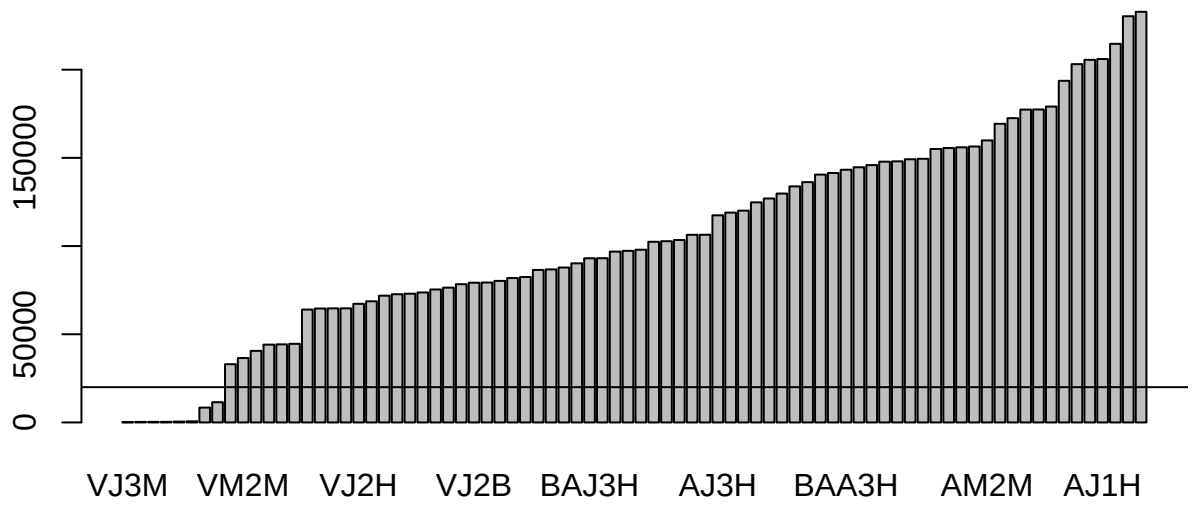

Figure 2.1: Number of sequences by sample

2.4 Filter OTUs by number of samples

First, we can visualize the number of OTU present 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))

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%).

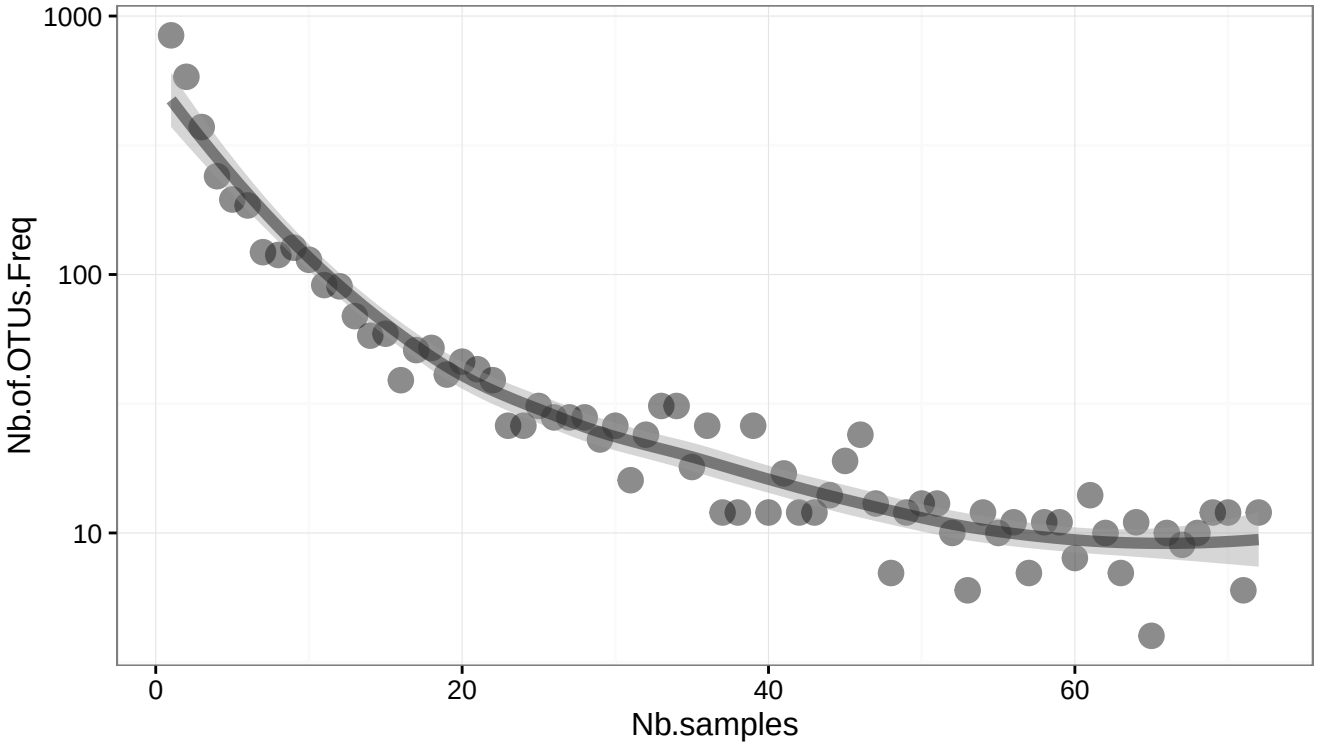

Figure 2.2: Number of OTU present in a given number of samples. Vertical bar illustrate the filtering parameter.

2.5 Filter OTUs by number of sequences

First, 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  773800
```

```
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)
```

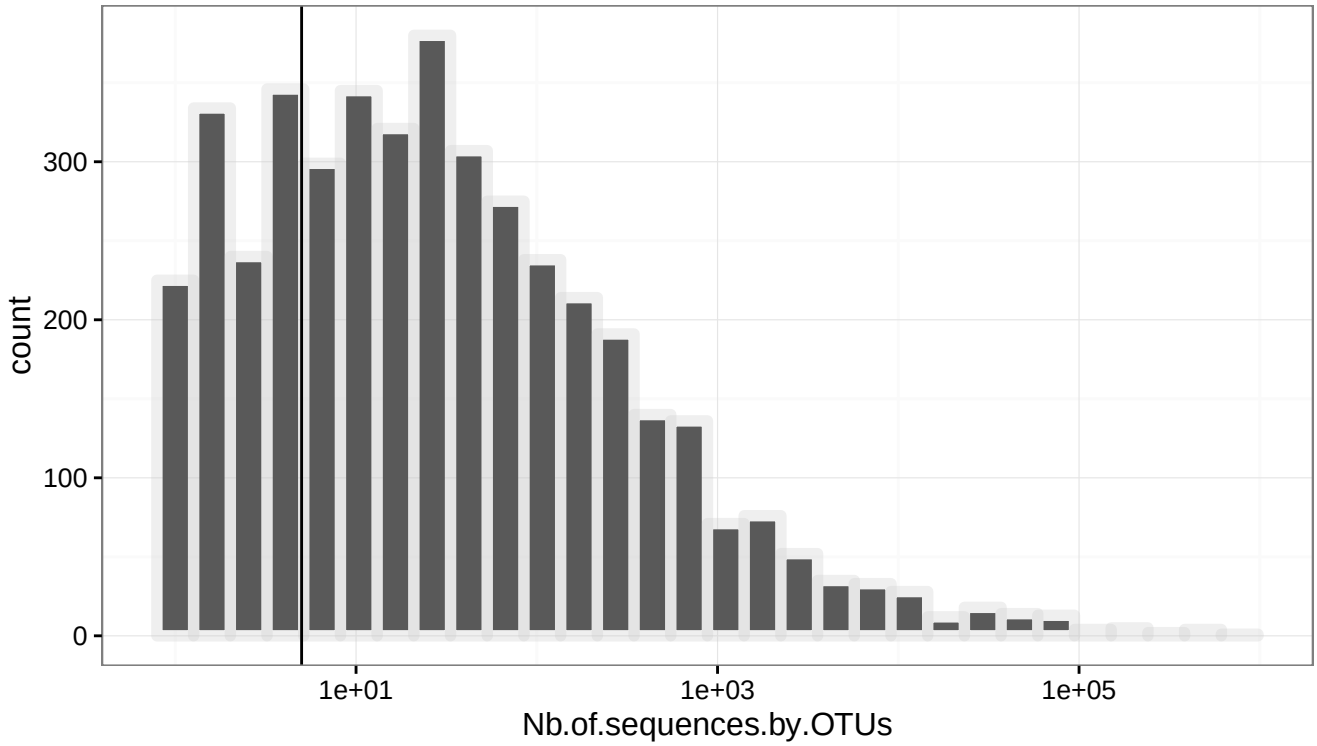

Figure 2.3: Number of sequences by OTU (log10 transformed). Horizontal bar illustrate the filtering parameter.

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') +
```

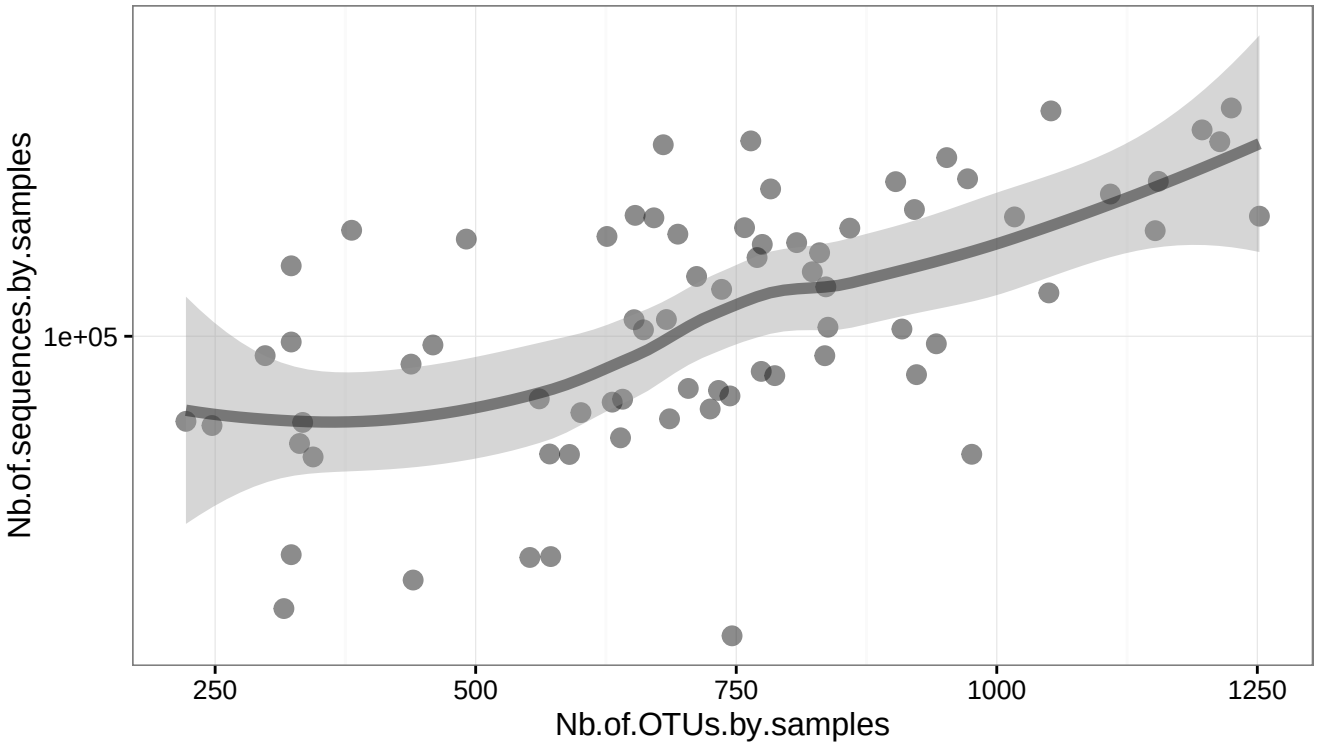

Figure 3.1: Number of OTUs by samples in fonction the number of sequences by samples (log10 axe). The tendency is represented by the line obtain from loess (Local Polynomial Regression Fitting).

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

```
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"))
```

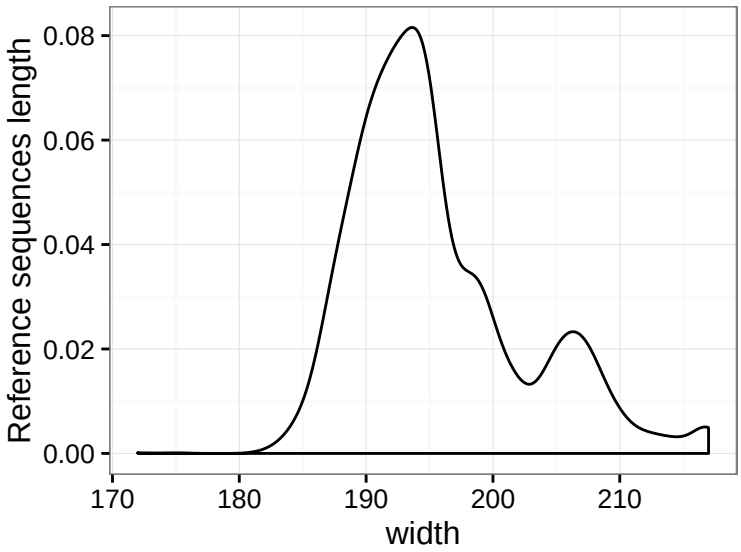

Figure 3.2: Distribution of reference sequences length.

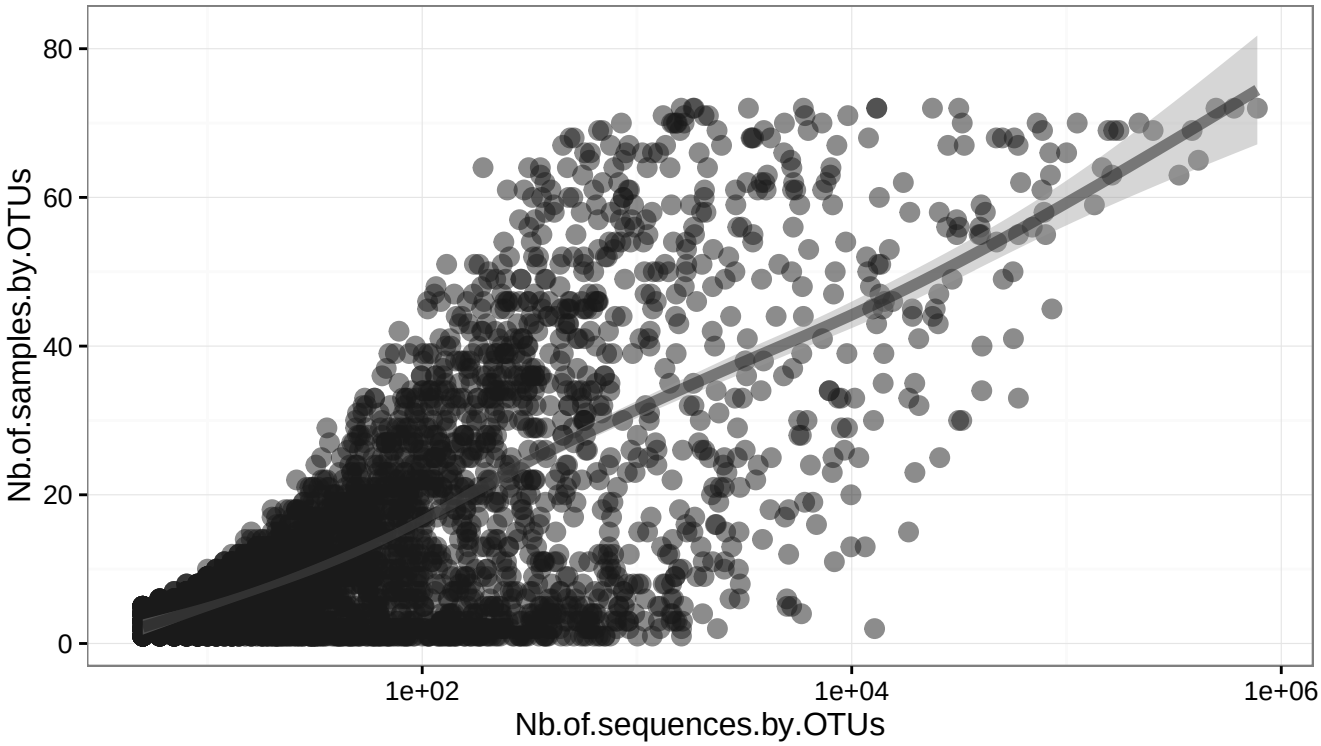

Figure 3.3: Number of sequences by OTUs (log10 axe) in fonction the number of samples where OTUs were found. The tendency is represented by the line obtain from gam (Generalized additive models with integrated smoothness estimation).

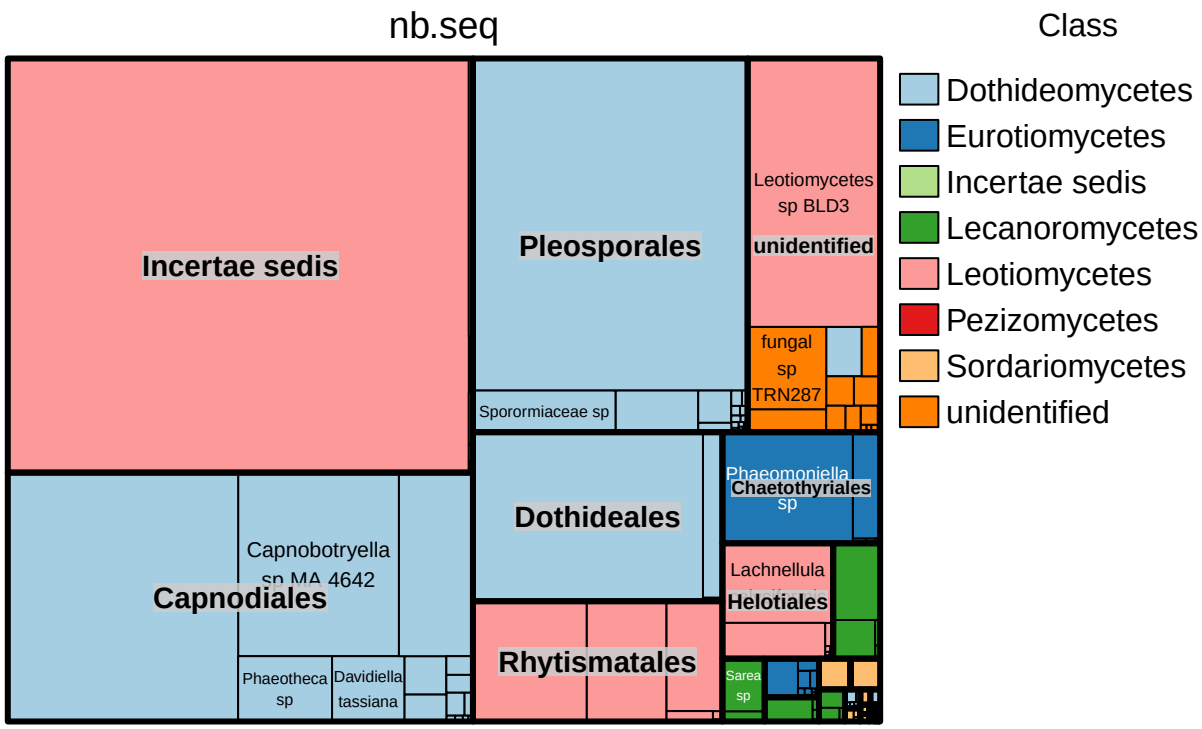

Figure 3.4: Distribution of the number of sequences in the taxonomy. Color represent Class, bold lines delimit Order and thick line delimit species.

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)
```

```
data.f3_MINSEQ1000 <- subset_taxa(data.f3, rowSums(data.f3@otu_table)>999)
sankey_phyloseq(data.f3_MINSEQ1000, tax2remove =
                  c("Incertae sedis", "unidentified", "Xylariales", "NA"),
                nbSeq = TRUE, taxa = c(1:6))
```

```
sankey_phyloseq(data.f3, tax2remove = c("Incertae sedis", "unidentified", "Xylariales"),
                nbSeq = FALSE, taxa = c(1:5), min.prop.tax = 0.01)
```

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)
```

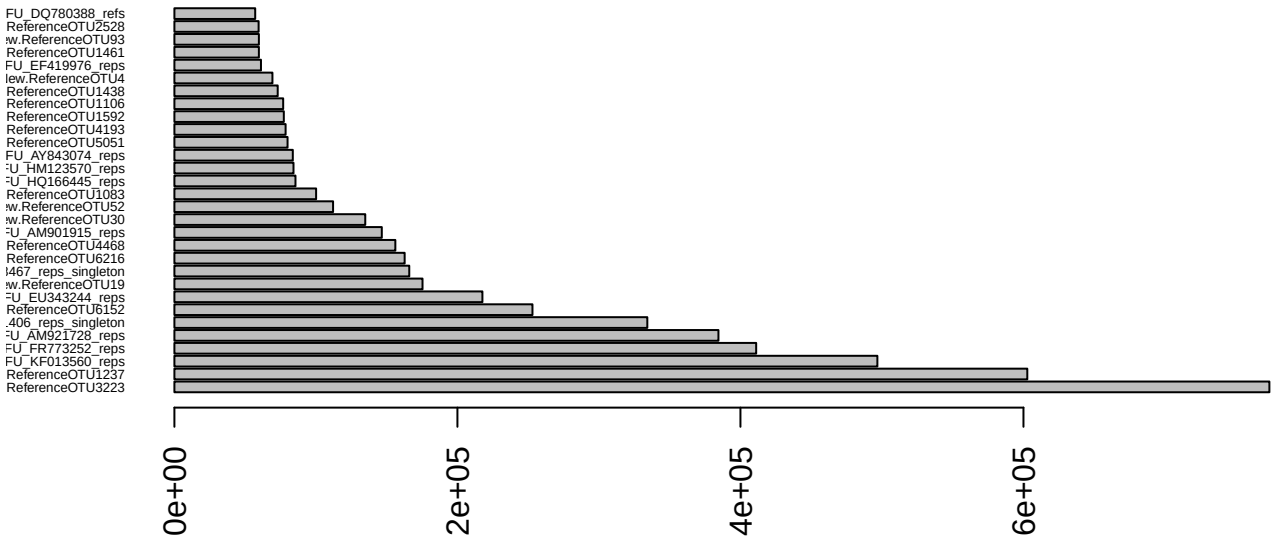

Figure 3.5: Number of sequences of the 30 more abundant OTUs (number of sequences).

```
print(xtable(df_the30mostfrequents[, c(1:8, 12)], auto = T,
             caption = "Taxonomie of the 30 more
             frequent 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	-	-	773785
Ascomycota	Leotiomycetes	Incertae sedis	Incertae sedis	Cyclaneusma	Cyclaneusma minus	-	-	602702
Ascomycota	Leotiomycetes	Incertae sedis	Incertae sedis	Cyclaneusma	Cyclaneusma minus	-	-	496770
Ascomycota	Dothideomycetes	Pleosporales				-	-	411106
Ascomycota	Dothideomycetes	Dothideales	Dothioraceae	unidentified	Dothioraceae sp	-	-	384457
Ascomycota	Dothideomycetes	Capnodiales	Incertae sedis	Capnobotryella	Capnobotryella sp MA 4642	Saprotroph	Undefined Saprotroph	334181
Ascomycota	Dothideomycetes	Pleosporales				-	-	253023
Ascomycota	Dothideomycetes	Capnodiales				-	-	217662
Ascomycota	Dothideomycetes	Pleosporales				-	-	175339
Ascomycota	Leotiomycetes	unidentified	unidentified	unidentified	Leotiomycetes sp BLD3	-	-	165973
Ascomycota	Leotiomycetes	unidentified	unidentified	unidentified	Leotiomycetes sp BLD3	-	-	162700
Ascomycota	Dothideomycetes	Pleosporales				-	-	156134
Ascomycota	Dothideomycetes	Capnodiales	Mycosphaerellaceae	Phaeothecoidea	Phaeothecoidea sp	Saprotroph	Undefined Saprotroph	146545
						-	-	134866
						-	-	112125
Ascomycota	Leotiomycetes	Rhytismatales	Rhytismataceae	Lophodermium	Lophodermium conigenum	Pathotroph	Plant Pathogen	100144
						-	-	85558
Ascomycota	Eurotiomycetes	Chaetothyriales	Herpotrichiellaceae	Phaeomoniella	Phaeomoniella sp	Saprotroph	Undefined Saprotroph	84162
Ascomycota	Dothideomycetes	Capnodiales				-	-	83661
						-	-	79960
						-	-	78582
Ascomycota	Leotiomycetes	Incertae sedis	Incertae sedis	Cyclaneusma	Cyclaneusma minus	-	-	77326
Ascomycota	Leotiomycetes	unidentified	unidentified	unidentified	Leotiomycetes sp BLD3	-	-	76859
Ascomycota	Leotiomycetes	Incertae sedis	Incertae sedis	Cyclaneusma	Cyclaneusma minus	-	-	73000
unidentified	unidentified	unidentified	unidentified	unidentified	fungal sp TRN287	-	-	69266
Ascomycota	Eurotiomycetes	Chaetothyriales	Herpotrichiellaceae	Phaeomoniella	Phaeomoniella sp	Saprotroph	Undefined Saprotroph	61167
Ascomycota	Dothideomycetes	Capnodiales				-	-	59701
						-	-	59694
Ascomycota	Leotiomycetes	Rhytismatales	Rhytismataceae	Lophodermium		Pathotroph	Plant Pathogen	59517

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

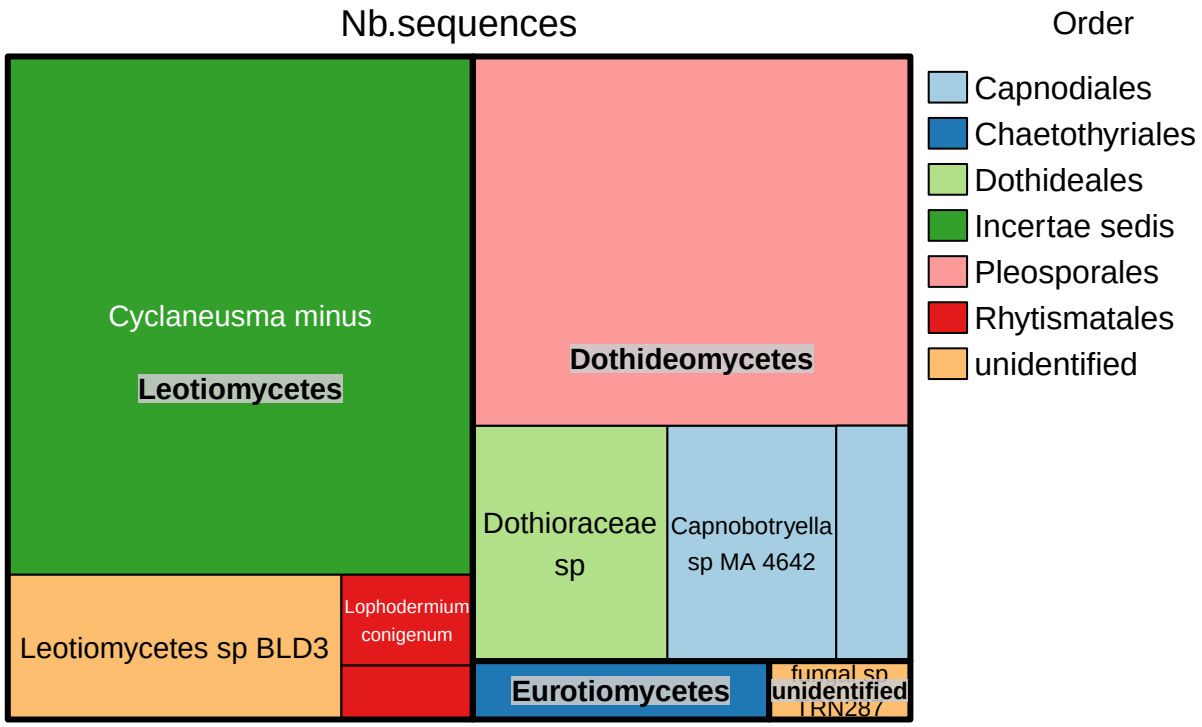

Figure 3.6: Number of sequences of the 30 most abundant OTUs (number of sequences). Colors indicate Order, bold lines delimit Class and thick lines delimit species.

```
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)
```

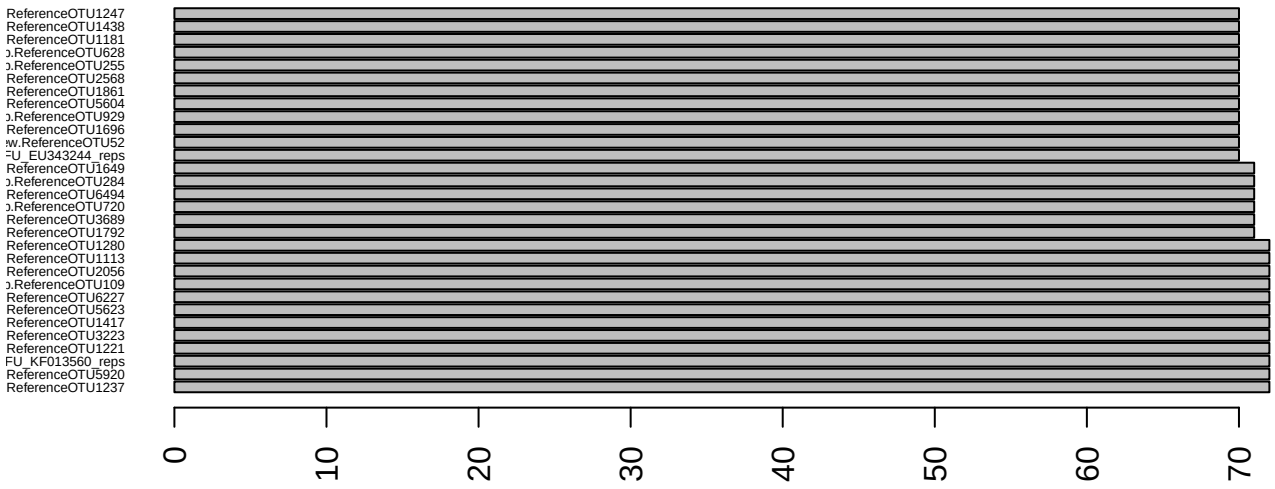

Figure 3.7: Number of samples of the 30 more frequent OTUs (number of samples).

```r
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	-	-	72
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	-	-	72
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	-	-	72
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	-	-	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
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

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

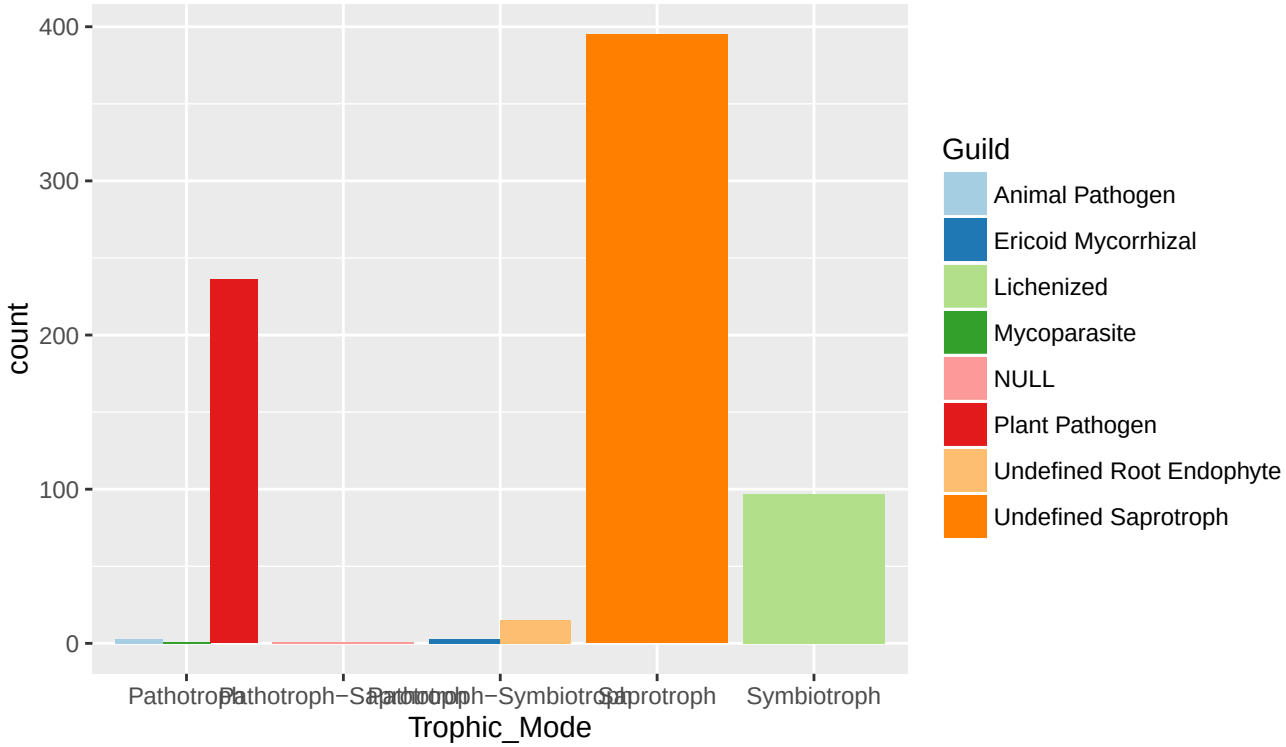

Figure 4.1: Distribution of OTUs into functional Guild.

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()
```

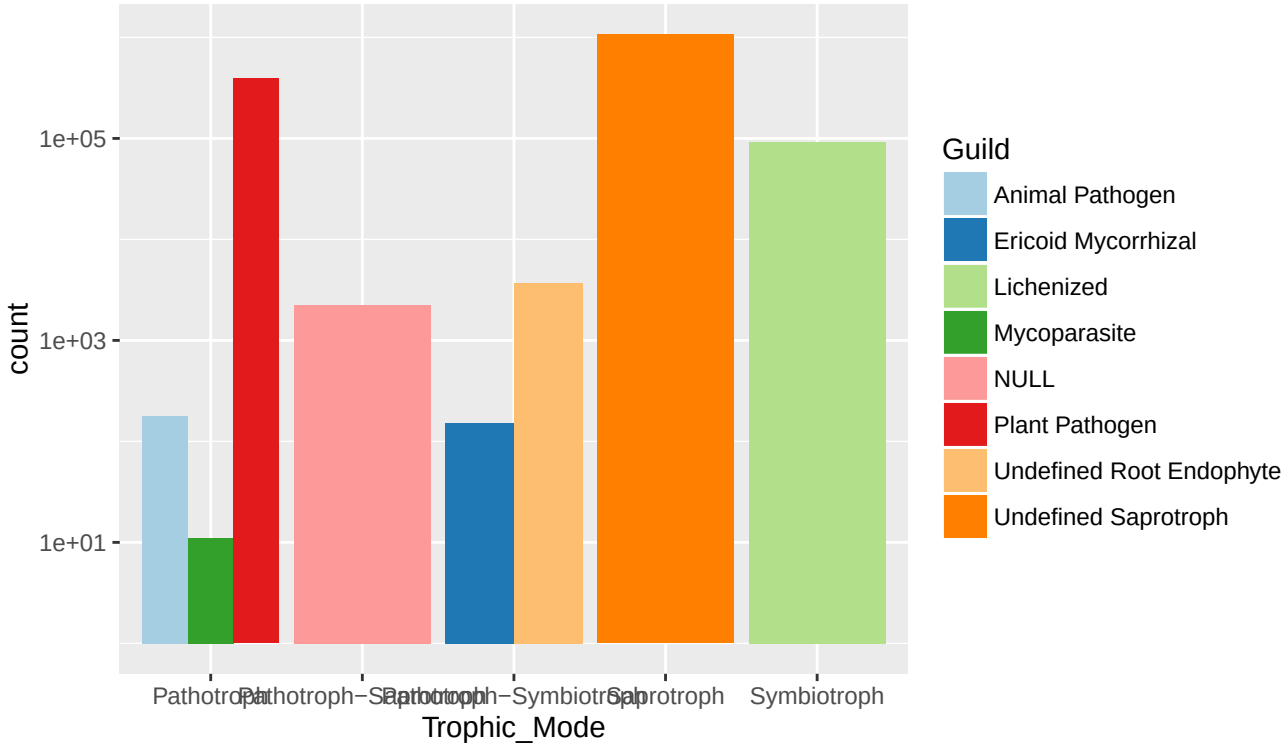

Figure 4.2: Distribution of sequences (log10 transformed) into functional Guild.

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)
```

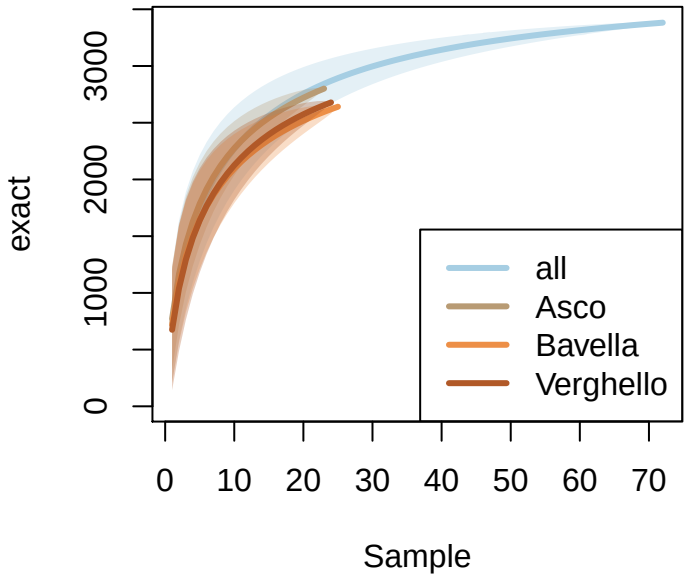

Figure 5.1: Rarefaction curves for each sites. Notes that if singletons were removed, these curves are biaised.

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)
```

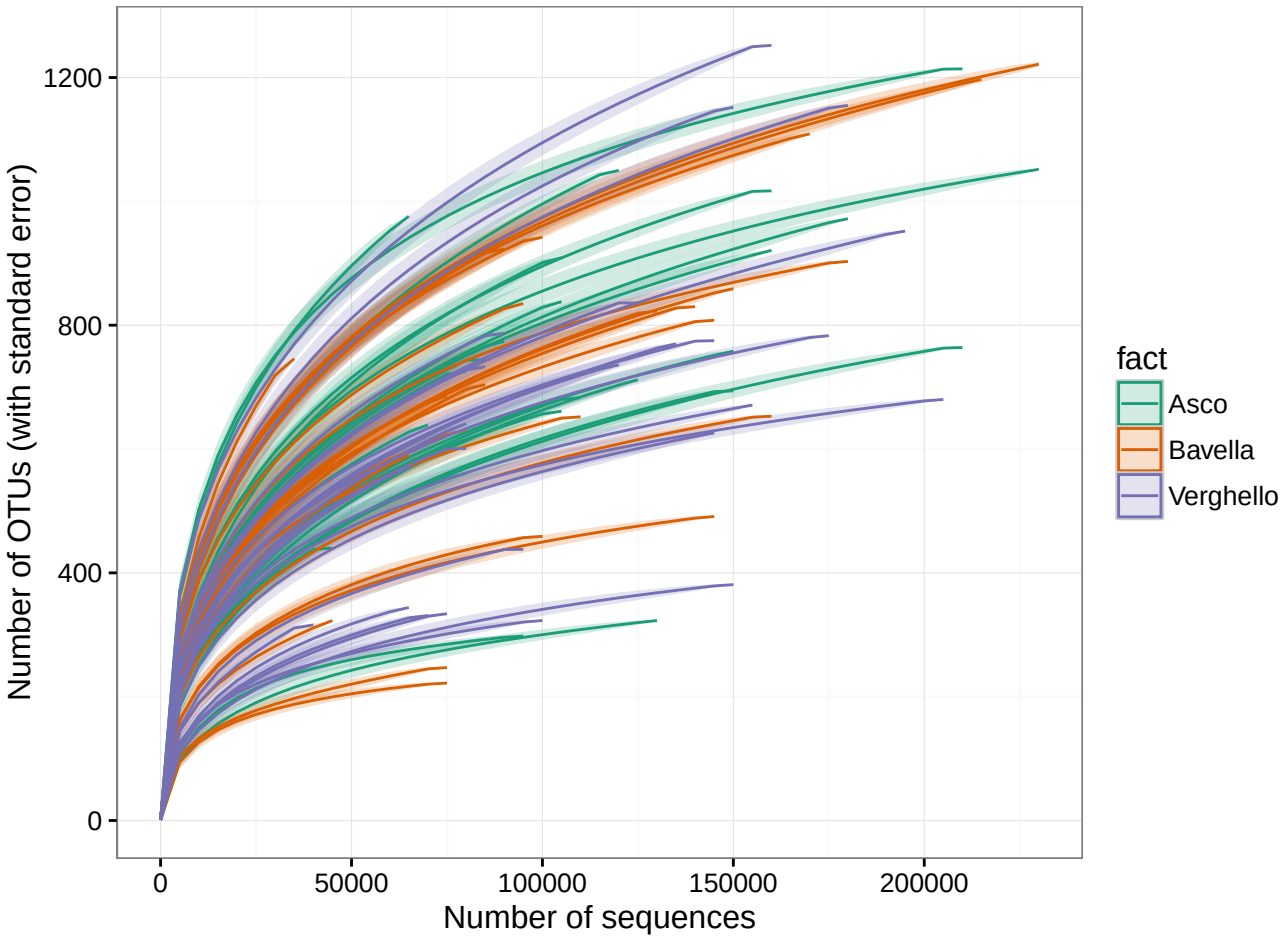

Figure 5.2: Rarefaction curves for each samples using sequences number on x-axes. Notes that if singletons were removed, these curves are biaised.

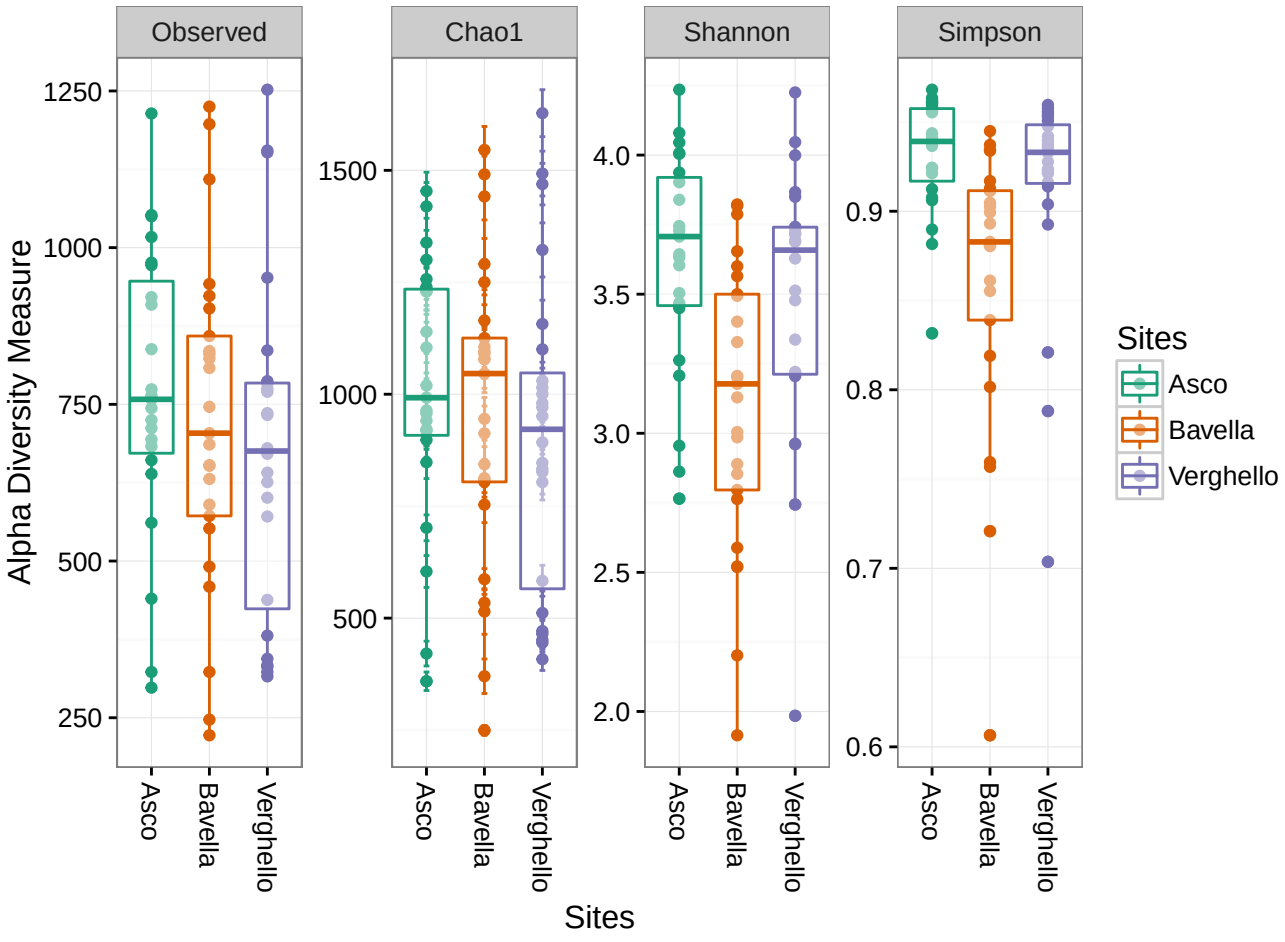

Figure 5.3: Diversity of each sites

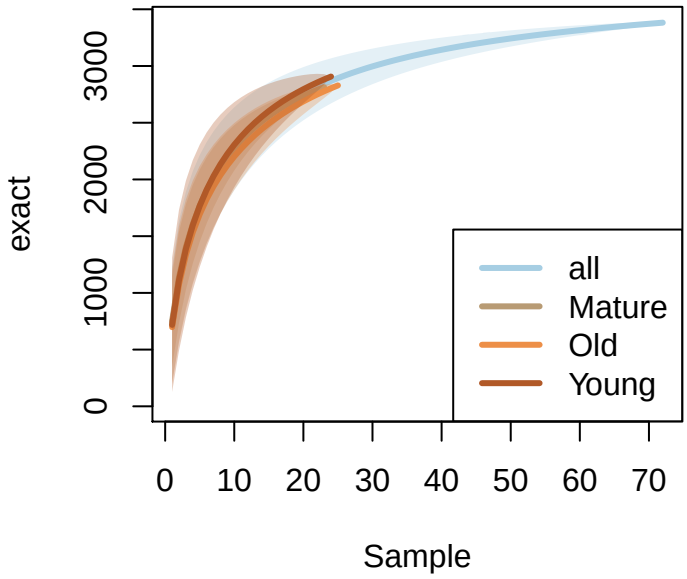

Figure 5.4: Rarefaction curves for each tree age modalities. Notes that if singletons were removed, these curves are biaised.

```
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)	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 1 (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))
```

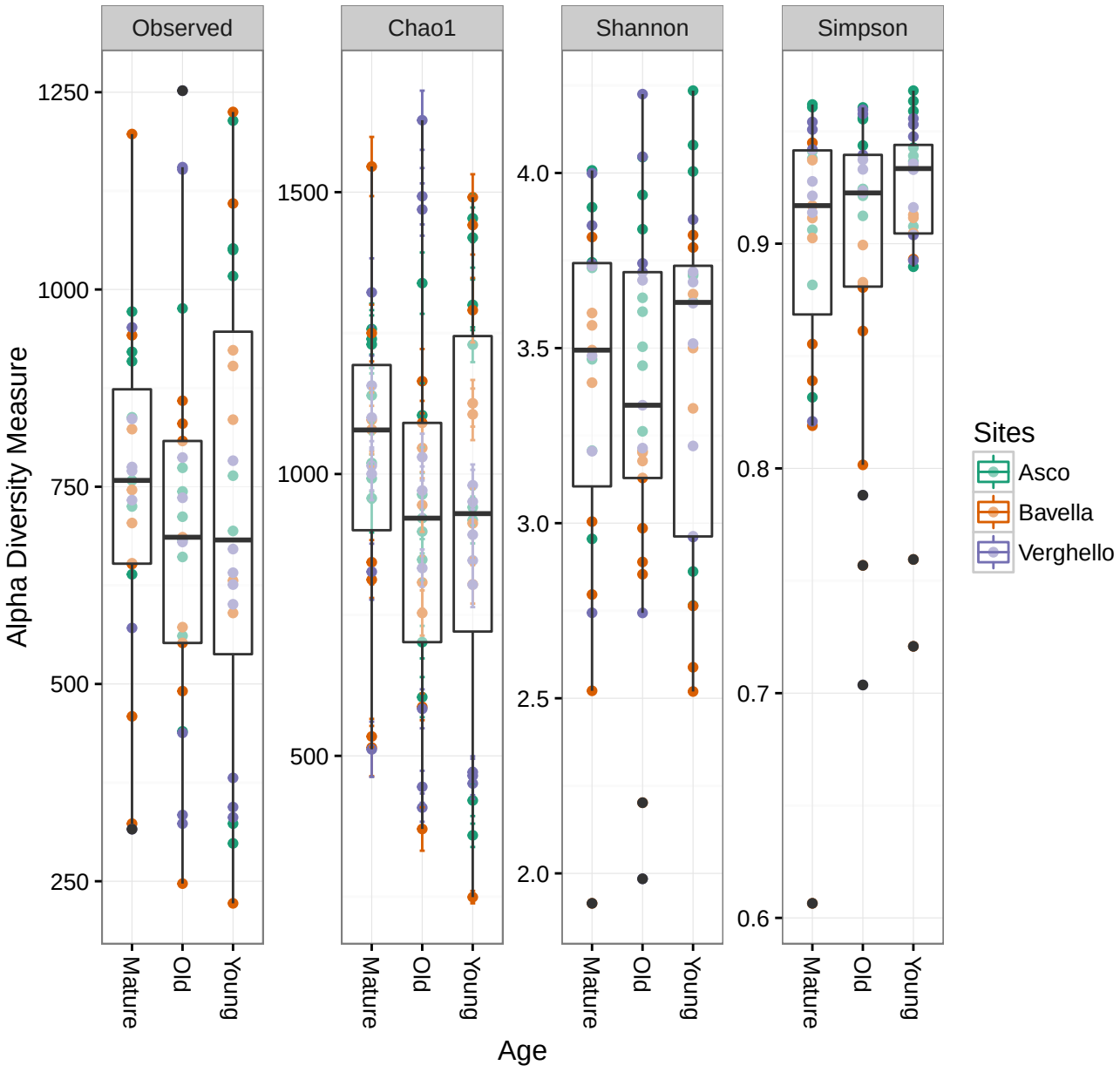

Figure 5.5: Diversity in function of tree age. Color represent sites.

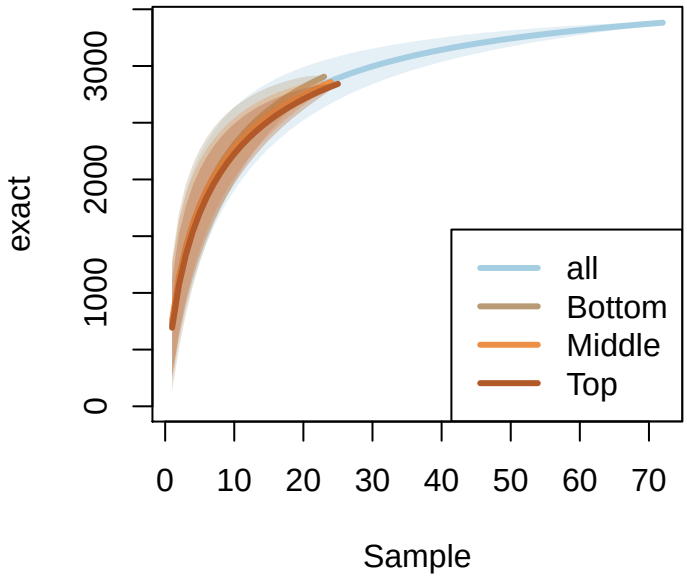

Figure 5.6: Rarefaction curves for each elevation. Notes that if singletons were removed, these curves are biaised.

```
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("")
```

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 Ordination

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

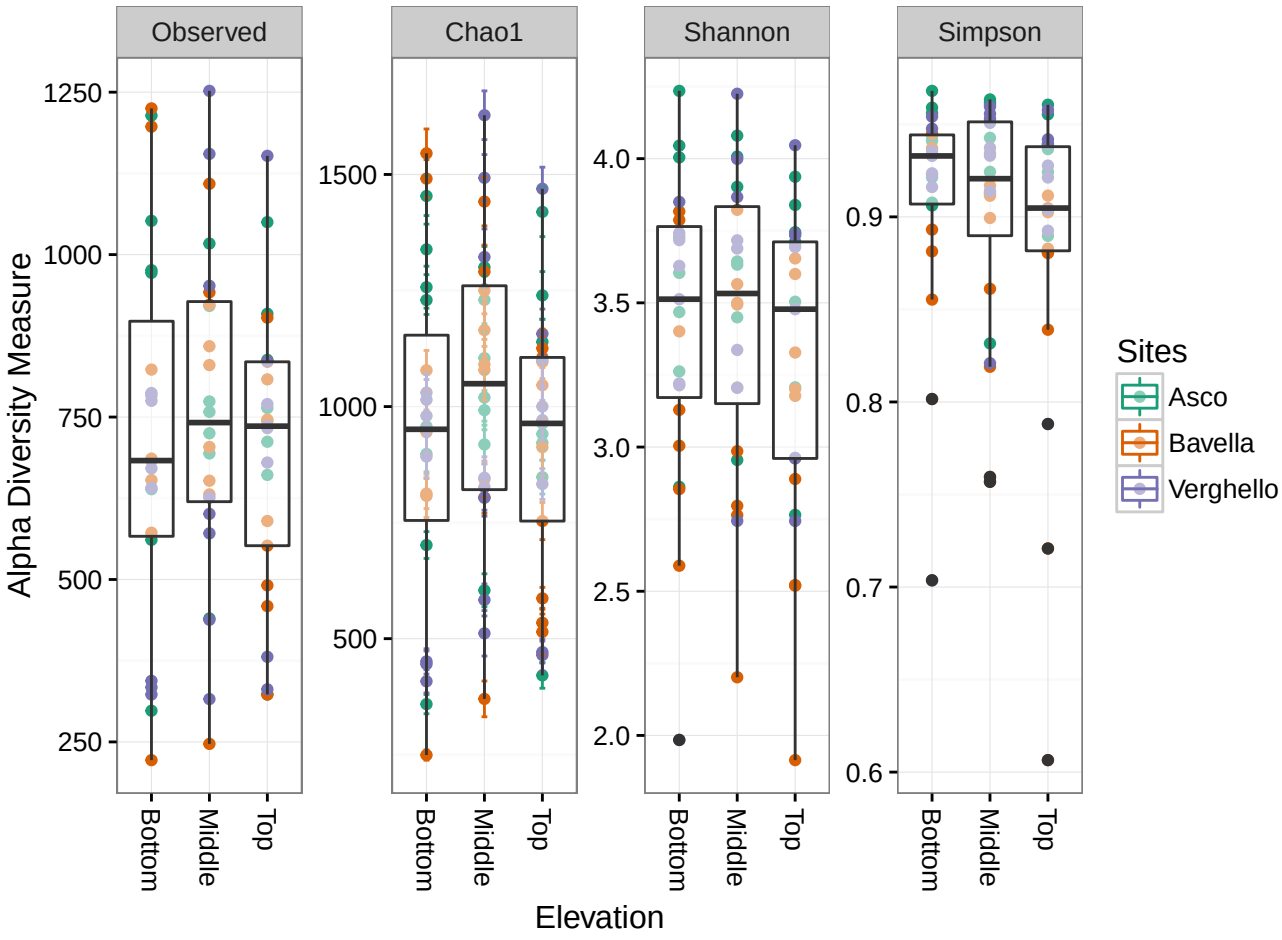

Figure 5.7: Diversity in function of elevation. Color represent sites.

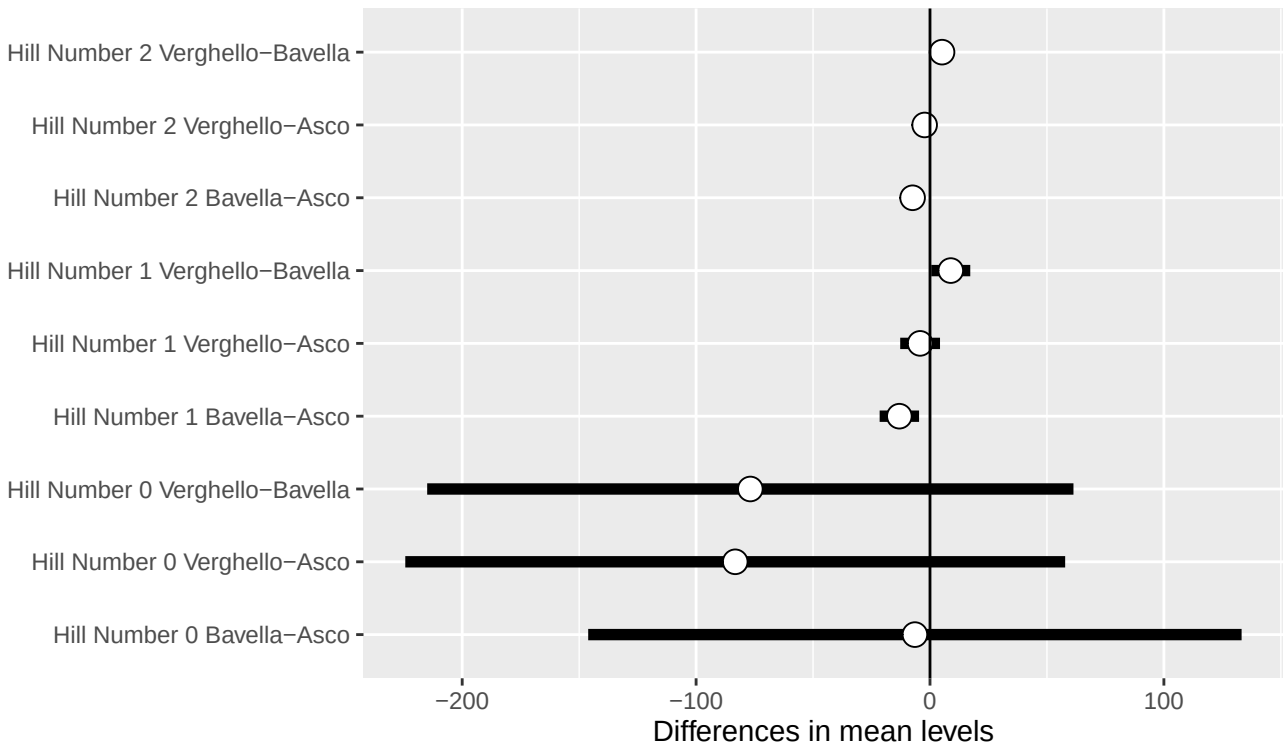

Figure 5.8: Results of the Tuckey HSD testing for differences in mean Hill numbers among pairs of modalities

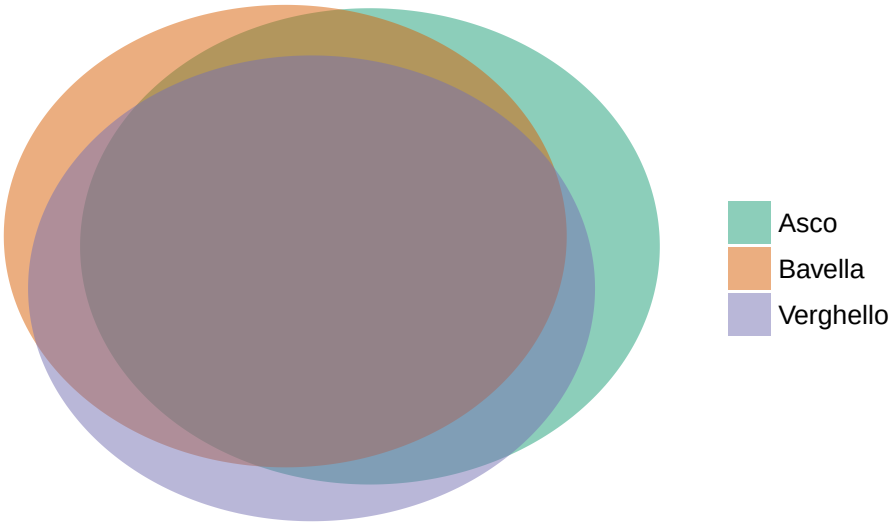

Figure 6.1: Venn diagramm ef the distribution of OTUs among Sites

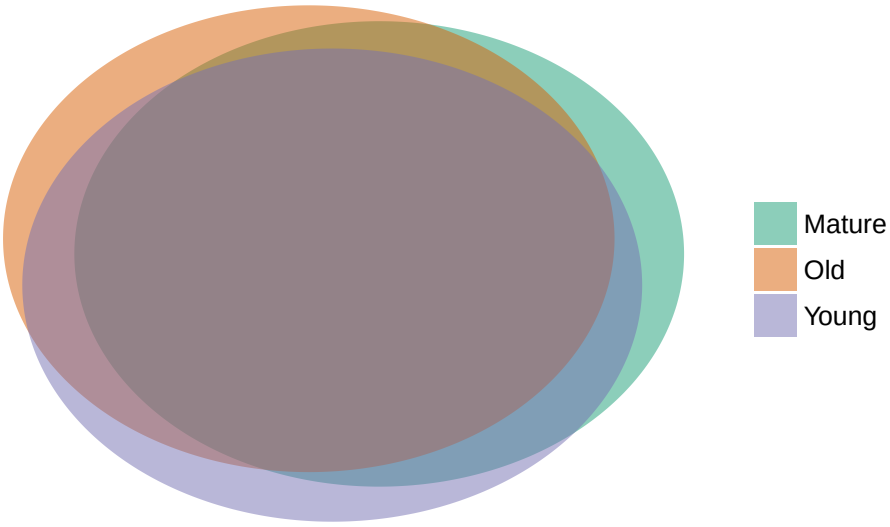

Figure 6.2: Venn diagramm ef the distribution of OTUs among host age

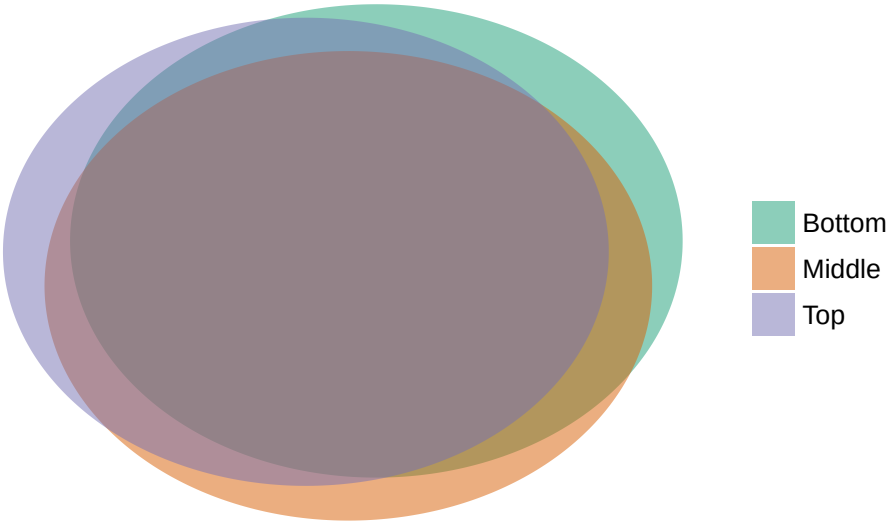

Figure 6.3: Venn diagramm ef the distribution of OTUs among elevation of samples

	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 Shannons entropy index (Hill number 2 (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 Simpsons concentration index (Hill number 3 (q = 2)

```
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 0.2076191
## Run 1 stress 0.2123561
## Run 2 stress 0.2121729
## Run 3 stress 0.2130255
## Run 4 stress 0.2083093
## Run 5 stress 0.2121839
## Run 6 stress 0.2084524
```

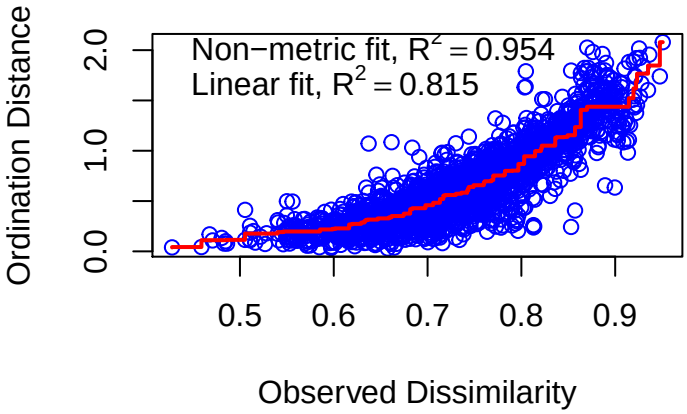

Figure 6.4: Stress plot of the NMDS

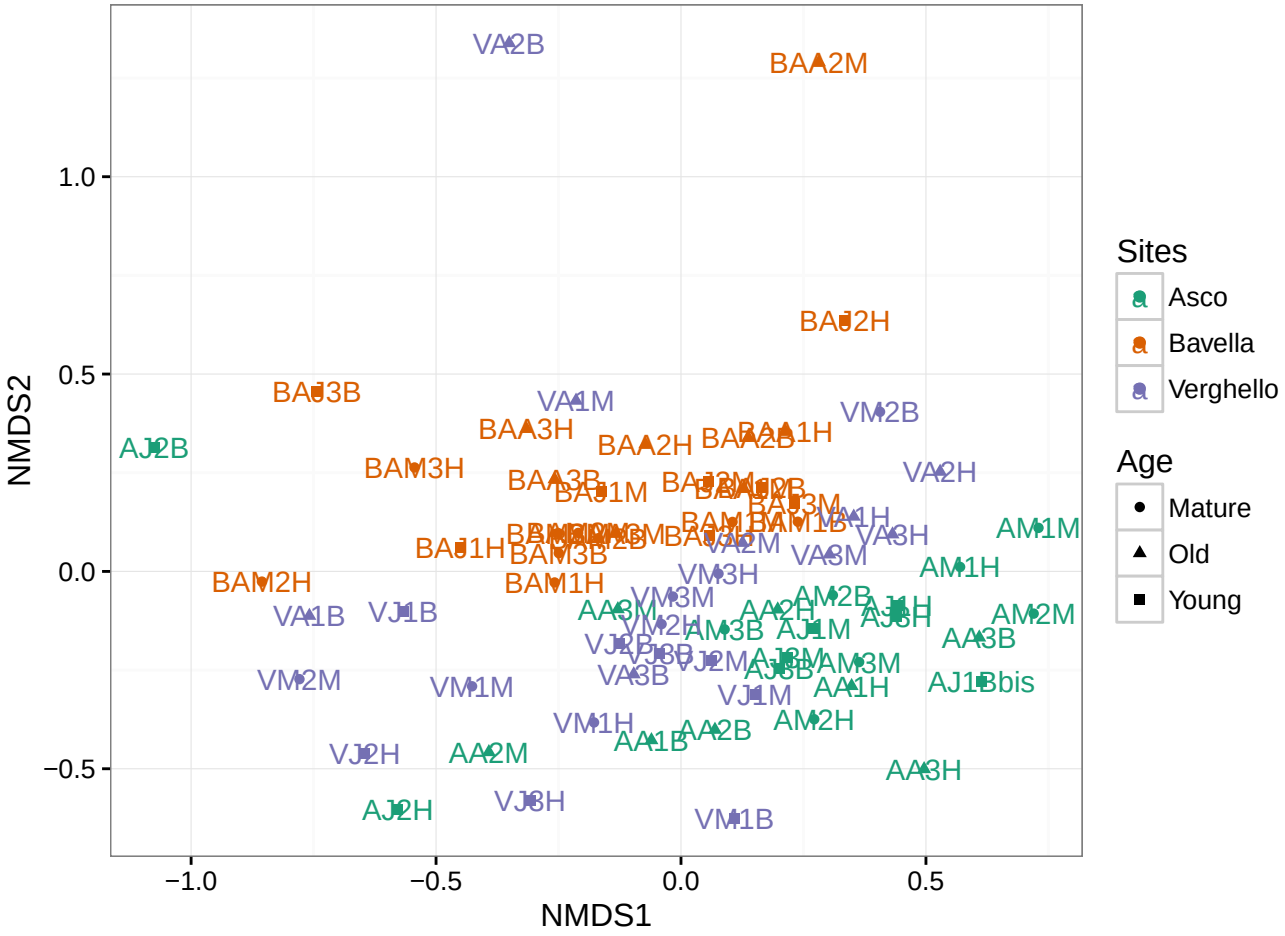

Figure 6.5: NMDS of OTU table. Colors represent sites and shape the age of tree.

```
## Run 7 stress 0.2117324
## Run 8 stress 0.2090304
## Run 9 stress 0.2117702
## Run 10 stress 0.2096642
## Run 11 stress 0.2130115
## Run 12 stress 0.2118757
## Run 13 stress 0.2116805
## Run 14 stress 0.2095071
## Run 15 stress 0.2122905
## Run 16 stress 0.2255298
## Run 17 stress 0.2117692
## Run 18 stress 0.2074187
## ... New best solution
## ... Procrustes: rmse 0.0144286  max resid 0.08090176
## Run 19 stress 0.2074905
## ... Procrustes: rmse 0.04257561  max resid 0.1669314
## Run 20 stress 0.2135174
## *** No convergence -- monoMDS stopping criteria:
##      4: no. of iterations >= maxit
##     16: 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)
```

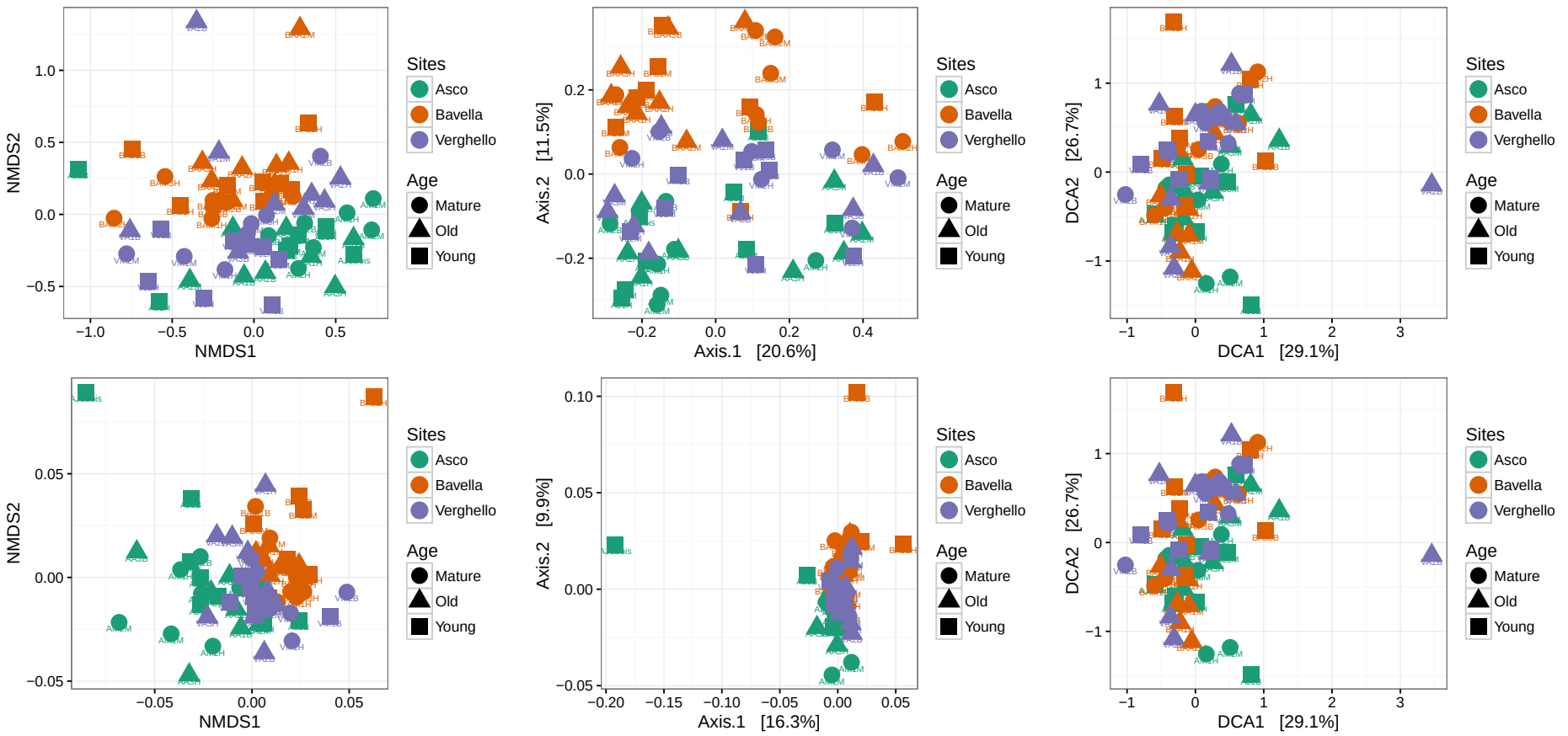

Figure 6.6: Comparison of different distances (bray (up) and gower (bottom)) and ordination methods (NMDS (left), PCoA (center) and DCA (right)).

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

6.3 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)
```

```
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	0.0374
Elevation	2	0.60	0.30	1.33	0.03	0.1125
Sites:Age	4	1.67	0.42	1.84	0.09	0.0007
Sites:Elevation	4	0.90	0.22	0.99	0.05	0.4788
Age:Elevation	4	1.15	0.29	1.27	0.06	0.0896
Sites:Age:Elevation	8	1.96	0.25	1.09	0.10	0.2564
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")
```

	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	0.0371
Elevation	2	0.59	0.29	1.37	0.03	0.1076
Sites:Age	4	1.63	0.41	1.90	0.09	0.0011
Sites:Elevation	4	0.86	0.21	1.00	0.05	0.4508
Age:Elevation	4	1.12	0.28	1.30	0.06	0.0833
Sites:Age:Elevation	8	1.87	0.23	1.09	0.10	0.2652
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

```
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)).")
```

	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	0.0080
Elevation	2	0.29	0.15	1.04	0.03	0.3762
Sites:Age	4	0.84	0.21	1.49	0.08	0.0031
Sites:Elevation	4	0.48	0.12	0.84	0.04	0.8951
Age:Elevation	4	0.58	0.14	1.02	0.05	0.4031
Sites:Age:Elevation	8	1.06	0.13	0.94	0.10	0.7012
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)).

6.4 Permanova on sites, host ages and individual trees

```
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).")
```

	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	0.0209
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	0.0004
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).

```
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")
```

	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	0.0212
Sites:Age	4	1.64	0.41	2.09	0.09	0.0001
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

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

	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	0.0013
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)).

6.5 Differences in abundances and OTUs number by Order.

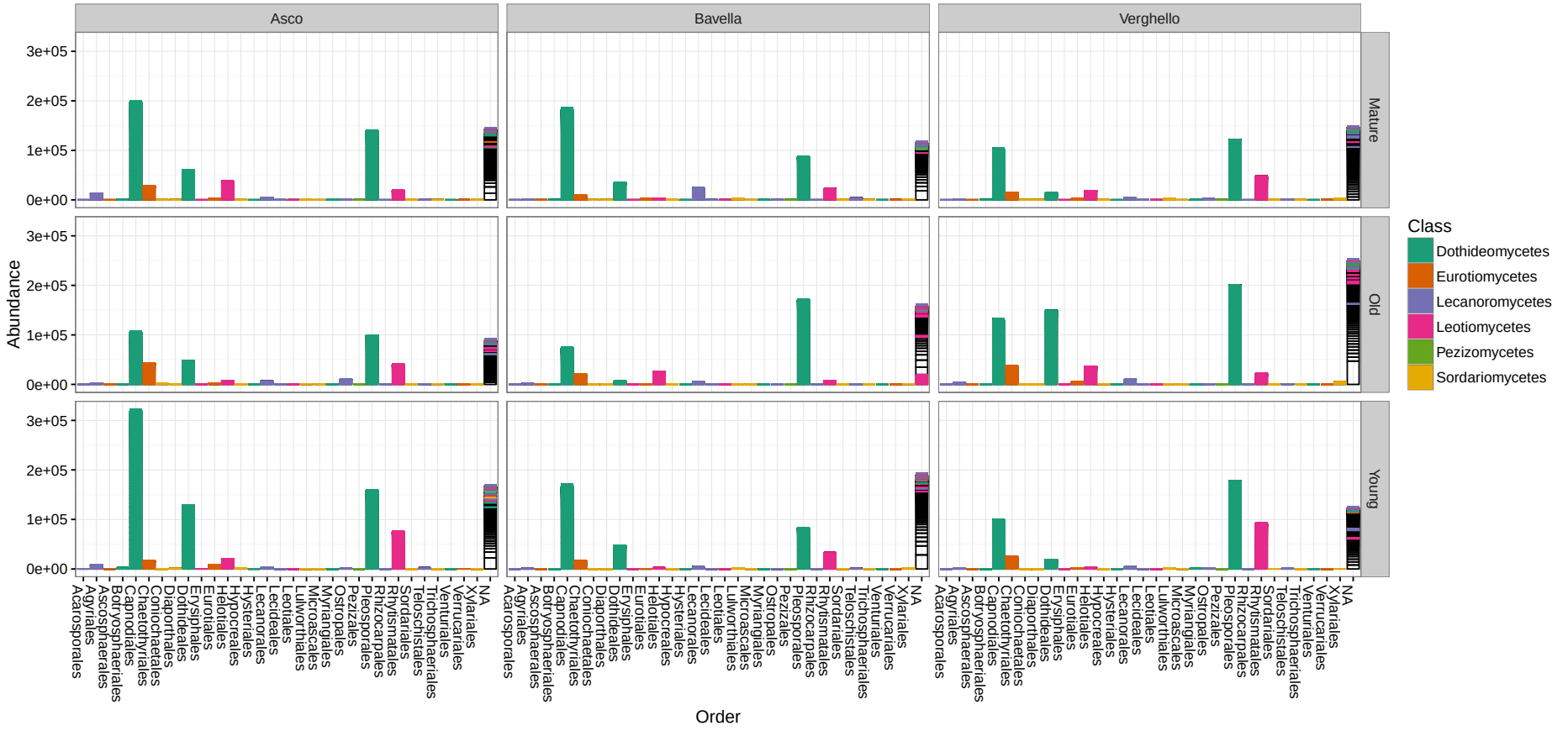

Figure 6.7: Taxonomic distribution of sequences in the different site * age combinaison.

```
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")
```

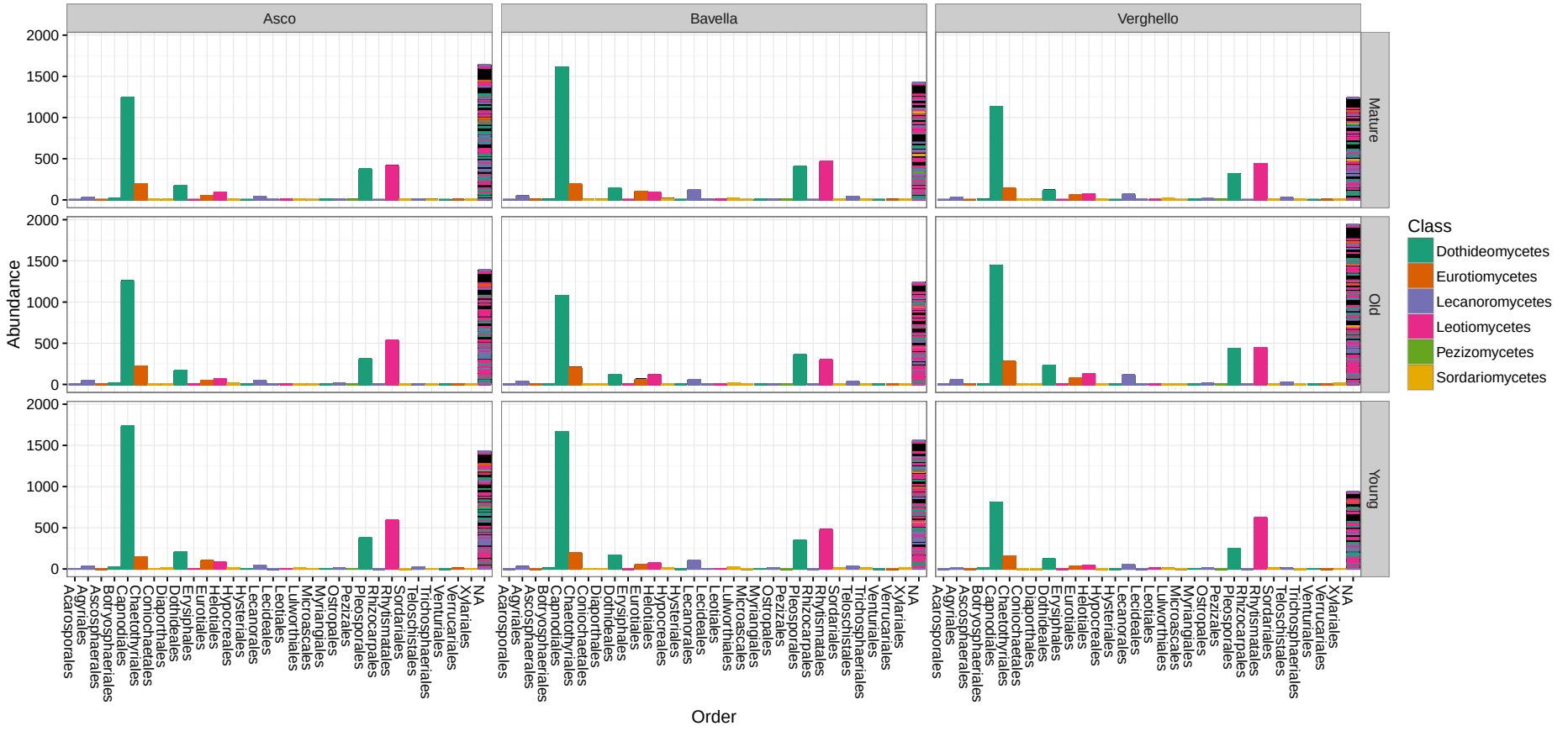

Figure 6.8: Taxonomic distribution of OTUs in the different site * age combinaison.

```
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")
```

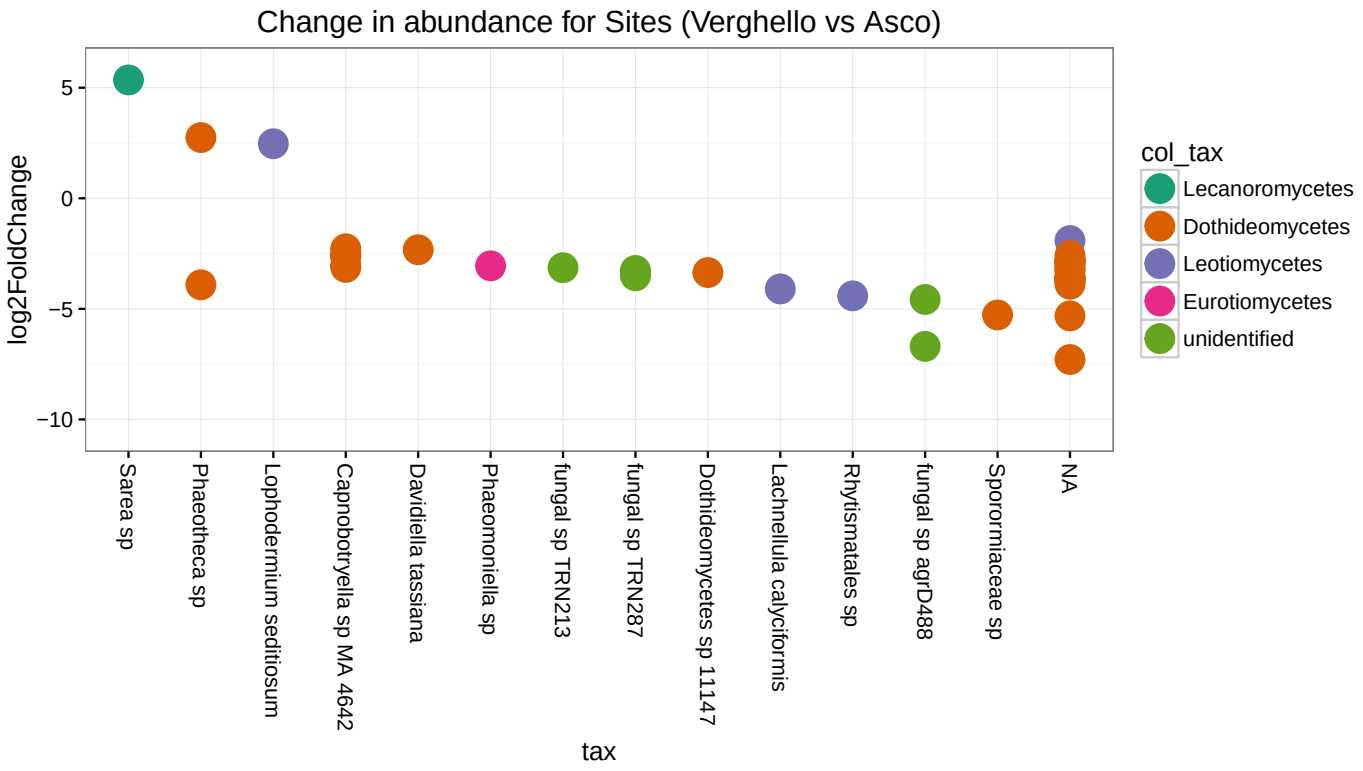

Figure 6.9: OTUs significantly different in terms of abundances between Verghello (positive values) and Asco (negative values)

6.6 Differences in abundances for each OTUs

6.6.1 Pairwise comparison of the OTUs composition by sites

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

## [1] '1.12.3'

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
```

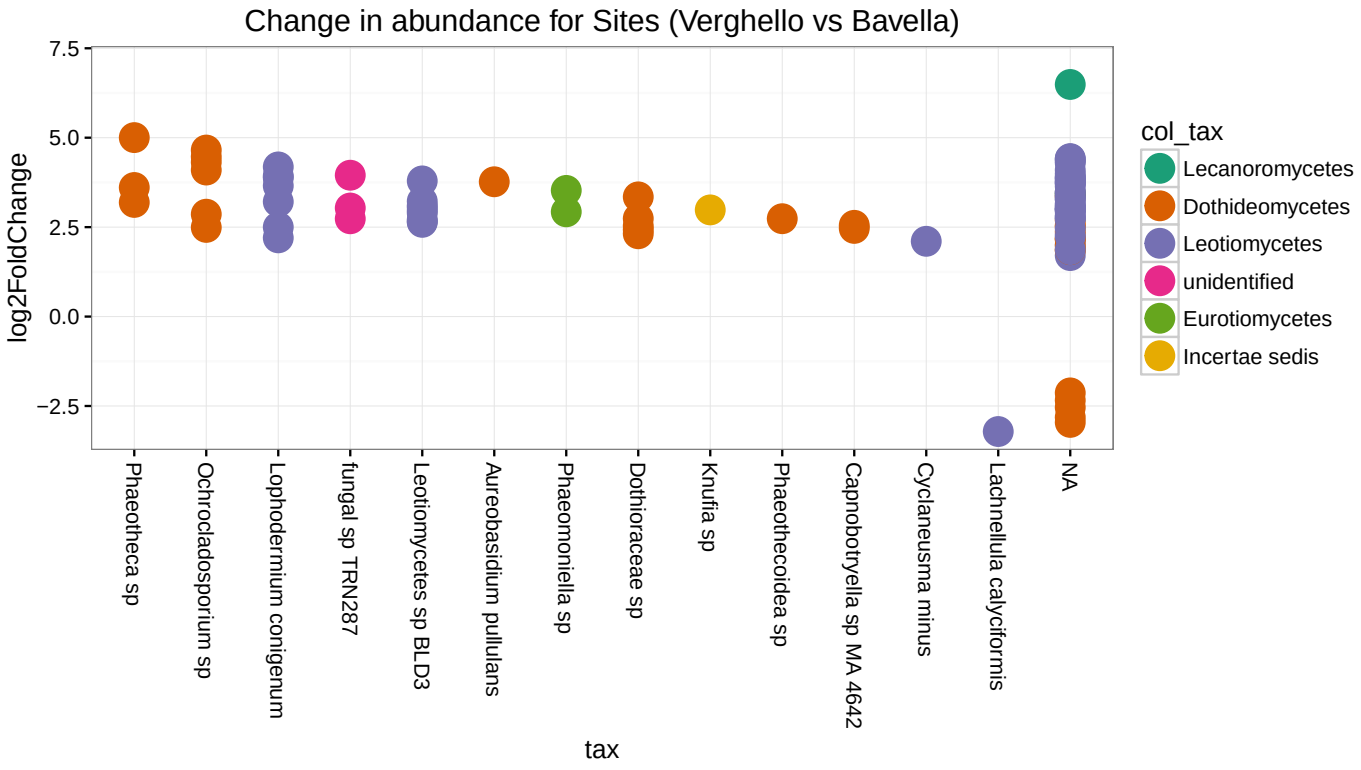

Figure 6.10: OTUs significantly different in terms of abundances between Verghello (positive values) and Bavella (negative values)

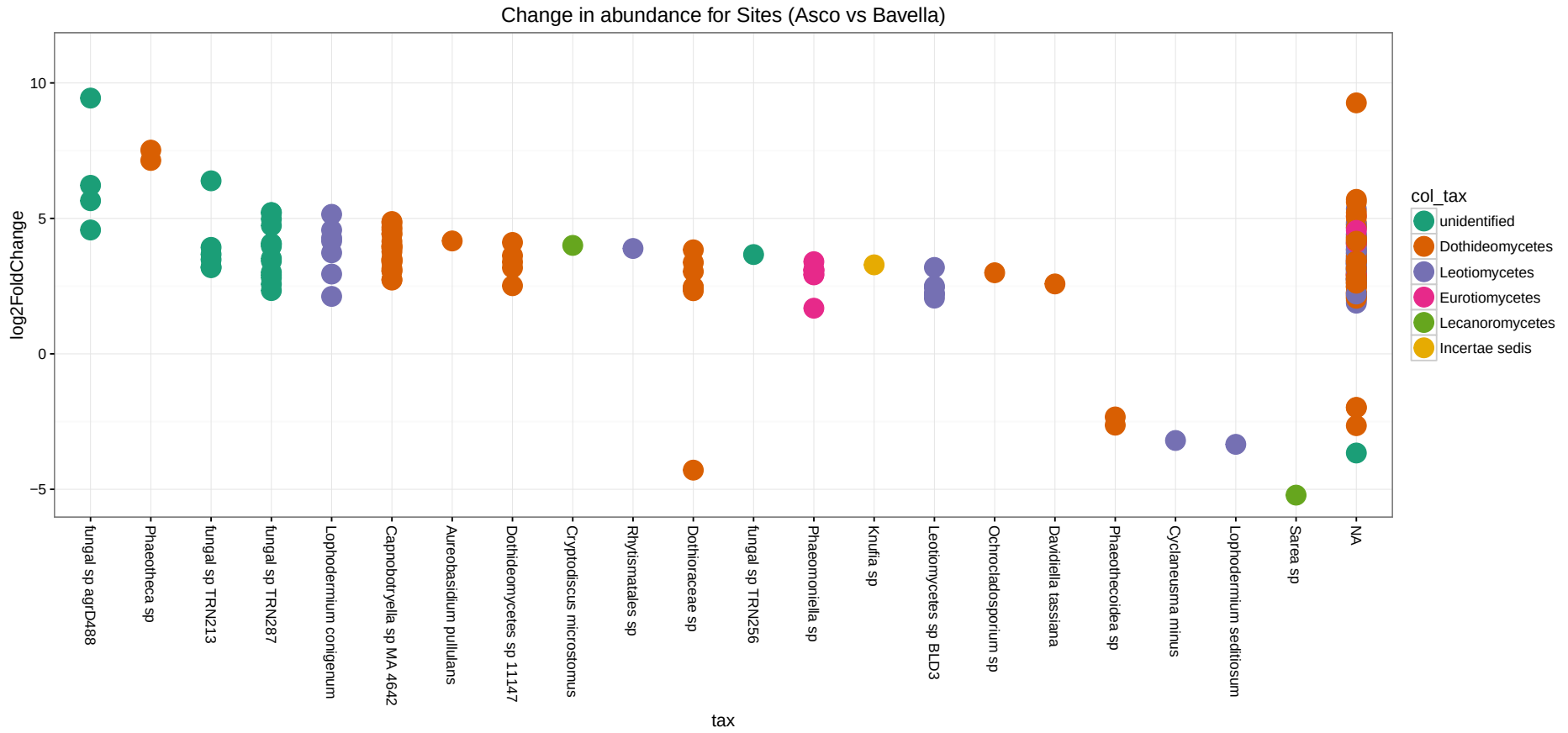

Figure 6.11: OTUs significantly different in terms of abundances between Asco (positive values) and Bavella (negative values)

```
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
```

6.6.2 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")
```

	Comparison	Species	Class	log2FoldChange (negative = more on second levels)
1	Verghello vs Asco			-3.77841397359728
2	Verghello vs Asco	Davidiella tassiana	Dothideomycetes	-2.3292426554423
3	Verghello vs Asco		Dothideomycetes	-3.61893267806188
4	Verghello vs Asco	fungal sp TRN213	unidentified	-3.13867374480923
5	Verghello vs Asco		Dothideomycetes	-3.75809969070393
6	Verghello vs Asco	Capnobotryella sp MA 4642	Dothideomycetes	-2.28440906171705
7	Verghello vs Asco			-3.39552445859464
8	Verghello vs Asco			-2.45061056468389
9	Verghello vs Asco		Dothideomycetes	-2.93876057054674
10	Verghello vs Asco			-2.1231609092589
11	Verghello vs Asco			-2.21785279122499
12	Verghello vs Asco			3.27030586093898
13	Verghello vs Asco	Dothideomycetes sp 11147	Dothideomycetes	-3.35876370715981
14	Verghello vs Asco	Phaeomoniella sp	Eurotiomycetes	-3.05588415677775
15	Verghello vs Asco			2.81017146488419
16	Verghello vs Asco		Dothideomycetes	-2.80304424723716
17	Verghello vs Asco	Capnobotryella sp MA 4642	Dothideomycetes	-2.58963016020498
18	Verghello vs Asco			-2.66392724218722
19	Verghello vs Asco		Leotiomycetes	-3.21880600708015
20	Verghello vs Asco			2.92930724447245
21	Verghello vs Asco		Leotiomycetes	-1.91162202262244
22	Verghello vs Asco	fungal sp TRN287	unidentified	-3.26119286517465
23	Verghello vs Asco		Dothideomycetes	-3.73051774614064
24	Verghello vs Asco		Dothideomycetes	-3.59888389077174
25	Verghello vs Asco			-2.83113825848711
26	Verghello vs Asco		Dothideomycetes	-2.5879763123238
27	Verghello vs Asco			3.35206655526952
28	Verghello vs Asco		Dothideomycetes	-2.9060336204246
29	Verghello vs Asco		Dothideomycetes	-3.67134758989312
30	Verghello vs Asco	fungal sp TRN287	unidentified	-3.50648758643916
31	Verghello vs Asco			3.49214798503904
32	Verghello vs Asco			-4.04018643085912
33	Verghello vs Asco		Dothideomycetes	-2.81488163023687
34	Verghello vs Asco	Phaeotheca sp	Dothideomycetes	-3.91229888934689
35	Verghello vs Asco		Dothideomycetes	-7.29191587441624
36	Verghello vs Asco			3.23002020397034
37	Verghello vs Asco	Capnobotryella sp MA 4642	Dothideomycetes	-2.63737112217892
38	Verghello vs Asco			-2.50206160823865
39	Verghello vs Asco			-2.82304112819981
40	Verghello vs Asco		Dothideomycetes	-3.27632848840878
41	Verghello vs Asco		Dothideomycetes	-2.85862844695998
42	Verghello vs Asco		Dothideomycetes	-3.07243778803445
43	Verghello vs Asco	Capnobotryella sp MA 4642	Dothideomycetes	-2.52643633508039
44	Verghello vs Asco			-3.39957820725733
45	Verghello vs Asco	Rhytismatales sp	Leotiomycetes	-4.41293144741018
46	Verghello vs Asco	Lachnellula calyciformis	Leotiomycetes	-4.09918269948416
47	Verghello vs Asco	fungal sp agrD488	unidentified	-6.69748501563904
48	Verghello vs Asco		Dothideomycetes	-3.52986856375373
49	Verghello vs Asco	Capnobotryella sp MA 4642	Dothideomycetes	-3.02821213013443
50	Verghello vs Asco	fungal sp agrD488	unidentified	-4.57678613187943
51	Verghello vs Asco			-4.70934332226896
52	Verghello vs Asco	fungal sp TRN287	unidentified	-3.33131311490098
53	Verghello vs Asco	Capnobotryella sp MA 4642	Dothideomycetes	-3.12702893768946
54	Verghello vs Asco		Dothideomycetes	-2.8017685199665
55	Verghello vs Asco			-10.6184764534919
56	Verghello vs Asco	Lophodermium seditiosum	Leotiomycetes	2.46708644436988
57	Verghello vs Asco	Phaeotheca sp	Dothideomycetes	2.74556852483771
58	Verghello vs Asco			3.33735059573991
59	Verghello vs Asco			3.33852254374311
60	Verghello vs Asco		Dothideomycetes	-3.59877394044465
61	Verghello vs Asco		Dothideomycetes	-2.81574973844836
62	Verghello vs Asco		Dothideomycetes	-5.31880689803234
63	Verghello vs Asco	Sarea sp	Lecanoromycetes	5.35004486549744
64	Verghello vs Asco	Sporormiaceae sp	Dothideomycetes	-5.27566358814124
65	Verghello vs Asco		Dothideomycetes	-3.88846863260061
66	Verghello vs Asco			5.97869844874395
67	Verghello vs Bavella	Phaeomoniella sp	Eurotiomycetes	2.92820119286806
68	Verghello vs Bavella		Dothideomycetes	1.84664832298385
69	Verghello vs Bavella	Lophodermium conigenum	Leotiomycetes	3.87776049046955
70	Verghello vs Bavella	fungal sp TRN287	unidentified	2.73394822182892
71	Verghello vs Bavella	Capnobotryella sp MA 4642	Dothideomycetes	2.46470174560228
72	Verghello vs Bavella		Dothideomycetes	2.69029588324704
73	Verghello vs Bavella	Leotiomycetes sp BLD3	Leotiomycetes	3.79199439132598
74	Verghello vs Bavella		Dothideomycetes	2.50907147594031
75	Verghello vs Bavella			3.67527127660293
76	Verghello vs Bavella		Leotiomycetes	3.90411109963067
77	Verghello vs Bavella		Leotiomycetes	3.76540742038509
78	Verghello vs Bavella			5.28089742779245
79	Verghello vs Bavella	Dothioraceae sp	Dothideomycetes	2.427402858263
80	Verghello vs Bavella	Phaeotheca sp	Dothideomycetes	5.00286308261026
81	Verghello vs Bavella			2.99842985863806
82	Verghello vs Bavella		Dothideomycetes	2.14743129703637
83	Verghello vs Bavella	Lophodermium conigenum	Leotiomycetes	3.93203634659419
84	Verghello vs Bavella		Dothideomycetes	2.16772856631734
85	Verghello vs Bavella	fungal sp TRN287	unidentified	3.03439420023129
86	Verghello vs Bavella			3.83667679584981
87	Verghello vs Bavella	Ochrocladosporium sp	Dothideomycetes	4.31719640571145
88	Verghello vs Bavella		Leotiomycetes	3.79823021765702
89	Verghello vs Bavella		Leotiomycetes	4.29395140158755
90	Verghello vs Bavella		Leotiomycetes	2.29421242892866
91	Verghello vs Bavella			6.5282066602472
92	Verghello vs Bavella		Dothideomycetes	-2.54157337757549
93	Verghello vs Bavella	Leotiomycetes sp BLD3	Leotiomycetes	3.22490791523471
94	Verghello vs Bavella	Ochrocladosporium sp	Dothideomycetes	4.46718046676545
95	Verghello vs Bavella	Knufia sp	Incertae sedis	2.98153304536828
96	Verghello vs Bavella	Capnobotryella sp MA 4642	Dothideomycetes	2.55071581339411
97	Verghello vs Bavella		Dothideomycetes	2.37437393788857
98	Verghello vs Bavella		Leotiomycetes	1.70475613364684
99	Verghello vs Bavella			4.61896240859855
100	Verghello vs Bavella	fungal sp TRN287	unidentified	3.9512454264957
101	Verghello vs Bavella		Dothideomycetes	3.35449916658489
102	Verghello vs Bavella	Lophodermium conigenum	Leotiomycetes	3.20990352420892
103	Verghello vs Bavella		Leotiomycetes	2.01227003493554
104	Verghello vs Bavella		Dothideomycetes	2.18720673517677
105	Verghello vs Bavella		Dothideomycetes	2.55080984617743
106	Verghello vs Bavella			4.94778862092973
107	Verghello vs Bavella		Leotiomycetes	3.02457651202801

	Comparison	Order	Class	log2FoldChange (negative = more on second levels)
1	Verghello vs Asco	Xylariales	Sordariomycetes	4.2444927272089
2	Verghello vs Bavella	Incertae sedis	Leotiomycetes	-1.3417319230562
3	Verghello vs Bavella	unidentified	unidentified	1.57935352182134
4	Asco vs Bavella	Botryosphaeriales	Dothideomycetes	6.05272114483586
5	Asco vs Bavella	Eurotiales	Eurotiomycetes	1.76179960851379
6	Asco vs Bavella	Incertae sedis	Leotiomycetes	-1.6831940181778
7	Asco vs Bavella	unidentified	unidentified	1.44997805431064
8	Asco vs Bavella	Xylariales	Sordariomycetes	-4.01099008009487

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

List of Figures

List of Tables