**Trait and phylogenetic diversity**

During the part of this course on community and metacommunity ecology, we have focused on species-level analyses. As we saw in the last lecture, looking at species traits can provide additional information on the processes structuring ecological communities and providing ecosystem services. But it also presents some important challenges, as we will see in this practical. Another approach that ecologists have used is based on species phylogeny. In today’s practical, we will see how to compute trait and phylogenetic gamma, alpha and beta diversity indices, and explore how they compare to each other.

**Species diversity**

We will use the PaciFlora database again, as for the practical on species-based community patterns. I will give you the code to import the data and compute gamma, alpha and beta diversity indices. We will do this for the islands belonging to the Society archipelago in French Polynesia.

library(tidyverse)

#######################

##Species communities##

#######################

dat.all <- read.csv("PaciFlora/Species\_list\_full\_2905.csv",sep=";")

dat.all <- dat.all[-which(is.na(dat.all$species)),]

dat.all <- dat.all[-which(is.na(dat.all$island)),]

dat.soc <- dat.all[which(dat.all$islandgroup=="Society"),]

length(unique(dat.soc$island))

dat.soc.red <- dat.soc[,c("species","island")]

dat.soc.red$presence <- 1

##reshape - pivot matrix

dat.soc.pa <- dat.soc.red %>%

pivot\_wider(names\_from=species,values\_from=c(presence))

list0 <- as.list(rep(0,ncol(dat.soc.pa)))

names(list0) <- names(dat.soc.pa)

dat.soc.pa <- as.data.frame(dat.soc.pa %>% replace\_na(list0))

row.names(dat.soc.pa) <- dat.soc.pa$island

dat.soc.pa <- dat.soc.pa[,-1]

dat.soc.pa <- dat.soc.pa[order(row.names(dat.soc.pa)),] ##ordering rows alphabetically

We can now compute gamma diversity as the total number of species across all island, and alpha diversity as the (average) number of species per island.

##Gamma

gamma.soc <- ncol(dat.soc.pa)

##Alpha

alpha.soc <- rowSums(dat.soc.pa)

alpha.soc.mean <- mean(alpha.soc)

For alpha diversity, we can also compute the ratio compared to gamma diversity, you will see why later:

alpha.soc.ratio <- alpha.soc/gamma.soc

alpha.soc.mean.ratio <- alpha.soc.mean/gamma.soc

We saw in class that we could partition total beta diversity, computed with the Jaccard or Sørensent indices, into two components: a replacement component, where the difference in species composition comes from a species being replaced by another, and a richness difference component, where species turnover is a result of species being added or subtracted from one site to another. There is nonetheless a debate in the literature about how to compute this partition.

Baselga (2010, 2015) uses the Simpson index of similarity to compute the replacement component, whereas Podani & Schmera (2011) and Cardoso et al. (2014) use Williams replacement index (see Table 1 below from Baselga 2015):

A table of scientific calculations

Description automatically generated with medium confidence

Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. Global ecology and biogeography, 19(1), 134-143.

Baselga, A., & Leprieur, F. (2015). Comparing methods to separate components of beta diversity. Methods in Ecology and Evolution, 6(9), 1069-1079.

Cardoso, P., Rigal, F., Carvalho, J.C., Fortelius, M., Borges, P.A.V., Podani, J. & Schmera, D. (2014) Partitioning taxon, phylogenetic and functional beta diversity into replacement and richness difference components. Journal of Biogeography, 41, 749-761.

Podani, J. & Schmera, D. (2011) A new conceptual and methodological framework for exploring and explaining pattern in presence-absence data. Oikos, 120, 1625-1638.

You can use the BAS (Simpson index) or the POD (Williams index) method (Table 1) using the functions betapart() and beta() from the betapart and the BAT R packages, respectively. Load the betapart and the BAT packages, and use the functions to compute the beta diversity indices for the two partitions (be careful of the parameters for function beta(), so that the outputs are comparable – the total beta diversity values should be the same for the two functions). Would they lead you to different conclusions?

If you managed to compute the beta diversity indices correctly, you can plot the outputs using the following code:

plot(beta.soc1$beta.sor,beta.soc2$Btotal,xlim=c(0,1),ylim=c(0,1))

lines(c(0,1),c(0,1),col="red")

plot(beta.soc1$beta.sim,beta.soc2$Brepl,xlim=c(0,1),ylim=c(0,1))

lines(c(0,1),c(0,1),col="red")

In the following, we will use the function beta() from the BAT R package, as it allows to compute species, trait and phylogenetic diversity.

**Trait diversity**

To compute trait diversity, we first need to get trait data for the species in the community. In practice, there can be two ways to do this: (i) using field studies, so that the trait values are specific to our study area and can account for phenotypic changes, or (ii) using general databases to get data at the species level. Here we will obviously use the second option.

But first, we need to select the traits we will include. Here we will use seed mass and plant height, which are amongst the most commonly available traits. It is possible to use more than two traits of course. In this case, it is usually convenient to go back to two dimensions using a principal component analysis or a principal coordinates analysis. Can you guess why?

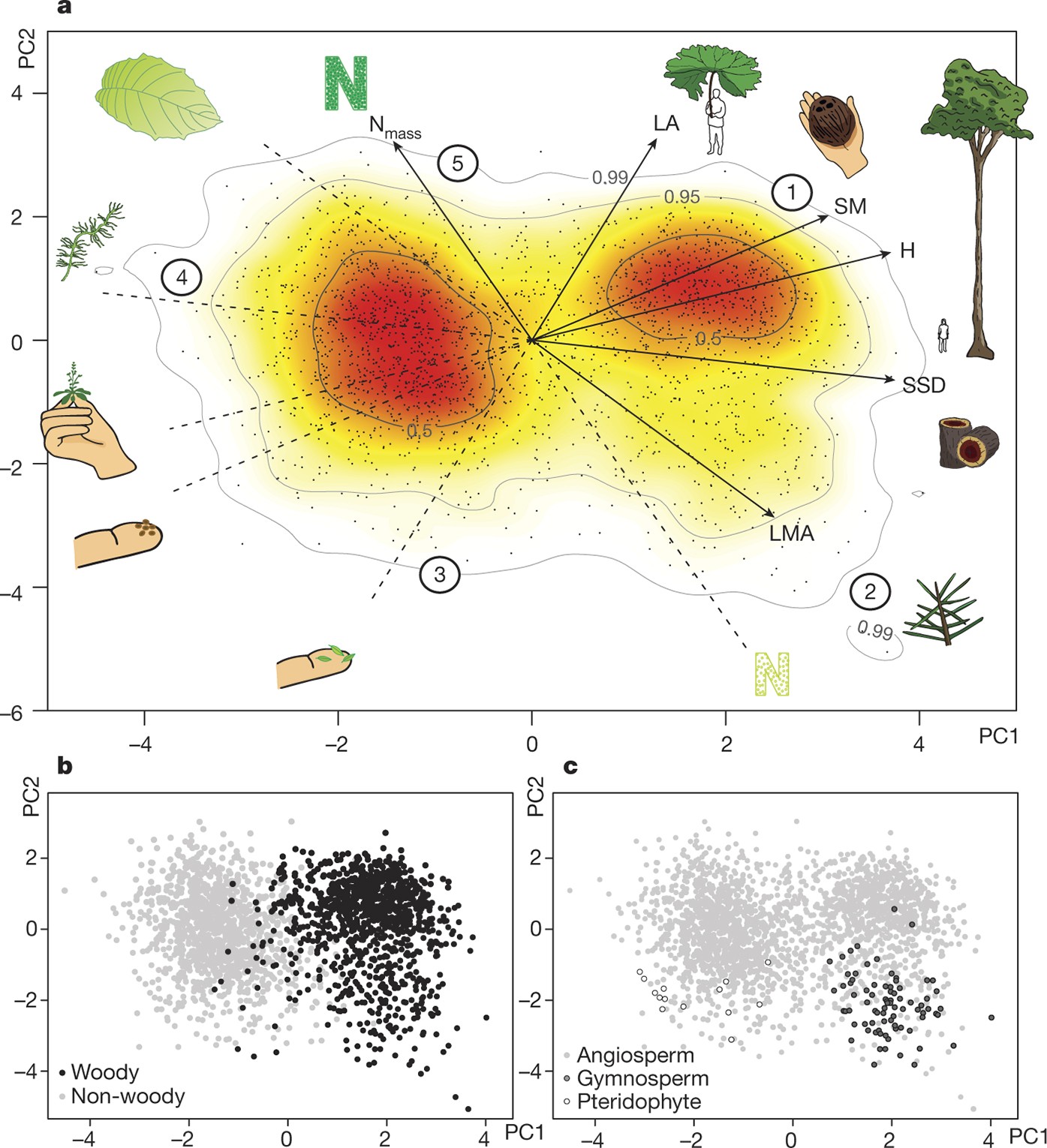


Figure 2: The global spectrum of plant form and function from Díaz et al. (2015). a, Projection of global vascular plant species (dots) on the plane defined by principal component axes (PC) 1 and 2 (details in Extended Data Table 1 and Extended Data Fig. 2). Solid arrows indicate direction and weighing of vectors representing the six traits considered; icons illustrate low and high extremes of each trait vector. Circled numbers indicate approximate position of extreme poles of whole-plant specialisation, illustrated by typical species (Extended Data Table 2). The colour gradient indicates regions of highest (red) to lowest (white) occurrence probability of species in the trait space defined by PC1 and PC2, with contour lines indicating 0.5, 0.95 and 0.99 quantiles (see Methods, kernel density estimation). Red regions falling within the limits of the 0.50 occurrence probability correspond to the functional hotspots referred to in main text. b, c, location of different growth-forms (b) and major taxa (c) in the global spectrum.

Díaz, S., Kattge, J., Cornelissen, J. et al. The global spectrum of plant form and function. Nature 529, 167–171 (2016). <https://doi.org/10.1038/nature16489>

Here we will use the TR8 R package, to extract this information from multiple plant databases. You can see all databases and the traits they contain included in the package by typing:

available\_traits()

As you can see, each database uses different names for the same traits. To get information on seed mass and plant height, we will extract the following variables:

names.spp.soc <- unique(dat.soc$species)

Soc.sp.tr8 <- tr8(species\_list = names.spp.soc, download\_list=c("seed\_mas\_cal","seed\_mass","seed\_wght","SeedMass","Height","h\_max","max\_height\_cal"),allow\_persistent=FALSE)

View(Soc.sp.tr8@results)

Look at your output. What quite obvious and problematic issues can you see?

The different databases can provide different values for the same trait, because they were estimated in different sites or regions, or using different methods. We will therefore average the values for each trait. First, we need to transform the character() data into numeric(). We have characters because some databases provide a range of values (e.g. “100;120” for *Abutilon theophrasti*). We will keep only the first value for simplification, using a regular expression:

library(stringi)

traits <- Soc.sp.tr8@results

traits$h\_max <- as.numeric(stri\_extract\_first\_regex(traits$h\_max, "[0-9]+"))

traits$seed\_wght <- as.numeric(stri\_extract\_first\_regex(traits$seed\_wght, "[0-9]+"))

traits$Height <- as.numeric(stri\_extract\_first\_regex(traits$Height, "[0-9]+"))

traits$SeedMass <- as.numeric(stri\_extract\_first\_regex(traits$SeedMass, "[0-9]+"))

View(traits)

We can now compute the mean for each trait from the different databases, when there are different values:

traits.mean <- data.frame(height=numeric(nrow(traits)),seed.mass=numeric(nrow(traits)))

row.names(traits.mean) <- row.names(traits)

traits.mean$height <- apply(traits[,c(1,4)],1,mean,na.rm=TRUE)

traits.mean$seed.mass <- apply(traits[,c(2,3,5)],1,mean,na.rm=TRUE)

As you can see, there are many trait values missing. This is actually a pretty common issue. In practice, you should use multiple databases (more than the ones included in TR8). Some databases will contain better data for your species than others. For example, AusTraits (<https://austraits.org/>) would have more data, as it was collated in a region of the world closer to French Polynesia, and where these species will be more common, as there will be a higher chance there will have also been introduced there (for native species, there would have been a higher chance they would be native in both regions). In Latombe et al. (2023) (<https://doi.org/10.1101/2023.03.28.534538>), we extracted information from TRY (Kattge et al., 2011, 2020), LEDA (Kleyer et al., 2008), PLANTATT (Hill et al., 2004), Austraits (Falster et al., 2021), BIEN (Maitner, 2022), EcoFlora (Fitter & Peat, 1994), and BROT (Tavşanoğlu & Pausas, 2018) (see references in the article).

We will nonetheless do our analyses for the 39 remaining species for the sake of this practical. First, we need to remove all columns with no data for both traits:

traits.mean <- traits.mean[-which(is.na(traits.mean$height) | is.na(traits.mean$seed.mass)),]

Also, the seed mass is 0g for a few species. This is obviously wrong. Here, for simplicity, we will just set these to 0.01g, assuming the seeds are very light. In practice, you may want to remove these aberrant values.

traits.mean[which(traits.mean==0,arr.ind = TRUE)] <- 0.01

We now need to remove all the species with no data in our community matrix too:

dat.soc.pa.trait <- dat.soc.pa[,which(colnames(dat.soc.pa) %in% row.names(traits.mean))]

And we need to remove the islands where we have fewer than 3 species with trait data present, so that we can compute a polygon in the trait space afterwards:

colSums(dat.soc.pa.trait)

rowSums(dat.soc.pa.trait)

dat.soc.pa.trait <- dat.soc.pa.trait[-which(rowSums(dat.soc.pa.trait)<3),]

Plot the traits of the 45 species in a two-dimensional space:

plot(traits.mean)

Can you see some issue for subsequent analyses?

The distributions are skewed with many small species with light seeds. We will therefore log-transform these values:

traits.mean$seed.mass <- log(traits.mean$seed.mass)

plot(traits.mean)

**Convex hull method**

We can now compute trait diversity (alpha and beta) using the convex hull method seen in class, using the function hull.build(), to compute the convex hulls, and the functions hull.alpha() and hull.beta() to compute trait alpha and beta diversity:

Soc.hull <- hull.build(comm = dat.soc.pa.trait, trait = traits.mean)

gamma.trait <- hull.gamma(comm = Soc.hull)

alpha.trait <- hull.alpha(comm = Soc.hull)

alpha.trait.ratio <- alpha.trait/gamma.trait

alpha.trait.mean <- mean(alpha.trait)

alpha.trait.mean.ratio <- alpha.trait.mean/gamma.trait

beta.trait <- hull.beta(comm = Soc.hull)

Let’s compare taxonomic and trait alpha diversity for the different islands. We can plot them in a nice way using the ggplot and ggrepel packages (which you will need to install if you don’t have them yet):

library(ggplot2)

library(ggrepel)

dat <- data.frame(island=names(alpha.trait.ratio),alpha.tax=alpha.soc.ratio[which(names(alpha.soc.ratio) %in% names(alpha.trait.ratio))],alpha.trait=alpha.trait.ratio)

ggplot(dat,aes(alpha.tax,alpha.trait,label=island))+

geom\_point()+

geom\_text\_repel()+

geom\_abline(intercept=0,slope=1,color="red")+

xlim(0,1)+

ylim(0,1)

And we can compare them quantitatively using Pearson and Spearman rank correlation indices:

cor(dat$alpha.tax,dat$alpha.trait,method="pearson")

cor(dat$alpha.tax,dat$alpha.trait,method="spearman")

Let’s compare taxonomic and trait beta diversity, for total and replacement beta diversity. First we need to only keep the islands that were considered in the trait analyses:

beta.soc2.temp <- lapply(beta.soc2,as.matrix)

beta.soc2.red <- list()

beta.soc2.red$Btotal <- c(as.dist(beta.soc2.temp$Btotal[which(row.names(beta.soc2.temp$Btotal) %in% names(beta.trait$Btotal)),which(colnames(beta.soc2.temp$Btotal) %in% names(beta.trait$Btotal))]))

beta.soc2.red$Brepl <- c(as.dist(beta.soc2.temp$Brepl[which(row.names(beta.soc2.temp$Brepl) %in% names(beta.trait$Brepl)),which(colnames(beta.soc2.temp$Brepl) %in% names(beta.trait$Brepl))]))

beta.soc2.red$Brich <- c(as.dist(beta.soc2.temp$Brich[which(row.names(beta.soc2.temp$Brich) %in% names(beta.trait$Brich)),which(colnames(beta.soc2.temp$Brich) %in% names(beta.trait$Brich))]))

We can now plot one against another:

par(mfrow=c(1,2)) plot(beta.soc2.red$Btotal,c(beta.trait$Btotal),xlim=c(0,1),ylim=c(0,1))

lines(c(0,1),c(0,1),col="red")

plot(beta.soc2.red$Brepl,c(beta.trait$Brepl),xlim=c(0,1),ylim=c(0,1))

lines(c(0,1),c(0,1),col="red")

We can also compute the correlations (Pearson and Spearman indices):

cor(beta.soc2.red$Btotal,c(beta.trait$Btotal),method="pearson")

cor(beta.soc2.red$Btotal,c(beta.trait$Btotal),method="spearman")

cor(beta.soc2.red$Brepl,c(beta.trait$Brepl),method="pearson")

cor(beta.soc2.red$Brepl,c(beta.trait$Brepl),method="spearman

What do you observe? How would you explain it?

**Dendrogram (tree-based) method**

The issue with using convex hulls, is that only species with extreme trait values in the community contribute to the computation of the diversity indices. Another approach is to use dendrograms. A dendrogram is a tree with the length of its branches representing distance between species or groups of species. The default algorithm is a neighbour-joining algorithm, but you can also use an average linkage clustering algorithm:

trait.tree1 <- tree.build(trait = traits.mean,func="nj") ##try different algorithms

trait.tree2 <- tree.build(trait = traits.mean, func="upgma") ##try different algorithms

par(mfrow=c(1,2))

plot(trait.tree1)

plot(trait.tree2)

This will obviously have an impact on the values of the gama, alpha and beta diversity indices:

gamma.dend1 <- gamma(comm=dat.soc.pa.trait,tree=trait.tree1)

alpha.dend1 <- alpha(comm=dat.soc.pa.trait,tree=trait.tree1)

alpha.dend.ratio1 <- alpha.dend1/c(gamma.dend1)

alpha.dend.mean.1 <- mean(alpha.dend1)

alpha.dend.mean.ratio1 <- alpha.dend.mean.1/gamma.dend1

beta.dend1 <- beta(comm=dat.soc.pa.trait,tree=trait.tree1)

gamma.dend2 <- gamma(comm=dat.soc.pa.trait,tree=trait.tree2)

alpha.dend2 <- alpha(comm=dat.soc.pa.trait,tree=trait.tree2)

alpha.dend.ratio2 <- alpha.dend2/c(gamma.dend2)

alpha.dend.mean.2 <- mean(alpha.dend2)

alpha.dend.mean.ratio2 <- alpha.dend.mean.2/gamma.dend2

beta.dend2 <- beta(comm=dat.soc.pa.trait,tree=trait.tree2)

par(mfrow=c(1,2))

plot(c(beta.dend1$Btotal),c(beta.dend2$Btotal),xlim=c(0,1),ylim=c(0,1))

lines(c(0,1),c(0,1),col="red")

plot(c(beta.dend1$Brepl),c(beta.dend2$Brepl),xlim=c(0,1),ylim=c(0,1))

lines(c(0,1),c(0,1),col="red")

How do the diversity values change for different algorithms used to generate the dendrogram, and how do they differ from the convex hull approach and to species diversity values? You can use the following code to compare the alpha diversity indices:

dat <- data.frame(island=names(alpha.trait.ratio),alpha.tax=alpha.soc.ratio[which(names(alpha.soc.ratio) %in% names(alpha.trait.ratio))],alpha.trait=c(alpha.dend.ratio1))

ggplot(dat,aes(alpha.tax,alpha.trait,label=island))+

geom\_point()+

geom\_text\_repel()+

geom\_abline(intercept=0,slope=1,color="red")+

xlim(0,1)+

ylim(0,1)

cor(dat$alpha.tax,dat$alpha.trait,method="pearson")

cor(dat$alpha.tax,dat$alpha.trait,method="spearman")

dat <- data.frame(island=names(alpha.trait.ratio),alpha.tax=alpha.soc.ratio[which(names(alpha.soc.ratio) %in% names(alpha.trait.ratio))],alpha.trait=c(alpha.dend.ratio2))

ggplot(dat,aes(alpha.tax,alpha.trait,label=island))+

geom\_point()+

geom\_text\_repel()+

geom\_abline(intercept=0,slope=1,color="red")+

xlim(0,1)+

ylim(0,1)

cor(dat$alpha.tax,dat$alpha.trait,method="pearson")

cor(dat$alpha.tax,dat$alpha.trait,method="spearman")

And you can use the following code to compare the beta diversity indices:

plot(beta.soc2.red$Btotal,c(beta.dend1$Btotal),xlim=c(0,1),ylim=c(0,1))

lines(c(0,1),c(0,1),col="red")

plot(beta.soc2.red$Brepl,c(beta.dend1$Brepl),xlim=c(0,1),ylim=c(0,1))

lines(c(0,1),c(0,1),col="red")

plot(beta.soc2.red$Btotal,c(beta.dend2$Btotal),xlim=c(0,1),ylim=c(0,1))

lines(c(0,1),c(0,1),col="red")

plot(beta.soc2.red$Brepl,c(beta.dend2$Brepl),xlim=c(0,1),ylim=c(0,1))

lines(c(0,1),c(0,1),col="red")

plot(c(beta.trait$Btotal),c(beta.dend1$Btotal),xlim=c(0,1),ylim=c(0,1))

lines(c(0,1),c(0,1),col="red")

plot(c(beta.trait$Brepl),c(beta.dend1$Brepl),xlim=c(0,1),ylim=c(0,1))

lines(c(0,1),c(0,1),col="red")

plot(c(beta.trait$Btotal),c(beta.dend2$Btotal),xlim=c(0,1),ylim=c(0,1))

lines(c(0,1),c(0,1),col="red")

plot(c(beta.trait$Brepl),c(beta.dend2$Brepl),xlim=c(0,1),ylim=c(0,1))

lines(c(0,1),c(0,1),col="red")

**Phylogenetic diversity**

Since trait values can be missing, as we just saw, an alternative is to compute phylogenetic diversity. Phylogenetic diversity is interesting on its own from an evolutionary perspective, but it can also be used as a proxy for taxonomic diversity, under the assumption that species that are phylogenetically close or distant, will have similar or different traits.

In this case, we will import the phylogeny for the PaciFlora data, which is included in the database, but you will first need to install the ggtree package:

library(BiocManager)

BiocManager::install("ggtree")library(ggtree)

library(ape)

PaciFlora.tree <- read.tree("PaciFlora/Smith\_Brown\_plus\_Wohlwend\_tree.tre")

After importing the phylogeny, we will keep only the species present the Society archipelago:

names.spp.soc <- unique(sub(" ", "\_", dat.soc$species)) ##species names need an underscore instead of a space in the tree

Soc.tree <- keep.tip(PaciFlora.tree,tip=which(PaciFlora.tree$tip.label %in% names.spp.soc))

And we can plot the tree:

ggtree(Soc.tree,layout="circular")# + geom\_tiplab(color='firebrick') ## the geom\_tiplab() function adds the species names, but it is nor really readable for that amount of species

We then must only keep the species that are included in the phylogeny (there are a few missing – data is rarely perfect!) in our presence-absence matrix:

dat.soc.pa.tree <- dat.soc.pa

colnames(dat.soc.pa.tree) <- sub(" ", "\_",colnames(dat.soc.pa.tree))

dat.soc.pa.tree <- dat.soc.pa.tree[,which(colnames(dat.soc.pa.tree) %in% Soc.tree$tip.label)]

We can then compute gamma, alpha and beta diversity. Since we have a tree, it is done exactly as for the tree-based method used for trait diversity:

gamma.phylo <- gamma(comm=dat.soc.pa.tree,tree=Soc.tree)

alpha.phylo <- alpha(comm=dat.soc.pa.tree,tree=Soc.tree)

alpha.phylo.ratio <- alpha.phylo/c(gamma.phylo)

alpha.phylo.mean <- mean(alpha.phylo)

alpha.phylo.mean.ratio <- alpha.phylo.mean/c(gamma.phylo)

beta.phylo <- beta(comm=dat.soc.pa.tree,tree=Soc.tree)

How do these indices compare to the species and trait diversity indices? You can use the following code to compare them:

##plot taxonomic vs phylogenetic alpha diversity

dat <- data.frame(island=names(alpha.soc.ratio),alpha.tax=alpha.soc.ratio,alpha.phylo=c(alpha.phylo.ratio))

ggplot(dat,aes(alpha.tax,alpha.phylo,label=island))+

geom\_point()+

geom\_text\_repel()+

geom\_abline(intercept=0,slope=1,color="red")+

xlim(0,1)+

ylim(0,1)

cor(alpha.soc.ratio,c(alpha.phylo.ratio))

cor(alpha.soc.ratio,c(alpha.phylo.ratio),method="spearman")

##plot taxonomic vs phylogenetic beta diversity

plot(beta.soc2$Btotal,(beta.phylo$Btotal),xlim=c(0,1),ylim=c(0,1))

lines(c(0,1),c(0,1),col="red")

plot(beta.soc2$Brepl,(beta.phylo$Brepl),xlim=c(0,1),ylim=c(0,1))

lines(c(0,1),c(0,1),col="red")

##remove islands for which we could not compute trait diversity

beta.phylo.temp <- lapply(beta.phylo,as.matrix)

beta.phylo.red <- list()

beta.phylo.red$Btotal <- c(as.dist(beta.phylo.temp$Btotal[which(row.names(beta.phylo.temp$Btotal) %in% names(beta.trait$Btotal)),which(colnames(beta.phylo.temp$Btotal) %in% names(beta.trait$Btotal))]))

beta.phylo.red$Brepl <- c(as.dist(beta.phylo.temp$Brepl[which(row.names(beta.phylo.temp$Brepl) %in% names(beta.trait$Brepl)),which(colnames(beta.phylo.temp$Brepl) %in% names(beta.trait$Brepl))]))

beta.phylo.red$Brich <- c(as.dist(beta.phylo.temp$Brich[which(row.names(beta.phylo.temp$Brich) %in% names(beta.trait$Brich)),which(colnames(beta.phylo.temp$Brich) %in% names(beta.trait$Brich))]))

##plot phylogenetic vs trait diversity

plot(c(beta.phylo.red$Btotal),c(beta.trait$Btotal),xlim=c(0,1),ylim=c(0,1))

lines(c(0,1),c(0,1),col="red")

plot(c(beta.phylo.red$Brepl),c(beta.trait$Brepl),xlim=c(0,1),ylim=c(0,1))

lines(c(0,1),c(0,1),col="red")

Based on these differences, would you draw different conclusions about plant invasions in French Polynesia, and would that affect management recommendations?

**Null models**

We can run null models on the phylogenetic analyses, to assess if there is an influence of phylogeny on how alien species are distributed across the islands. Here the idea is to keep the taxonomic alpha and beta diversity values constant, and to remove the effect of phylogeny, i.e. we will randomize the position of the species in the phylogenetic tree, while keeping the same tree structure and the same site-by-species matrix.For the sake of keeping computational times reasonable, we will only run 99 null models.

reps <- 99 ##number of null models

##initialise the matrices and vectors to store the data

gamma.phylo.rand <- alpha.phylo.mean.rand <- alpha.phylo.mean.ratio.rand <- numeric(reps)

alpha.phylo.rand <- alpha.phylo.ratio.rand <- matrix(NA,reps,length(alpha.phylo))

beta.phylo.total.rand <- beta.phylo.repl.rand <- matrix(NA,reps,length(c(beta.phylo$Btotal)))

##Compute the null models

for(i in 1:reps){

print(i)

dat.soc.pa.tree.temp <- dat.soc.pa.tree ##let’s duplicate the site-by-species matrix

names(dat.soc.pa.tree.temp) <- names(dat.soc.pa.tree.temp)[sample(ncol(dat.soc.pa.tree.temp),ncol(dat.soc.pa.tree.temp))] ##randomise the order of the species names

##compute the patterns for the null model

gamma.phylo.rand[i] <- gamma(comm=dat.soc.pa.tree.temp,tree=Soc.tree)

alpha.phylo.rand[i,] <- alpha(comm=dat.soc.pa.tree.temp,tree=Soc.tree)

alpha.phylo.ratio.rand[i,] <- alpha.phylo/c(gamma.phylo)

alpha.phylo.mean.rand[i] <- mean(alpha.phylo)

alpha.phylo.mean.ratio.rand[i] <- alpha.phylo.mean/c(gamma.phylo)

toto <- beta(comm=dat.soc.pa.tree.temp,tree=Soc.tree)

beta.phylo.total.rand[i,] <- c(toto$Btotal)

beta.phylo.repl.rand[i,] <- c(toto$Brepl)

}

mean(alpha.phylo) ##average observed alpha diversity over all islands

quantile(c(mean(alpha.phylo),rowMeans(alpha.phylo.rand)),c(0.025,0.05,0.95,0.975)) ##quantiles of the 99 average simulated phylogenetic alpha diversity and the average observed alpha diversity over all islands

mean(beta.phylo$Btotal) ##average observed Sorensen beta diversity over all islands

quantile(c(mean(beta.phylo$Btotal),rowMeans(beta.phylo.total.rand)),c(0.025,0.05,0.95,0.975)) ##quantiles of the 99 average simulated phylogenetic Sorensen beta diversity and the average observed Sorensen beta diversity over all islands

mean(beta.phylo$Brepl) ##average observed Williams beta diversity over all islands

quantile(c(mean(beta.phylo$Brepl),rowMeans(beta.phylo.repl.rand)),c(0.025,0.05,0.95,0.975)) ##quantiles of the 99 average simulated phylogenetic Williams beta diversity and the average observed Williams beta diversity over all islands

##Let’s plot the simulated histograms vs the observed values

par(mfrow=c(1,3))

hist(rowMeans(alpha.phylo.rand),xlim=c(5800,7300))

abline(v=mean(alpha.phylo),col="red")

hist(rowMeans(beta.phylo.total.rand),xlim=c(0.65,0.75))

abline(v=mean(beta.phylo$Btotal),col="red")

hist(rowMeans(beta.phylo.repl.rand),xlim=c(0.1,0.2))

abline(v=mean(beta.phylo$Brepl),col="red")

##We can also plot the simulated histograms vs the observed values for alpha diversity for each island separately

par(mfrow=c(4,4))

for(i in 1:14){

hist(alpha.phylo.rand[,i],xlim=c(min(c(alpha.phylo.rand[,i],alpha.phylo[i]))\*0.9,max(c(alpha.phylo.rand[,i],alpha.phylo[i]))\*1.1),main=row.names(alpha.phylo)[i])

abline(v=mean(alpha.phylo[i]),col="red")

}

##and we can plot this as a graph

dat <- data.frame(island=names(alpha.soc.ratio),alpha.tax=alpha.soc.ratio,alpha.phylo=c(alpha.phylo.ratio))

datt <- data.frame(island=rep(names(alpha.soc.ratio),each=99),alpha.tax=rep(alpha.soc.ratio,each=99),alpha.phylo=c(alpha.phylo.ratio.rand))

ggplot()+

geom\_point(datt,mapping=aes(alpha.tax,alpha.phylo))+

geom\_point(dat,mapping=aes(alpha.tax,alpha.phylo),colour="red")+

geom\_label\_repel(dat,mapping=aes(alpha.tax,alpha.phylo,label=island),direction = "x")+

geom\_abline(intercept=0,slope=1,color="red")+

xlim(0,1)+

ylim(0,1)

Have your conclusions changed after using null models?

**It was too easy and you have additional time?**

**Expand the analyses**

Compute the same indices for the other archipelagos seen in the practical on species community patterns (Hawai’i, Samoa, Marquesas, and Fiji). Do the patterns behave the same way for the different archipelagos?

**Think out of the box**

If you go back to the species community patterns from the previous practicals (e.g. species accumulation curves and occupancy-frequency distributions), can you think of how these could be adapted to traits or phylogeny?