

# Guide through the functions available in package `adiv`

Sandrine Pavoine

August, 2021

## 1 List of main functions

The functions in `adiv` have different objectives. Among the core functions of `adiv` are those dedicated to the measurement of diversity in an assemblage ( $\alpha$  diversity, Table 1.1).

### 1.1 $\alpha$ diversity

`adiv` contains two functions to calculate species diversity, but each function implements several mathematical formulas (or indices). Function `speciesdiv` permits the calculation of some of the most used diversity indices, including the Shannon index (Shannon, 1948). Even if it allows for the consideration of abundance data, function `speciesdiv` can also calculate species richness in assemblages, that is to say the number of present species. The second function, named `divparam`, implements diversity indices said parametric because they depend on a parameter that controls the importance given to rare compared to abundant species when measuring diversity.

In the `adiv` package, four main functions consider the phylogeny of species when measuring their diversity: `pIa`, `qDT`, `evodiv` and `evodivparam`. `pIa`, `qDT`, and `evodivparam` implement parametric indices. In a founding article, Faith (1992) extended species richness to consider the phylogeny of species when measuring their diversity. In `adiv` package, Faith's phylogenetic diversity (PD) index, which is the sum of the branch lengths on a phylogenetic tree, is available in a dedicated function named `EH` for evolutionary history to avoid any confusion with the concept of phylogeny-based diversity that is more commonly used in the package. Indeed, when the branch lengths in an ultrametric phylogenetic tree are expressed as divergence times, the sum of the branch lengths in that tree indicates the amount of evolutionary history. `adiv` also contains related functions for example to select  $k$  species out of  $n$  to optimize the amount of evolutionary history that would be preserved if they were saved and the others were driven to extinction (function `optimEH`, see also functions `randEH` and `orisaved`). Functions `pIa`, `qDT`, `evodiv` and `evodivparam` in `adiv` all generalize Faith's PD index to include information on species abundance when measuring phylogenetic diversity, although they do this generalization in different ways.

`evodiv` and `evodivparam` simply replace species in species diversity indices by features, which are branch length units on the phylogenetic tree where species are the tips. A branch of length  $L$  is assumed to support  $L$  features. A feature is assumed to be present in a community if the community contains at least one of the species that descend from that feature on the phylogenetic tree. The abundance of a feature in the community is the summed abundance of all species descending from that feature. For indices requiring relative abundance rather than absolute abundance, the abundance of each feature is normalized by the summed abundance of all features (see Pavoine and Ricotta, 2019a, for a detailed description of that framework). Contrary to the indices developed in `evodiv` and `evodivparam`, in `pIa` and `qDT`, traditional diversity indices are applied to the abundance of (phylogenetic) features normalized by the total abundance of all species at a site (Pavoine et al., 2009; Chao et al., 2010). With this normalization, these functions still focus on the species as the target unit to measure diversity, even if they consider phylogenetic data. The selection among these different approaches (`evodiv` and `evodivparam` versus `pIa` or `qDT`) requires deciding which unit is of most interest for a given study: species (characterized by their phylogeny) or features (branch units on the phylogenetic tree).

Trait-based diversity can be quantified by the same functions as those dedicated to phylogenetic diversity (`pIa`, `qDT`, `evodiv` and `evodivparam`) if a trait-based dendrogram is obtained and used to

Table 1.1: `adiv` functions dedicated to the measurement of diversity and evenness.

Main function	Measured quantities	Associated functions
Species diversity		
<code>speciesdiv</code>	diversity indices that rely on relative or absolute species abundance (Gini, 1912; Shannon, 1948; Simpson, 1949; Menhinick, 1964; McIntosh, 1967; Margalef, 1972; Magurran, 2004) and species richness (the number of species)	
<code>divparam</code>	diversity indices where a parameter controls the relative importance given to rare versus abundant species (Renyi, 1960; Havrda and Charvat, 1967; Daroczy, 1970; Hill, 1973; Tsallis, 1988)	<code>plot.divparam</code>
Species evenness		
<code>specieeve</code>	evenness indices that rely on relative or absolute species abundance (McIntosh, 1967; Heip, 1974; Pielou, 1975; Smith and Wilson, 1996; Magurran, 2004)	
<code>eveparam</code>	evenness indices where a parameter controls the relative importance given to rare versus abundant species (Hill, 1973)	<code>plot.eveparam</code>
Tree-based diversity (e.g. phylogenetic diversity)		
<code>EH</code>	sum of branch lengths on a phylogenetic tree (Faith, 1992)	<code>optimEH</code> , <code>randEH</code> , <code>orisaved</code>
<code>evodiv</code>	diversity indices that rely on the presence, or the relative or absolute abundance of features on a phylogenetic tree, with the assumption that the number of features on a given branch of a phylogenetic tree is equal to the length of this branch (Pavoine, 2016; Pavoine and Ricotta, 2019a)	
<code>evodivparam</code>	parametric diversity indices applied to phylogenetic diversity (Pavoine and Ricotta, 2019a)	<code>plot.evodivparam</code>
<code>FPdivparam</code>	Parametric Indices of Functional and Phylogenetic Diversity (Pavoine and Ricotta, 2021)	<code>plot.FPdivparam</code>
<code>pIa</code>	index $I_a$ by Pavoine et al. (2009)	
<code>qDT</code>	${}^qD(T)$ index developed by Chao et al. (2010) as the mean diversity of order $q$ over $T$ years in a phylogenetic tree. In function <code>qDT</code> , the index is computed over the whole tree from root to tips. The function uses the formula of the ${}^qD(T)$ index extended to non-ultrametric trees (where the distance from tip to root can vary).	
Tree-based evenness (e.g. phylogenetic evenness)		
<code>evoveparam</code>	Hill, HCDT and Renyi indices applied to phylogenetic evenness (Pavoine and Ricotta, 2019a)	<code>plot.evoveparam</code>
(Dis)similarity-based diversity		
<code>QE</code>	Rao's quadratic entropy within communities (Rao, 1982)	<code>discomQE</code> , <code>rare_Rao</code> , <code>Rentropy*</code>
<code>qHdiv</code>	parametric index of functional and phylogenetic diversity that includes consistent interspecific and intraspecific components (Pavoine and Izsak, 2014)	<code>twoHmax</code>
<code>Rentropy</code>	Pavoine et al. (2017) functional or phylogenetic entropy	<code>QE</code>
<code>FPdivparam</code>	Parametric Indices of Functional and Phylogenetic Diversity (Pavoine and Ricotta, 2021)	<code>plot.FPdivparam</code>

\* The quadratic entropy is also the core index used in functions `decdiv`, `wapqe`, `EqRao`, `EqRS`, `EqRSintra` (see Table 1.3).

---

replace the phylogenetic tree. The first step to obtain the dendrogram consists in calculating trait-based dissimilarities between species. For that, several functions can be used depending on the type of traits considered. For example, function `dsimFun` in package `adiv`, function `gowdis` in package `FD`, and `dist.ktab` in package `ade4` can handle multiple types of traits (e.g. quantitative, nominal, binary, and also fuzzy, circular for function `dist.ktab`). Then the function `hclust` of package `stats` can be applied to trait-based dissimilarities between species. Alternatively, the `adiv` package implements diversity indices that include a matrix of (trait-based) dissimilarities or a matrix of (trait-based) similarities between species in addition to abundance data on the species (functions `QE`, `qHdiv`, `Rentropy`, Table 1.1) without the need to build a trait-based dendrogram.

Functions in `adiv` can also evaluate the evenness in species abundances in a community (Table 1.1). A modification of function `evodivparam` also allows for the evaluation of the evenness in feature abundance (`evoveparam` function, Table 1.1).

Other functions allow the measurement of trait-based, taxonomic or phylogenetic (dis)similarities between species and species-based, trait-based, taxonomic or phylogenetic dissimilarities between communities (Table 1.2).

## 1.2 On the species-to-species dissimilarity indices available in `adiv`

The few functions in `adiv` dedicated to the calculation of dissimilarities or similarities between species using trait, taxonomic or phylogenetic data (Table 1.2) complement others found in other packages like function `gowdis` in package `FD` (Laliberte et al., 2014), or `dist.ktab` in package `ade4` (Thioulouse et al., 2018) for functional dissimilarities, like function `cophenetic.phylo` in package `ape` (Paradis et al., 2004) or `distTips` in package `ade4phylo` (Jombart and Dray, 2008) for phylogenetic dissimilarities, or like function `dist.taxo` in package `ade4` for taxonomic dissimilarities.

Compared to these alternative functions, the functions `CFprop`, `multiCFprop`, `CFbinary`, `multiCFbinary`, `dsimFun`, `dsimTax`, and `dsimTree` present in package `adiv` have the added benefit of leading to positive semi-definite matrices of similarities and, for `dsimFun`, `dsimTax`, and `dsimTree`, to dissimilarities that are squared Euclidean. A matrix  $\mathbf{A} = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n}$  is positive semi-definite (= non-negative definite) if, for any vector  $\mathbf{x} = (x_1 \dots x_n)^t$ ,  $\sum_{i=1}^n \sum_{j=1}^n x_i x_j a_{ij} \geq 0$ . Here  $(x_1 \dots x_n)^t$  denotes the transpose of  $(x_1 \dots x_n)$ . A matrix  $\mathbf{D} = (d_{ij})_{1 \leq i \leq n, 1 \leq j \leq n}$  is squared Euclidean if one can find  $n$  points  $M_1, \dots, M_n$  in a Euclidean space, so that the Euclidean distance between any two points  $M_i, M_j$  is  $\sqrt{d_{ij}}$  (Gower and Legendre, 1986).

These properties are exploited in function `dsimcom` of `adiv` to define the phylogenetic or functional dissimilarity between two communities, as indicated in Pavoine and Ricotta (2014). They are also exploited in diversity indices, for example in the parametric diversity coefficient  ${}^qH$  developed by Pavoine and Izsak (2014) (function `qHdiv` in `adiv`) to measure trait-based and phylogenetic diversities and in ordination analyses (see for example function `dsimpca` in `adiv`).

## 1.3 Plot-to-plot dissimilarities and $\beta$ diversity

Only few approaches exist in `adiv` to calculate dissimilarity between communities on the basis of the presence-absence or abundance of species (see e.g., functions `Jac`, `distMS`, Table 1.2 and functions `betas-tatjac` and `betastatsor`, Table 1.3). This is because other packages of R implement such approaches, see e.g., function `vegdist` in package `vegan` (Oksanen et al., 2019) or `dist.binary` in package `ade4`. Package `adiv` focuses on trait-based (functional) and phylogenetic dissimilarities between communities. It can calculate dissimilarities between communities using species presences-absences or abundances in the communities and a matrix of (trait-based or phylogenetic) (dis)similarities between species (e.g., functions `discomQE`, `dissRicotta`, and `dsimcom`, Table 1.2). It can also calculate dissimilarities between communities using species presences-absences or abundances in the communities and a phylogenetic (or trait-based) tree with species as tips (functions `DP`, `evodiss`, and `evodiss_family`, Table 1.2).

Table 1.2: `adiv` functions dedicated to the measurement of (dis)similarities between species and between communities.

Main function	Aim	Associated functions
Similarities between species based on trait, taxonomic or phylogenetic data		
<code>multiCFprop</code>	intra- and inter-specific similarities applied to traits expressed as proportions ( <a href="#">Pavoine and Izsak, 2014</a> )	<code>CFprop</code>
<code>multiCFbinary</code>	intra- and inter-specific similarities applied to binary traits ( <a href="#">Pavoine and Izsak, 2014</a> )	<code>CFbinary</code>
<code>dsimFun</code>	pair-wise functional (dis)similarities between species ( <a href="#">Pavoine and Ricotta, 2014</a> )	
<code>dsimTax</code>	pair-wise taxonomic (dis)similarities between species ( <a href="#">Pavoine and Ricotta, 2014</a> )	
<code>dsimTree</code>	pair-wise (dis)similarities between species that rely on a (phylogenetic) tree with species as tips ( <a href="#">Pavoine and Ricotta, 2014</a> )	
Dissimilarities between plots based on species presence-absence data		
<code>Jac</code>	additive decomposition of the Jaccard index into turnover and richness difference ( <a href="#">Ricotta, Podani and Pavoine, 2016</a> )	<code>DJac</code> , <code>PADDis</code>
Dissimilarities between plots based on species abundance data		
<code>distMS</code>	Marczewski-Steinhaus coefficient of dissimilarity ( <a href="#">Orloci, 1978</a> ; <a href="#">Legendre and Legendre, 1998</a> )	
(Dis)similarities between plots based on dissimilarities between species and species' presence/absence		
<code>dissABC</code> , <code>generalized_Traddiss</code> and <code>PADDis</code>	coefficients of similarity between communities that extend compositional similarity indices such as Jaccard and Sørensen indices to functional and phylogenetic similarity ( <a href="#">Ricotta and Pavoine, 2015b</a> ; <a href="#">Ricotta, Podani and Pavoine, 2016</a> ; <a href="#">Pavoine and Ricotta, 2019b</a> )	<code>Jac</code> , <code>DJac</code>
Dissimilarities between plots based on dissimilarities between species and species' abundance		
<code>discomQE</code>	Rao's coefficient of the dissimilarity between communities (DISC index, <a href="#">Rao, 1982</a> )	<code>QE</code>
<code>dissRicotta</code>	Ricotta et al. coefficient of the dissimilarities between communities ( <a href="#">Ricotta et al., 2015</a> )	
<code>dsimcom</code>	<a href="#">Pavoine and Ricotta (2014)</a> generalization of compositional similarity indices such as Jaccard and Sørensen indices to include information on functional and phylogenetic similarity between species and abundance data for each species	<code>sQ</code>
<code>dislptransport</code>	<a href="#">Kosman (1996)</a> and <a href="#">Gregorius et al. (2003)</a> coefficient of the dissimilarities between communities ( <a href="#">Ricotta et al., 2021</a> )	
Dissimilarities between plots based on a (phylogenetic) tree with species as tips		
<code>DP</code>	<a href="#">Ricotta et al. (2020)</a> plot-to-plot functional or phylogenetic dissimilarity (index named $D_F$ for functional data and $D_P$ for phylogenetic data, calculated with equation 2 in <a href="#">Ricotta et al., 2020</a> )	
<code>evodiss</code>	PD-dissimilarity indices described and/or discussed in <a href="#">Pavoine (2016)</a>	
<code>evodiss_family</code>	family of indices defined by <a href="#">Nipperess et al. (2010)</a> (see <a href="#">Pavoine, 2016</a> )	<code>evodiss_ternaryplot</code>

---

## 1.4 Partitioning of diversity ( $\alpha$ , $\beta$ , $\gamma$ )

The measurement of the diversity within communities ( $\alpha$  diversity) and that of the diversity between communities ( $\beta$  diversity) complements each other as both contribute to the global diversity of all combined communities ( $\gamma$  diversity). `adiv` contains functions to partition  $\gamma$  diversity into a component of  $\alpha$  diversity and one or several components of  $\beta$  diversity (Table 1.3). For example several components of  $\beta$  diversity are needed if diversity is studied across nested spatial scales (plot, site, region, continent, etc.). In the current version of `adiv`, these partitioning approaches depend on Rao's quadratic entropy (see function `QE` and Rao, 1982).

Other approaches allow:

1. partitioning phylogenetic diversity between evolutionary periods, indicating which amount of independent evolutionary history the species in a community represent (function `aptree` and related functions, Table 1.3).
2. partitioning phylogenetic  $\alpha$ ,  $\beta$ ,  $\gamma$  diversities between evolutionary periods to reveal at which evolutionary periods two or more communities differ in their composition. For example, if two communities have different species and if these species speciated a long time ago, then the partitioning approach will reveal that the compositions of the communities have profound phylogenetic differences; in contrast, if two communities have different species but if these species splitted very recently, then the partitioning approach will reveal that the evolutionary period that best represent the differences between the compositions of the communities is recent (function `abgaptree`, Table 1.3). This approach was defined by Pavoine et al. (2009), see also Chiu et al. (2014) for an alternative approach.
3. partitioning species  $\beta$  diversity into species nestedness and turnover (functions `betastatjac` and `betastatsor`, table 1.3).

## 1.5 Ordination analyses

`adiv` contains functions to display and depict trait-based and phylogenetic differences between communities thanks to ordination analyses (Table 1.4). Function `rlqESLTP` implements Pavoine et al. (2011) extension of the RLQ ordination approach (Doledec et al., 1996). It searches connections of species' traits and phylogenetic positions with the environmental conditions and spatial positions where they have been observed. Functions `evoCA`, `evoNSCA`, `evopcachord`, and `evopcahellinger` use a matrix of presences-absences or abundances of species in communities and a phylogenetic tree with species as tips to apply correspondence analysis (function `evoNSCA`), non-symmetric correspondence analysis (Lauro and D'Ambra, 1984; Kroonenberg and Lombardo, 1999, function `evoNSCA`), or principal component analysis (functions `evopcachord` and `evopcahellinger`) to the abundance of phylogenetic features in communities. As for phylogenetic diversity, a feature here represents a unit of branch length on a phylogenetic tree. A feature on a given branch is assumed to be present in a community if one of the species descending from the branch occurs in the community. The abundance of a feature on a branch of the phylogenetic tree is the summed abundance of all species descending from that branch. By this feature-based approach, functions `evoCA`, `evoNSCA`, `evopcachord`, and `evopcahellinger` allow depicting phylogenetic dissimilarities between communities and identifying the species and the branches of the phylogenetic tree that drive these dissimilarities. The crossed double principal coordinate analysis (functions `crossdpcoa_maineffect`, `crossdpcoa_version1` and `crossdpcoa_version2`) allow to disentangle the effect of two crossed factors (e.g., space and time) on the trait-based (functional) or phylogenetic compositions of communities. Finally, the double similarity principal component analysis (function `dsimpca`) can focus either on trait-based diversity or on phylogenetic diversity. It allows the description of functional or phylogenetic similarities between communities, and the identification of species and their traits or phylogenetic positions that best characterize a community in comparison with the other communities but also in reference to its own functional or phylogenetic diversity.

Table 1.3: `adiv` functions dedicated to the partitioning of diversity at different scales

Main function	Aim	Associated functions
$\alpha$ , $\beta$ , $\gamma$ diversities and the nested apportionment of diversity		
<code>abgdivparam</code>	apportionment of parametric indices of species diversity	<code>plot.abgdivparam</code>
<code>abgevodivparam</code>	apportionment of parametric indices of phylogenetic diversity (Pavoine, 2016)	<code>plot.abgevodivparam</code>
<code>wapqe</code>	additive apportionment of quadratic entropy (Rao, 1986)	<code>rtestwapqe</code>
<code>EqRao</code>	apportionment of quadratic entropy (Rao, 1986) for unbalanced sampling designs and uneven weights for samples (Pavoine et al., 2016)	<code>rtestEqRao</code>
<code>EqRSintra</code>	apportionment of quadratic entropy (Rao, 1986) for unbalanced sampling designs and even weights for samples (Pavoine et al., 2016)	<code>rtestEqRSintra</code>
<code>EqRS</code>	apportionment of quadratic entropy (Rao, 1986) for balanced sampling designs and even weights for samples (Pavoine et al., 2016)	<code>rtestEqRS</code>
Evolutionary partitioning in phylogenetically-structured species assemblages		
<code>aptree</code>	apportionment of phylogenetic diversity within one or several communities between evolutionary periods (Pavoine et al., 2009)	<code>plot.aptree</code> , <code>rtestaptree</code> , <code>plot.rtestaptree</code> , <code>tecAptree</code>
<code>decdiv</code>	apportionment of trait-based diversity within one or several communities between phylogenetic clades (Pavoine et al., 2010)	<code>plot.decdiv</code> , <code>rtest-decdiv</code>
Evolutionary partitioning of $\alpha$ , $\beta$ , $\gamma$ diversities		
<code>abgaptree</code>	apportionment of $\alpha$ , $\beta$ , $\gamma$ diversities between evolutionary periods (Pavoine et al., 2009)	
Partitioning of $\beta$ diversity into species nestedness and species turnover		
<code>betastatjac</code>	multiple-site dissimilarity derived from Jaccard coefficient of similarity (Ricotta and Pavoine, 2015a) (presence-absence data for species)	
<code>betastatsor</code>	multiple-site dissimilarity derived from Sørensen coefficient of similarity (Ricotta and Pavoine, 2015a) (presence-absence data for species)	

Note: Functions `wapqe`, `EqRao`, `EqRS` and `EqRSintra` can handle any number of nested factors characterizing species communities. For example, they can be used to evaluate the diversity within plots ( $\alpha$  diversity), the diversity among plots but within regions (a first level of  $\beta$  diversity), the diversity among regions within a continent (a second level of  $\beta$  diversity) and the diversity within the continent ( $\gamma$  diversity).



Table 1.4: `adiv` functions for ordination approaches

Main function	Aim	Associated functions
<code>crossdpcoa_maineffect</code>	the crossed double principal coordinate analysis typically analyzes the phylogenetic or functional compositions of communities according to two factors affecting the communities (e.g. space and time; habitat and region) (Pavoine et al., 2013)	<code>crossdpcoa_version1</code> , <code>crossdpcoa_version2</code>
<code>dspca</code>	double similarity principal component analysis (DSPCA) (Pavoine, 2019): an ordination approach to analyse functional or phylogenetic similarities between communities	<code>plot.dspca</code>
<code>evoCA</code>	evolutionary (phylogenetic) correspondence analysis (Pavoine, 2016)	<code>plot.evoCA</code>
<code>evoNSCA</code>	evolutionary non-symmetric correspondence analysis (Pavoine, 2016)	
<code>evopcachord</code>	evolutionary principal component analysis based on the chord distance (Pavoine, 2016)	
<code>evopcahellinger</code>	evolutionary principal component analysis based on Hellinger distance (Pavoine, 2016)	
<code>rlqESLTP</code>	linking patterns in phylogeny, traits, abiotic variables and space (Pavoine et al., 2013)	<code>plot.rlqESLTP</code> , <code>summary.rlqESLTP</code>

Note: Function `rlqESLTP` integrates, altogether, trait-based information and phylogenetic data to describe species, and spatial data and environmental data to describe sampled sites.

## 1.6 Originality, distinctiveness, uniqueness and redundancy

In complement, `adiv` contains functions to evaluate the contribution each species has to trait-based diversity or to phylogenetic diversity. These contributions are reflected by originality indices (also named uniqueness, isolation, distinctness or distinctiveness indices) (see functions `distinctDis`, `distinctTopo`, `distinctTree`, `distinctUltra`, Table 1.5). If the species of a community are trait-based or phylogenetically redundant, then the trait-based or phylogenetic diversity of the community is expected to be much lower than species diversity. Species diversity here represents a scenario where species would be maximally dissimilar. The ratio of trait-based (or phylogenetic) diversity over species diversity thus represents the trait-based (or phylogenetic) uniqueness ( $U_\alpha$ ) in the community and its complement ( $R_\alpha = 1 - U_\alpha$ ) represents the trait-based (or phylogenetic) redundancy in the community (Ricotta et al., 2016, 2018). `adiv` contains such indices of the trait-based (or phylogenetic) uniqueness of a community (functions `uniqueness`, `treeUniqueness`, and `evouniparam`, Table 1.5). Similarly, if two communities have different species but if each species in the first community has a sibling species with close trait values or a close phylogenetic position, then the trait-based (or phylogenetic)  $\beta$  diversity between the two community is expected to be much lower than the species-based  $\beta$  diversity. Species-based  $\beta$  diversity here represents a scenario where species are assumed to be maximally dissimilar. The ratio of trait-based (or phylogenetic)  $\beta$  diversity over species-based  $\beta$  diversity thus could represent the trait-based (or phylogenetic)  $\beta$  uniqueness ( $U_\beta$ ) between communities and its complement ( $R_\beta = 1 - U_\beta$ ) would represent trait-based (or phylogenetic)  $\beta$  redundancy provided the trait-based diversity measure always is lower than or equal to the species-based  $\beta$  diversity (Ricotta et al., 2020). `adiv` contains such indices of the  $\beta$  trait-based (or phylogenetic) uniqueness across two communities (functions `betaUniqueness` and `betaTreeUniqueness`, Table 1.5).

## 1.7 Phylogenetic signal

Some other functions perform tests for phylogenetic signal in traits (Table 1.6): the fact that closely-related species in a phylogenetic tree tend to share similar or close trait values whereas distantly-related species tend to have different trait values. These functions complement those available in other packages

Table 1.5: **adiv** functions of originality, uniqueness and redundancy.

Main function	Aim	Associated functions
Originality (=distinctiveness) of each species at a given place		
<b>distinctDis</b>	dissimilarity-based species' originality ( <a href="#">Eiswerth and Haney, 1992</a> ; <a href="#">Ricotta, 2004</a> ; <a href="#">Schmera et al., 2009</a> ; <a href="#">Pavoine et al., 2017</a> )	
<b>distinctTopo</b>	species' originality calculated from the topology of a (phylogenetic) tree with species as tips ( <a href="#">May, 1990</a> ; <a href="#">Vane-Wright et al., 1991</a> ; <a href="#">Pavoine et al., 2008</a> )	
<b>distinctTree</b>	species' originality that rely on the structure and branch lengths of (phylogenetic) trees ( <a href="#">Redding, 2003</a> ; <a href="#">Redding and Mooers, 2006</a> ; <a href="#">Isaac et al., 2007</a> )	
<b>distinctUltra</b>	species' originality that rely on the structure and branch lengths of ultrametric (phylogenetic) trees ( <a href="#">Pavoine et al., 2005</a> ; <a href="#">Pavoine and Izsak, 2014</a> )	<b>orisaved</b>
Abundance-weighted originality (=distinctiveness) of each species at a given place		
<b>distinctAb</b>	dissimilarity- and abundance-based species' originality ( <a href="#">Pavoine and Ricotta, 2021</a> )	
Uniqueness versus redundancy in the whole community*		
<b>uniqueness</b>	community-level functional uniqueness and redundancy ( <a href="#">Ricotta et al., 2016</a> )	
<b>treeUniqueness</b>	community-level phylogenetic (or tree-based) redundancy taking into account the branching pattern of the underlying phylogenetic tree (or any other tree, like a functional dendrogram) ( <a href="#">Ricotta et al., 2018</a> ; <a href="#">Pavoine and Ricotta, 2019a</a> )	
<b>betaUniqueness</b>	<a href="#">Ricotta et al. (2021)</a> plot-to-plot functional $\beta$ uniqueness	
<b>betaTreeUniqueness</b>	<a href="#">Ricotta et al. (2020)</a> plot-to-plot functional or phylogenetic $\beta$ uniqueness (index named $U_F$ for functional data and $U_P$ for phylogenetic data in <a href="#">Ricotta et al., 2020</a> )	
<b>evouniparam</b>	parametric diversity indices applied to phylogenetic uniqueness ( <a href="#">Pavoine and Ricotta, 2019a</a> )	<b>plot.evouniparam</b>

\* compared to a scenario where species would be maximally dissimilar.



Table 1.6: Tests for phylogenetic signal and graphical display of traits in front of the phylogeny

Main function	Aim	Associated functions
<code>apd</code>	Hardy (2008)'s test for phylogenetic structure in species abundance distribution	
<code>rtestdecdiv</code>	tests, for one community (with presence/absence or abundance data), if a representation of trait diversity on a phylogenetic tree highlights a nonrandom pattern (Pavoine et al., 2010)	<code>decdiv</code> , <code>plot.decdiv</code>
<code>K</code> and <code>Kstar</code>	Blomberg et al. (2003) $K$ and $K^*$ statistics and permutation test	
<code>Kw</code>	Pavoine and Ricotta (2013) statistic $K_w$ and permutation test	
<code>barp4d*</code>	barplots of trait values associated with the tips of a phylogenetic tree	
<code>dotp4d*</code>	dotplots of trait values associated with the tips of a phylogenetic tree	
<code>gridp4d*</code>	gridplots of trait values associated with the tips of a phylogenetic tree	
<code>plot.phylo4d*</code>	general interface for functions <code>barp4d</code> , <code>dotp4d</code> , and <code>gridp4d</code>	

\* As written in the help file of the functions `plot.phylo4d`, `barp4d`, `dotp4d`, `gridp4d`, "the four functions were written by Francois Keck in the package named `phylosignal`. Functions were there named as follows: `multiplot.phylo4d`, `barplot.phylo4d`, `dotplot.phylo4d`, and `gridplot.phylo4d`. At the end of 2019, the package was orphaned and the functions were integrated in package `adiv`. The versions of the functions have been slightly modified compared to those developed by Francois Keck."

like `picante` (Kembel et al., 2010). Among the functions available in `adiv`, function `rtestdecdiv` can handle multiple numbers and multiple types of traits (e.g., nominal, quantitative, ordinal, fuzzy, circular). Functions `K`, `Kstar` and `Kw` focus on quantitative traits and are all related to Blomberg et al. (2003) statistics. As written in the help file of these three functions, "Blomberg et al. (2003) introduced two statistics of phylogenetic signal:  $K = MSE0/MSE$  and  $K^* = MSE^*/MSE$ , where  $MSE$  is the mean squared error of the trait values calculated using the variance-covariance matrix derived from the phylogenetic tree,  $MSE0$  is the mean squared error of the tip trait values, measured from a phylogenetically correct mean of tip trait values and  $MSE^*$  is the mean squared error of the tip trait values, measured from the estimate of the mean of the raw tip trait values. In both statistics  $K$  and  $K^*$ , the value of  $MSE$  will be relatively small if the phylogenetic tree accurately describes the variance-covariance pattern observed in the data, leading to high values for  $K$  and  $K^*$  (meaning high phylogenetic signal). In functions `K` and `Kstar`,  $K$  and  $K^*$ , respectively, are divided (normalized) by their expected value if the trait evolved under a Brownian motion along the branches of the phylogenetic tree (this expected value is invariant under permutation of trait values among the tips of the phylogeny).

To test for phylogenetic signal, Blomberg et al. (2003) actually considered neither  $K$  nor  $K^*$  but  $MSE$  as the core statistic associated with random permutations of trait values among tips of the phylogenetic tree. Although the literature on phylogenetic signal has currently mostly ignored  $K^*$  focusing on statistic  $K$ ,  $K^*$  could thus actually have been considered as the core statistic of Blomberg et al. (2003) test for phylogenetic signal. Indeed, as  $MSE^*$  is independent of permutations of trait values among the tips of the phylogeny while  $MSE0$  is, Blomberg et al. (2003) approach corresponds to considering  $K^*$  and not  $K$  as the statistic of the test for phylogenetic signal in traits. This test is also equivalent to an alternative implemented via phylogenetically independent contrasts also proposed by Blomberg et al. (2003).

Function `Kw` implements index  $K_w$ , a modified version of  $K^*$  that grants a higher importance in the calculation of phylogenetic signal to the tips that have many closely related tips (Pavoine and Ricotta, 2013).

In functions, `K`, `Kstar` and `Kw`, I considered the same permutation scheme as in Blomberg et al. (2003) but used  $K$ ,  $K^*$  and  $K_w$ , as the core statistic, respectively. The test developed by Blomberg et al. (2003) thus corresponds to function `Kstar`."

---

## 1.8 Graphics

Package `adiv` contains a few graphical functions to:

1. display profiles of parametric species or phylogenetic diversity, evenness, and uniqueness indices (see functions `plot.divparam`, `plot.eveparam`, `plot.evodivparam`, `plot.evoeveparam`, `plotevouni-param` in Table 1.1, and function `plot.abgdivparam`, `plot.abgevodivparam` in Table 1.3);
2. display results of ordination methods (functions `plot.dspca`, `plot.evoCA`, `plot.rlqESLTP` in Table 1.4);
3. display data in front of a phylogenetic tree (functions `plot.phylo4d`, `barp4d`, `dotp4d`, `gridp4d` in Table 1.6) or at the nodes of a phylogenetic tree (function `decdiv`, Tables 1.6 and 1.3);
4. display an ultrametric phylogenetic tree with vertical lines at each speciation event (limits of evolutionary periods, function `plot.aptree` in Table 1.3);
5. display results of permutation tests (functions `plot.rtestaptree` in Table 1.3);
6. display three key components of the phylogenetic dissimilarities between communities on a ternary plot (function `evodiss_ternaryplot`, Table 1.2).

The help files of `adiv` functions also contain suggestions and examples for the use of other packages to have more options in order to display their results.

## 2 Applications

Install package `adiv`:

```
> install.packages("adiv")
```

Then, load the package:

```
> library(adiv)
```

For the illustrations below, you will also be required to install other packages:

```
> install.packages("ape")
> install.packages("phylobase")
> install.packages("ade4")
```

Load the packages:

```
> library(ape)
> library(phylobase)
> library(ade4)
```

These packages will allow us to read, manipulate and display phylogenetic trees (packages `ape` and `phylobase`, Paradis et al., 2004; R Hackathon et al., 2019), estimate the contribution of each node of the phylogenetic tree to the factorial map of the evolutionary correspondence analysis (package `ade4`, Thioulouse et al., 2018).

Load the data set named `batcomm` on bat communities. It is a list that contains bat phylogenies (below we use `batcomm$tre`) and bat abundances (below we use `batcomm$ab`) along a disturbance gradient in a neotropical rainforest. The abundance data were collected by [Medellin et al. \(2000\)](#) in the Selva Lacandona of Chiapas, Mexico, in four habitats: rainforest (F), cacao plantations (P), oldfields (O) and cornfields (C). The phylogeny we use below is [Fritz et al. \(2009\)](#) phylogeny pruned to retain only the species present in Medellin et al. data set.

```
> data(batcomm)

> batab <- batcomm$ab
> rownames(batab)

[1] "F" "P" "O" "C"

> rownames(batab) <- c("rainforest", "cacao plantations", "oldfields", "cornfields")
```

The following instruction can be used to obtain a species diversity profile in each habitat type thanks to Hill numbers ([Hill, 1973](#))

```
> plot(divparam(batab, q=seq(0, 3, le=50)), pch=1:4, col=1:4,
+       axisLABEL = "Species diversity")
```

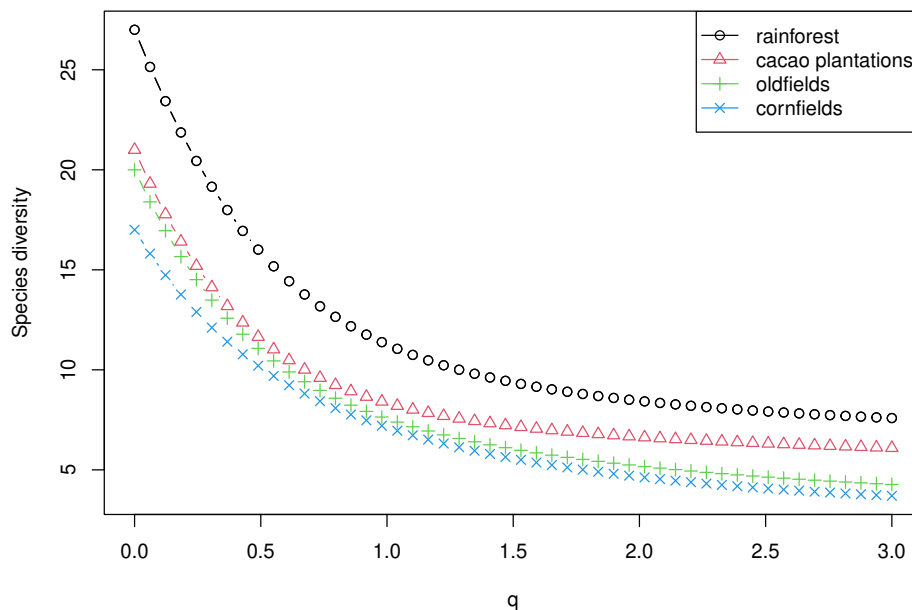


Figure 2.1: Species diversity profiles

Species diversity profiles (Figure 2.1) show that, when abundance data are discarded (when parameter  $q$  of the Hill numbers is equal to 0), the rainforest largely dominates in terms of the number of species, with cornfields having the lowest richness. However, when diversity measures give slightly greater importance to abundance ( $q \approx 0.5$ ), the cacao plantations and the oldfields reach diversity levels almost as low as that of the cornfields. When this importance increases again ( $q > 1$ ), the cacao plantations then recover to medium-level biodiversity, while that of oldfields remains at the same low level as cornfields. Our interpretation of the impact of environmental disturbance on species diversity may thus depend on the way abundance data are used in diversity indices.

The following associated instruction provides species evenness profiles, using the ratio of Hill numbers to their maximum possible value (equal to the number of species in each habitat):

```
> plot(eveparam(batab, q=seq(0, 3, le=50)), pch=1:4, col=1:4,
+      axisLABEL = "Species evenness")
```

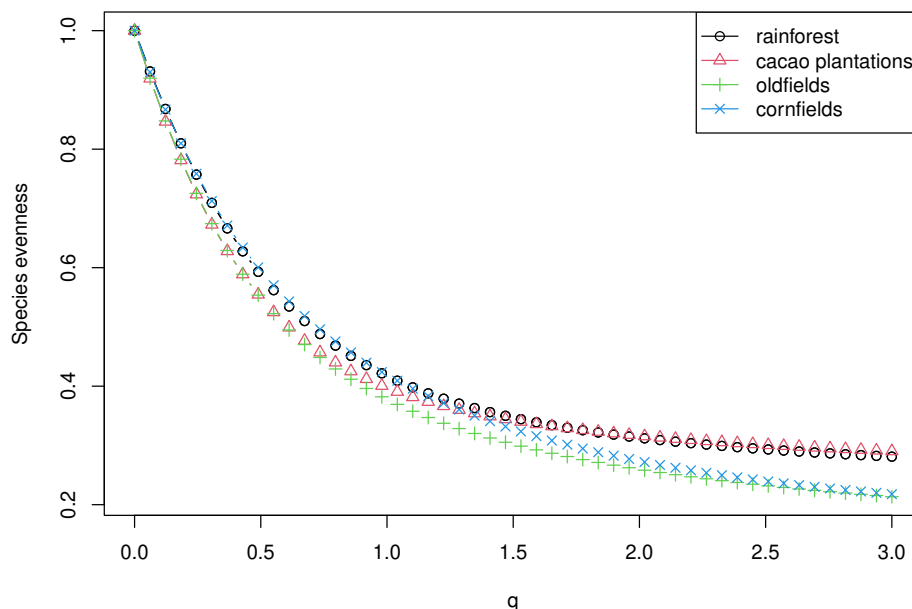


Figure 2.2: Species evenness profiles

Parametric evenness indices indicate how evenly weighted species are in diversity measurement. Evenness is maximum ( $=1$ ) when their parameter ( $q$ )  $= 0$  because species are given equal weight in the diversity index: independently of abundance, species weights are even, and, with Hill numbers, diversity is equal to the number of species. When  $q$  increases, abundant species are overweighted compared to rare species, we thus expect that the evenness in species weights decreases. With Hill numbers, when  $q$  tends to infinity, evenness varies between  $1/n$  ( $n$ =number of species) if a species' relative abundance approaches unity and 1 if species have even abundances. Applying Hill numbers to bat communities (Figure 2.2), we can observe that the rate of decrease in species evenness, with parameter  $q$  varies according to the habitat considered. With Hill numbers, such differences could, in theory, be due to the fact that habitats differ in their level of species richness. However, this is not the case here as the rainforest with the highest species richness has one of the lowest rate of decrease. Instead, the differences in the rate of decrease are here due to the shape of abundance distribution where a single species dominates in abundance in oldfields and cornfields but not in rainforest and cacao plantations as shown by the dotplots of species abundances below (Figures 2.3, 2.4, 2.5, 2.6).

```
> dotchart(sort(unlist(batcomm$ab[1,])), font=3, main="Rainforest", xlab="Abundance")
```

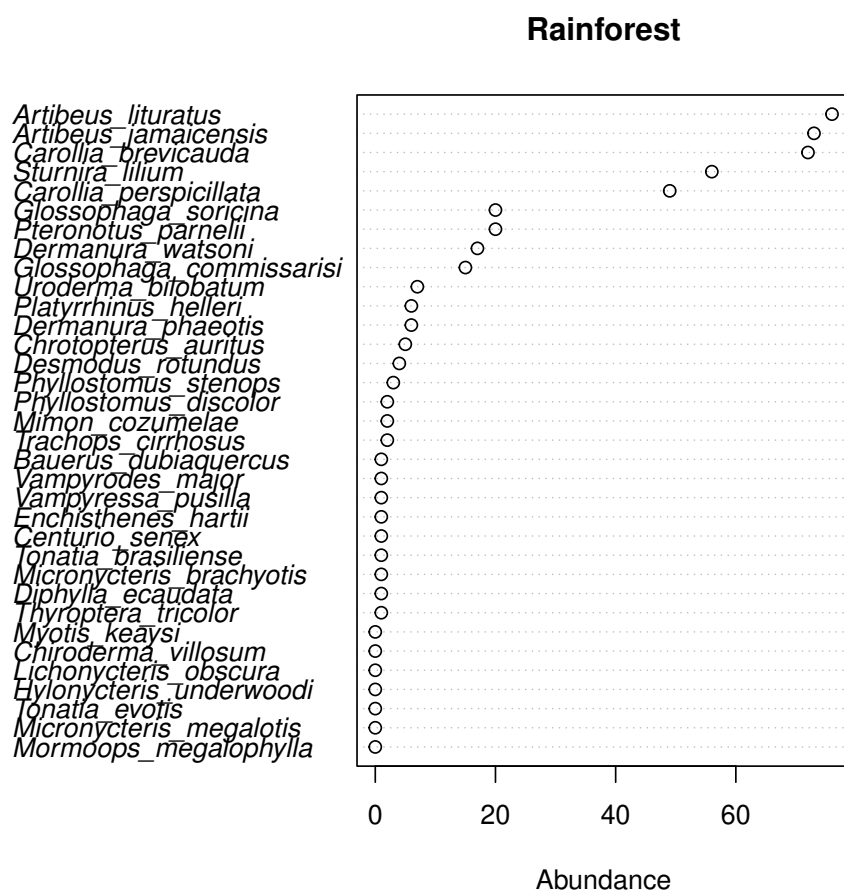


Figure 2.3: Distribution of species abundances in the rainforest

---

```
> dotchart(sort(unlist(batcomm$ab[2,])), font=3, main="Cacao plantations", xlab="Abundance")
```

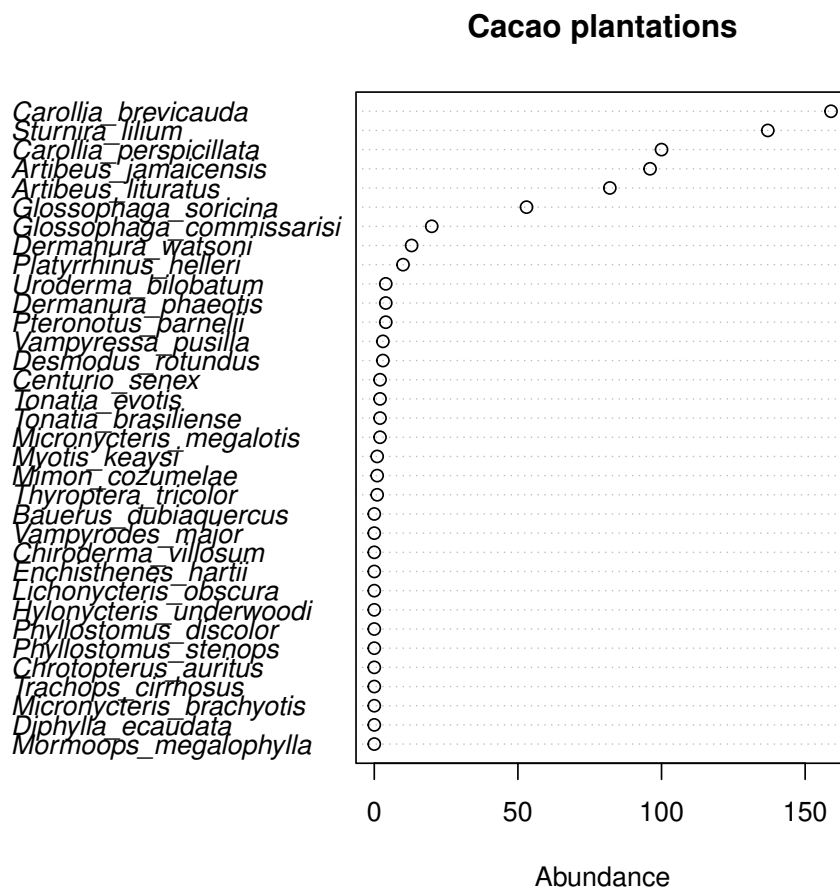


Figure 2.4: Distribution of species abundances in cacao plantations

```
> dotchart(sort(unlist(batcomm$ab[3,])), font=3, main="Oldfields", xlab="Abundance")
```

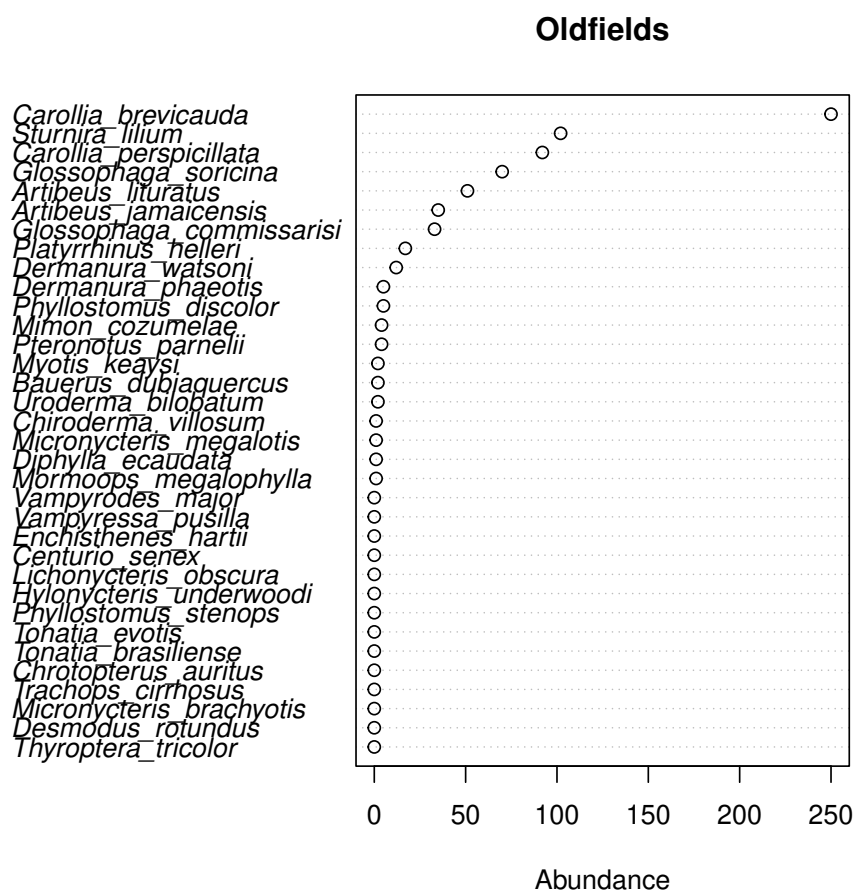


Figure 2.5: Distribution of species abundances in oldfields



```
> dotchart(sort(unlist(batcomm$ab[4,])), font=3, main="Cornfieds", xlab="Abundance")
```

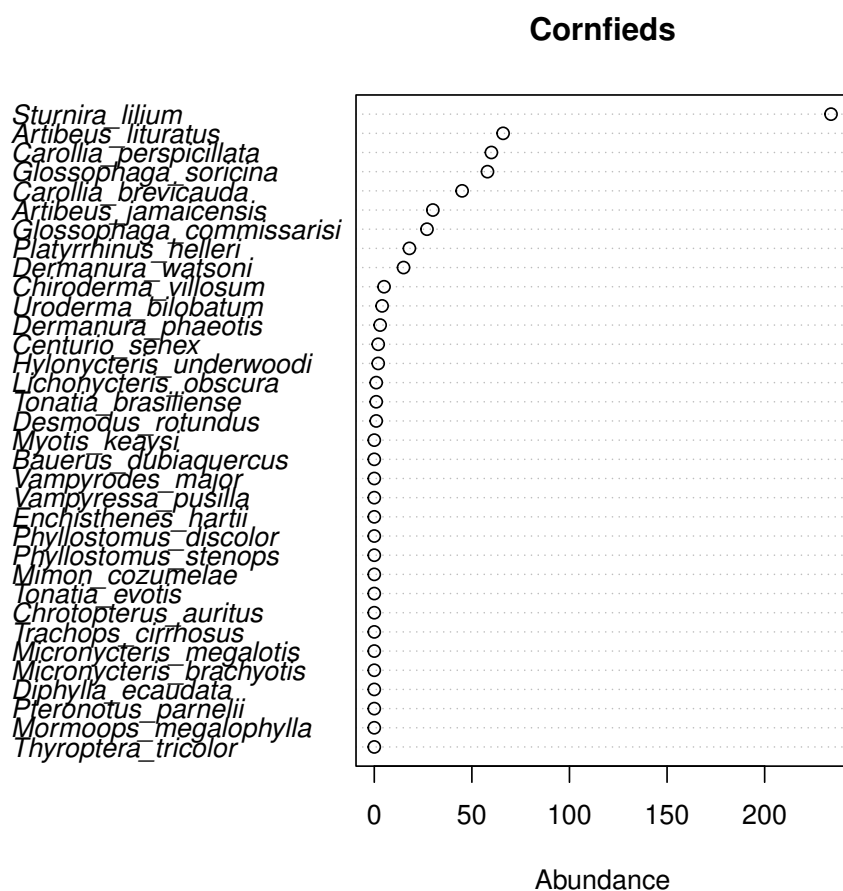


Figure 2.6: Distribution of species abundances in the cornfields

Phylogenetic data can be added as described in Figure 2.7 to obtain the phylogenetic diversity profile in each habitat type thanks to Hill numbers applied to phylogenetic features (Pavoine and Ricotta, 2019a). According to Figure 2.7, when rare species are given high weight in the measurement of phylogenetic diversity ( $q < 2$ ), the level of phylogenetic diversity decreases from the rainforest (highest level) to cacao plantations and oldfields (medium) and finally cornfields (lowest). Then all habitats reach similar levels of phylogenetic diversity.

```
> phy <- read.tree(text=batcomm$tre) # Bat phylogenetic tree
> ab <- batab[, phy$tip.label]
> # Species in the abundance table are ordered as in the phylogenetic tree
> plot(evodivparam(phy, ab, q=seq(0, 3, le=50)), pch=1:4, col=1:4)
> # Resulting diversity profiles
```

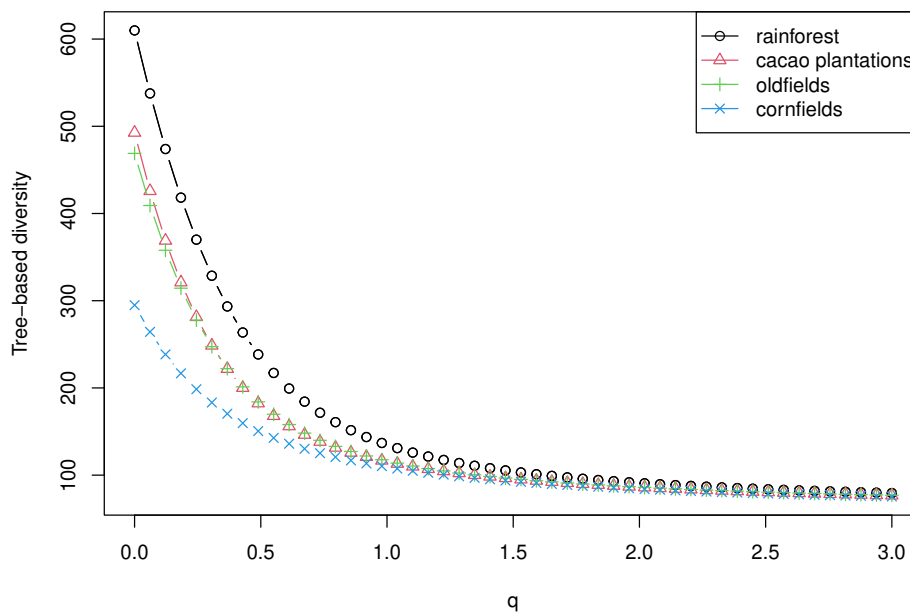


Figure 2.7: Phylogenetic diversity profiles

Non-parametric phylogenetic diversity indices could also be used. Among them, some are dedicated to "tree-based" diversity, that is to say diversity among species knowing that a hierarchical structure, such as a phylogenetic tree, links the species. Other indices are dedicated to "dissimilarity-based" diversity, that is to say diversity among species knowing that there are different degrees of dissimilarities between two species. Function `evodiv` of `adiv` contains indices of "tree-based" diversity. Functions `QE` and `Rentropy` implement indices of "dissimilarity-based" diversity as illustrated in Figure 2.8 below:

---

```

> par(mfrow=c(1,2))
> barplot(QE(ab, as.dist(cophenetic(phy)))$diversity)
> barplot(Rentropy(ab, as.dist(cophenetic(phy)))$diversity)
> par(mfrow=c(1,1))

```

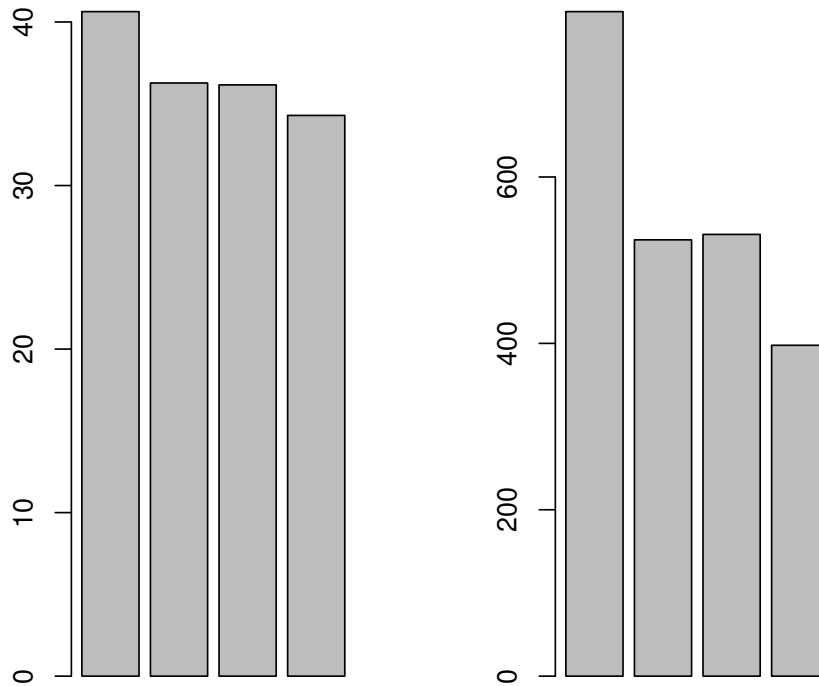


Figure 2.8: Phylogenetic diversity in habitats according to the quadratic entropy (left, [Rao, 1982](#)) and R entropy (right, [Pavoine et al., 2017](#))

Species abundances can be displayed in front of the phylogenetic tree (log-transformed abundance) as shown in [Figure 2.9](#).

```
> bat.4d <- phylo4d(phy, log(t(ab)+1))
> barp4d(bat.4d, center = FALSE, scale = FALSE, tip.cex = 0.75, data.xlim = c(0, 5))
```

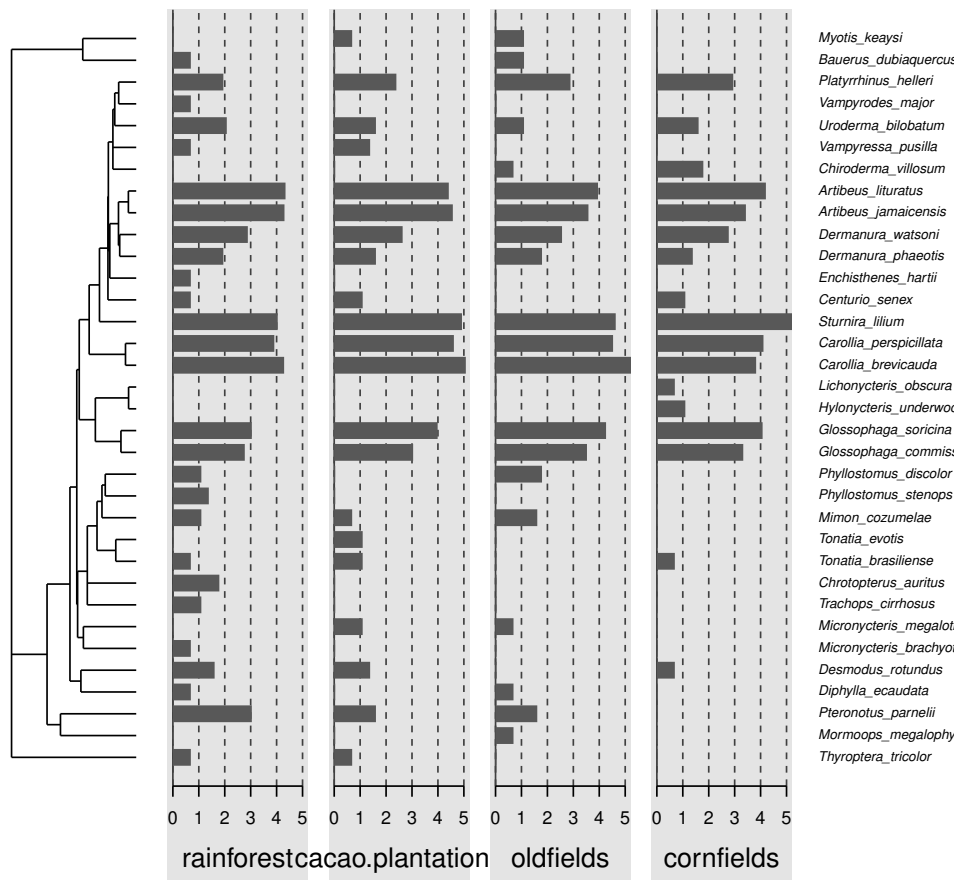


Figure 2.9: (Log-transformed) species abundances in each habitat

In the phylogenetic tree some species have more closely related sister species than others. Originality indices (also named distinctiveness or uniqueness) can be used to measure the degree of uniqueness of a species. A species is unique if it has no close relatives in the phylogenetic tree. In Figure 2.10, I used functions `distinctTree` and `distinctUltra` to measure the phylogenetic originality of each bat species and function `barp4d` to display it.

```

> ori <- cbind.data.frame(distinctTree(phy, c("ED", "ES")),
+   distinctUltra(phy, c("Qb", "2Hb")))
> bat.4d <- phylo4d(phy, ori)
> barp4d(bat.4d, center = FALSE, scale = FALSE)

```

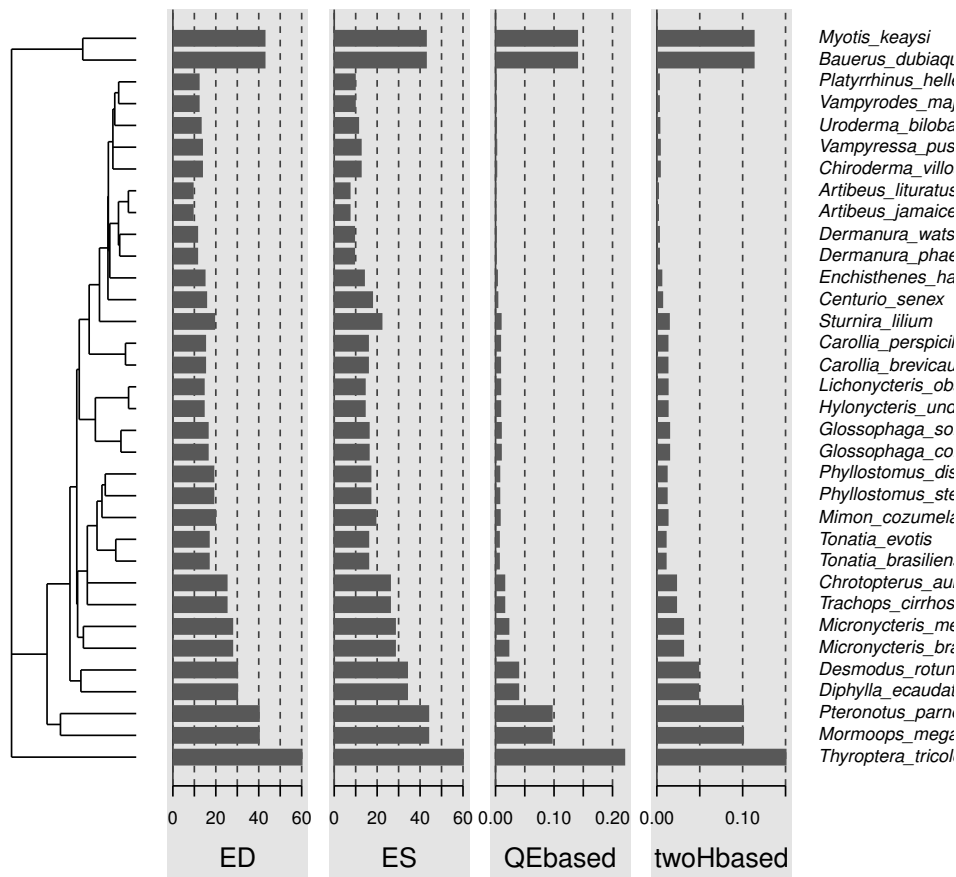


Figure 2.10: Species originalities (phylogenetic distinctiveness)

Tests can be run to evaluate if differences between the phylogenetic compositions of the habitats could have been obtained randomly:

```

> pa <- ab; pa[ab>0] <- 1 # Matrix of the presence/absence of the species in the four habitats
> dphy <- cophenetic(phy) # Matrix of phylogenetic distances between species
> # P-value of the test with abundance data:
> rtestEqRS(ab, as.dist(dphy), nrep=999)$pvalue

```

```
[1] 0.085
```

```

> # P-value of the test with presence/absence data:
> rtestEqRS(pa, as.dist(dphy), nrep=999)$pvalue # P-value of the test

```

```
[1] 0.064
```

The p-value of the test with abundance is higher than 5% and that with presence-absence data is close to 5%, suggesting that the phylogenetic differences between the four habitats are weak and could have been obtained at random.

Several functions exist in `adiv` to evaluate the amount of phylogenetic differences in the composition of several communities. For example below I use the chord distance applied to the abundance of phylogenetic features in communities:

```
> evodiss(phy, ab, "Chord")
```

	rainforest	cacao plantations	oldfields
cacao plantations	0.1186973		
oldfields	0.2412351	0.1461271	
cornfields	0.2395816	0.2073270	0.3044903

Parametric measures of  $\alpha$ ,  $\beta$ ,  $\gamma$  phylogenetic diversity can also be obtained with function `abgevodivparam` (see Figure 2.11 below).

```
> plot(abgevodivparam(phy, ab, q=seq(0, 3, le=50)))
```

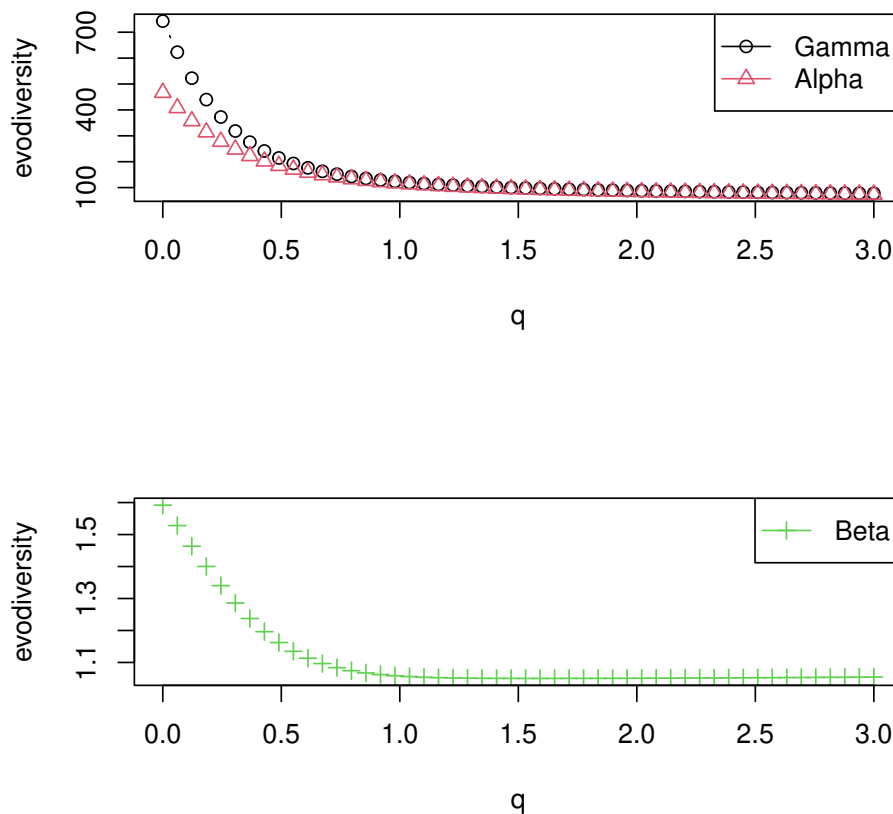


Figure 2.11:  $\alpha$ ,  $\beta$ ,  $\gamma$  components of diversity in bat communities within, between and all over the different levels of the disturbance gradient (rainforest, cacao plantations, oldfields, cornfields), respectively

Ordination analyses can also help to study the potential phylogenetic differences between the communities of the four habitats. Below is an example with the evolutionary correspondence analysis (`evoCA`, [Pavoine, 2016](#)) applied to bat species' presences-absences and to their phylogenetic tree:

```
> evoca <- evoCA(phy, pa, scannf=FALSE, nf=2) # Two axes are here retained.
> # % of inertia (eigenvalues) expressed by each axis:
> evoca$eig/sum(evoca$eig)
```

---

```
[1] 0.4155098 0.3696947 0.2147955
```

Factorial map with coordinates for the habitats (Figure 2.12):

```
> plot(evoca$li[, 1], evoca$li[, 2], xlab = "Axis1", ylab="Axis2", col="red", pch=19, asp=1)
> abline(h=0, col="grey"); abline(v=0, col="grey")
> text(evoca$li[, 1], evoca$li[, 2], rownames(evoca$li), pos = c(2,2,4,2), offset=0.2)
```

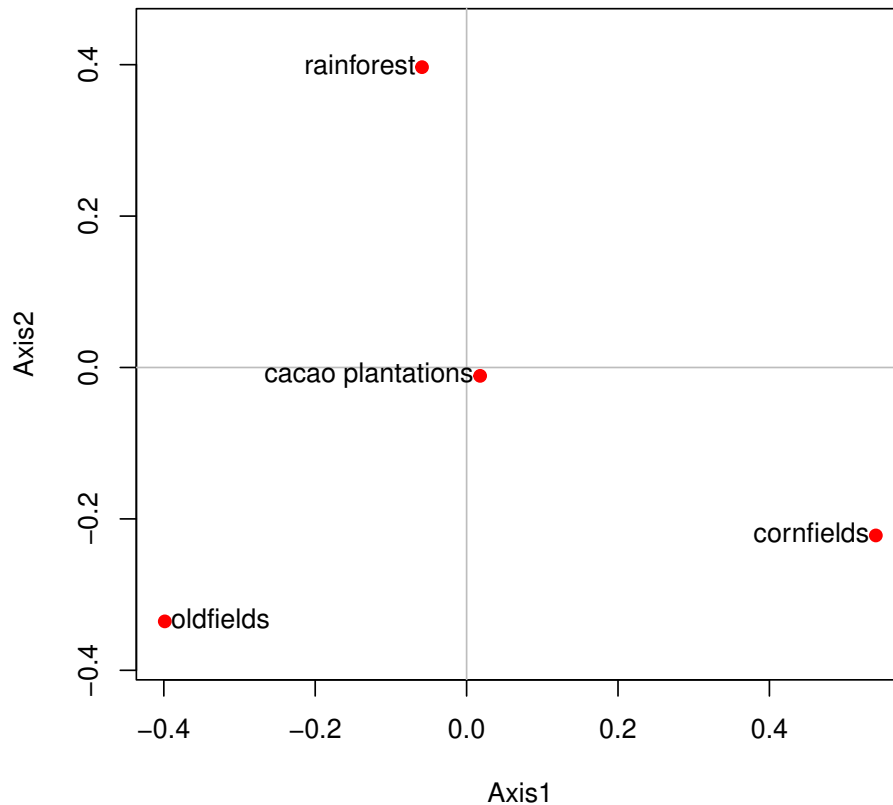


Figure 2.12: Scores for habitats on the first two axes of evoCA

Factorial map with coordinates for the nodes of the phylogenetic tree (only the labels of the nodes contributing to more than 5% of the inertia of one of the axes are displayed) (Figure 2.13):

```
> kept <- rownames(inertia.dudi(evoca, col=TRUE)$col.abs)[apply(inertia.dudi(evoca,
+   col=TRUE)$col.abs, 1, max)>5]
> keptc <- kept[-10] # graphical improvements to avoid the labels superimpose
> keptc[9] <- paste(kept[9:10], collapse="\n")
```



---

```

> plot(evoca$co[, 1], evoca$co[, 2], xlab = "Axis1", ylab="Axis2", col="grey", asp = 1)
> abline(h=0, col="grey"); abline(v=0, col="grey")
> points(evoca$co[kept, 1], evoca$co[kept, 2], col = "red", pch = 19)
> positions <- c(2,2,2,4,1,2,4,3,1,3)
> text(evoca$co[kept[-10], 1], evoca$co[kept[-10], 2], gsub("_", " ", keptc),
+      pos = positions, offset = 0.75, font=3)

```

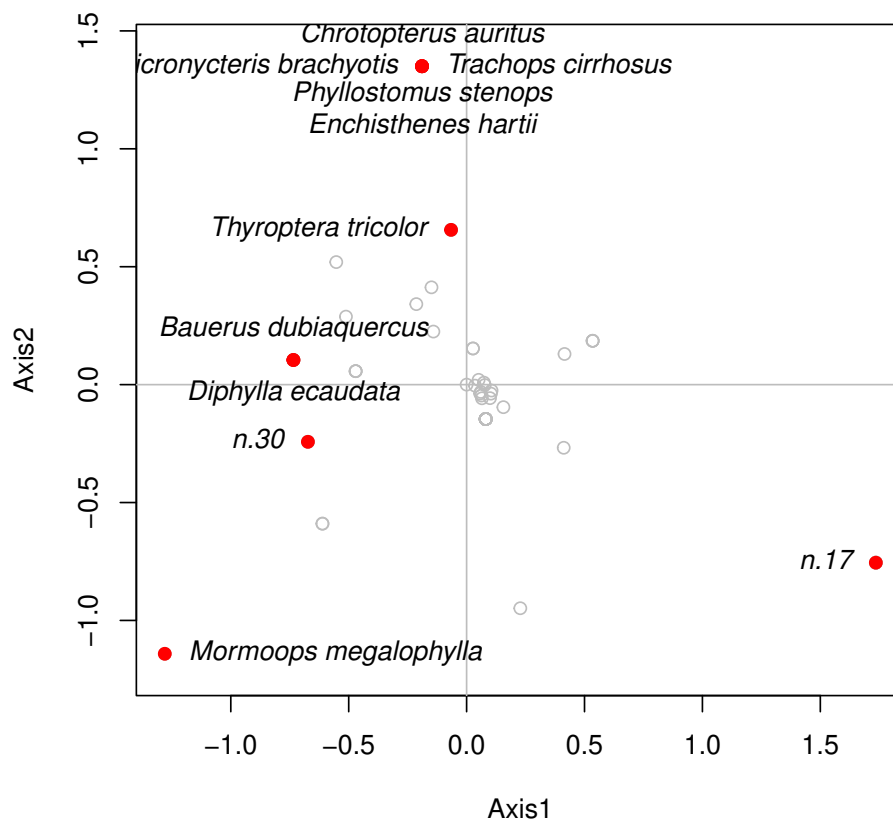


Figure 2.13: Scores for species (tips) and nodes (of the phylogenetic tree) on the first two axes of evoCA

Note that several functions exist in R to display factorial maps with optimized positions for labels (see e.g. package `adegraphics`, function `adegpar` argument `plabels$optim`, Siberchicot et al., 2017). Those functions are used or suggested in package `adiv` for graphical display. However, with presence/absence data, in phylogenetic or even simple correspondence analysis the number of labels that superimpose may be too high hampering these functions to be efficient, because several species may have identical patterns of occurrence across the sampled sites. For example, here species *Chrotopterus auritus*, *Enchisthenes hartii*, *Micronycteris brachyotis*, *Phyllostomus stenops*, and *Trachops cirrhosus* were only observed in the rainforest. Manual adjustments of labels thanks to basic R scripts (as done in the R scripts above), in that case, allow optimum visualization.

Codes for nodes on the phylogenetic tree can be obtained by the following instructions (Figure 2.14):

```

> # phylogenetic tree with some of the tip labels colored
> tipcolors <- rep(1, 34)
> tipcolors[phy$tip.label %in% kept] <- "red"
> tipcolors[phy$tip.label == "Bauerus dubiaquercus"] <- "violet"
> tipcolors[phy$tip.label %in% c("Myotis keaysi", "Hylonycteris underwoodi", "Lichonycteris obscura")] <- "blue"
> plot(phy, show.node.label=TRUE, tip.color=tipcolors, cex=0.75)

```

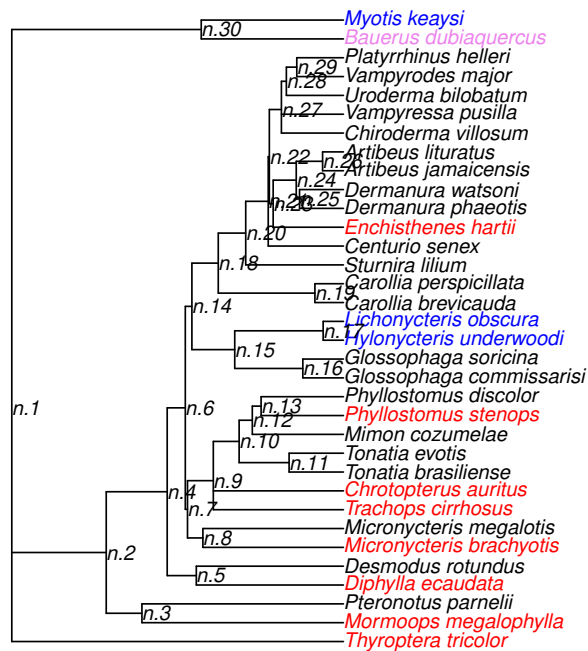


Figure 2.14: Phylogenetic tree with species names colored according to their contribution to the first two axes of evoCA; top-contributing species (red), species descending from top-contributing nodes (blue), top-contributing species that descend from top-contributing nodes (violet).

Species coordinates on the evoCA axes can also be put in front of the phylogenetic tree using the following instruction (Figure 2.15):

```
> bat.4d <- phylo4d(phy, evoca$co[phy$tip.label, ])
> dotp4d(bat.4d, center = FALSE, scale = FALSE)
```

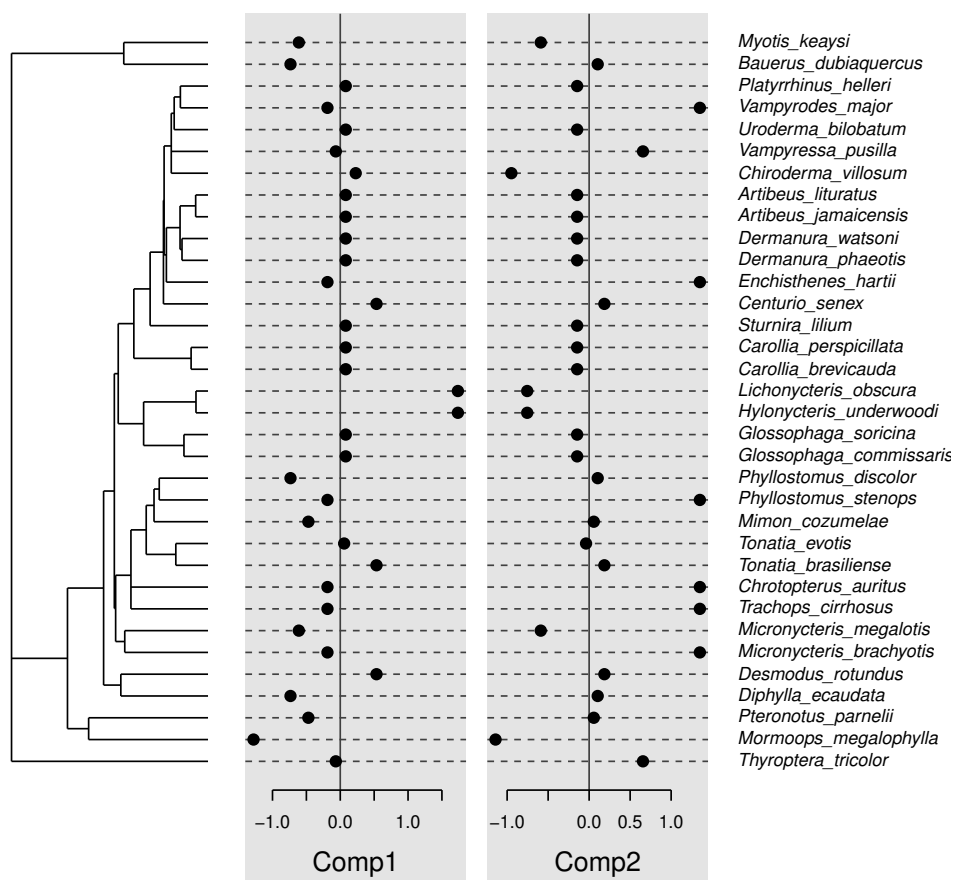


Figure 2.15: Species scores on the first two axes of evoCA (Comp1 and Comp2) in front of the phylogenetic tree

A dynamic 3d plot with the phylogenetic tree plotted on the factorial map can be obtained as follows (Figure 2.16):

```
> plot(evoca)
```

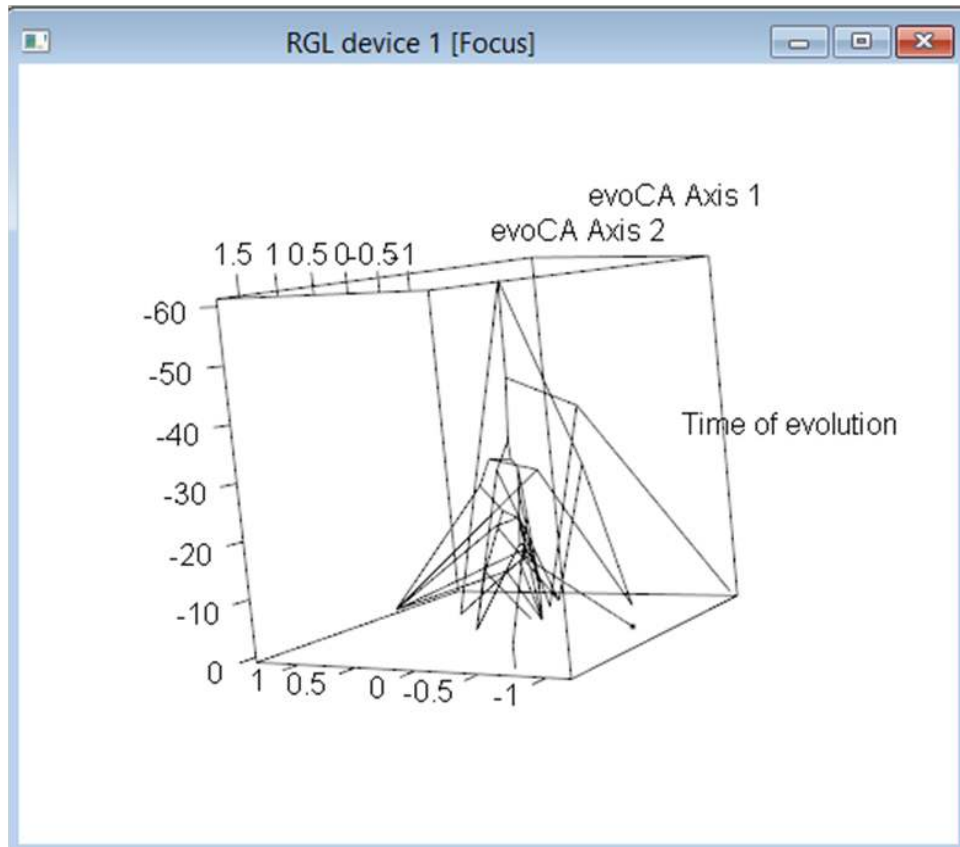


Figure 2.16: Three-dimensional factorial map where species and nodes have coordinates on the first two axes of evoCA and the third axis represents time of evolution and allows the phylogenetic tree to be visualized. The clearer the image of the phylogenetic tree, the stronger the phylogenetic differences between the compared habitats.

All these displays of the results of evoCA applied to bat communities show low phylogenetic differences between the habitats: differences in the species compositions of the habitats and in the abundances of these species imply species that are rather closely related on the phylogenetic tree.

## References

- Blomberg, S.P., Garland, T., Ives, A.R. (2003). Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution*, 57, 717–745. <https://doi.org/10.1111/j.0014-3820.2003.tb00285.x>
- Chao, A., Chiu, C.-H., Jost, L. (2010). Phylogenetic diversity measures based on Hill numbers. *Philosophical Transactions of the Royal Society London Series B*, 365, 3599–3609. <https://doi.org/10.1098/rstb.2010.0272>
- Chiu, C.-H., Jost, L., Chao, A. (2014). Phylogenetic beta diversity, similarity, and differentiation measures based on Hill numbers. *Ecological Monographs*, 84, 21–44. <https://doi.org/10.1890/12-0960.1>
- Daroczy, Z. (1970). Generalized information functions. *Information and Control*, 16, 36–51. [https://doi.org/10.1016/S0019-9958\(70\)80040-7](https://doi.org/10.1016/S0019-9958(70)80040-7)
- Doledec, S., Chessel, D., Ter Braak, C.J.F., Champely, S. (1996) Matching species traits to environmental variables: a new three-table ordination method. *Environmental and Ecological Statistics*, 3, 143–166. <https://doi.org/10.1007/BF02427859>

- 
- Eiswerth, M.E., Haney, J.C. (1992). Allocating conservation expenditures: accounting for inter-species genetic distinctiveness. *Ecological Economics*, 5, 235–249. [https://doi.org/10.1016/0921-8009\(92\)90003-B](https://doi.org/10.1016/0921-8009(92)90003-B)
- Faith, D.P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation*, 61, 1–10. [https://doi.org/10.1016/0006-3207\(92\)91201-3](https://doi.org/10.1016/0006-3207(92)91201-3)
- Fritz, S.A., Bininda-Emonds, O.R.P., Purvis, A. (2009). Geographic variation in predictors of mammalian extinction risk: big is bad, but only in the tropics. *Ecology Letters*, 12, 538–549. <https://doi.org/10.1111/j.1461-0248.2009.01307.x>
- Gini, C. (1912). *Variabilita e mutabilita*. Studi economicoaguridici delle facoltta di giurizprudenza dell, Universite di Cagliari III, Parte II.
- Gower, J.C., Legendre, P. (1986). Metric and Euclidean properties of dissimilarity coefficients. *Journal of Classification*, 3, 5–48.
- Gregorius, H.-R., Gillet, E.M., Ziehe, M. (2003) Measuring differences of trait distributions between populations. *Biometrical Journal*, 8, 959–973 <https://doi.org/10.1002/bimj.200390063>
- Hardy, O.J. (2008). Testing the spatial phylogenetic structure of local communities: statistical performances of different null models and test statistics on a locally neutral community. *Journal of Ecology*, 96, 914–926. <https://doi.org/10.1111/j.1365-2745.2008.01421.x>
- Havrda, M., Charvat, F. (1967). Quantification method of classification processes: concept of structural alpha-entropy. *Kybernetika*, 3, 30–35.
- Hill, M.O. (1973). Diversity and evenness: a unifying notation and its consequences. *Ecology*, 54, 427–432. <https://doi.org/10.2307/1934352>
- Heip, C. (1974). A new index measuring evenness. *Journal of the Marine Biological Association UK*, 54, 555–557. <https://doi.org/10.1017/S0025315400022736>
- Isaac, N.J., Turvey, S.T., Collen, B., Waterman, C., Baillie, J.E. (2007). Mammals on the EDGE: conservation priorities based on threat and phylogeny. *PloS ONE*, 2, e296. <https://doi.org/10.1371/journal.pone.0000296>
- Jombart, T., Dray, S. (2008). adephylo: exploratory analyses for the phylogenetic comparative method. *Bioinformatics*, 26, 1907–1909. <https://doi.org/10.1093/bioinformatics/btq292>
- Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., Blomberg, S.P., Webb, C.O. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, 26, 1463–1464. <https://doi.org/10.1093/bioinformatics/btq166>
- Kosman, E. (1996) Difference and diversity of plant pathogen populations: a new approach for measuring. *Phytopathology*, 86, 1152–1155.
- Kroonenberg, P.M., Lombardo, R. (1999). Nonsymmetric correspondence analysis: a tool for analysing contingency tables with a dependence structure. *Multivariate Behavioral Research*, 34, 367–396. [https://doi.org/10.1207/S15327906MBR3403\\_4](https://doi.org/10.1207/S15327906MBR3403_4)
- Laliberte, E., Legendre, P., Shipley, B. (2014). *FD: measuring functional diversity from multiple traits, and other tools for functional ecology*. R package version 1.0-12. <https://CRAN.R-project.org/package=FD>
- Lauro, N., D’Ambra, L. (1984). Non-symmetrical correspondence analysis. In: Tomassone, R. (Ed.), *Data Analysis and Informatics*, III, pp. 433–446. North-Holland, Amsterdam: Elsevier.
- Legendre, P., Legendre, L. (1998). *Numerical Ecology*. Amsterdam: Elsevier.
- Magurran, A.E. (2004). *Measuring biological diversity*. Oxford, U.K.: Blackwell Publishing.
- Margalef, R. (1972). Homage to Evelyn Hutchinson, or why is there an upper limit to diversity? *Transactions of the Connecticut Academy of Arts and Sciences*, 44, 211–235.

- 
- May, R.M. (1990). Taxonomy as destiny. *Nature*, 347, 129–130. <https://doi.org/10.1038/347129a0>
- McIntosh, R.P. (1967). An index of diversity and the relation of certain concepts to diversity. *Ecology*, 48, 392–404. <https://doi.org/10.2307/1932674>
- Medellin, R., Equihua M., Amin, M.A. (2000). Bat diversity and abundance as indicators of disturbance in Neotropical rainforest. *Conservation Biology*, 14, 1666–1675. <https://doi.org/10.1111/j.1523-1739.2000.99068.x>
- Menhinick, E.F. (1964). A comparison of some species-individuals diversity indices applied to samples of field insects. *Ecology*, 45, 859–861. <https://doi.org/10.2307/1934933>
- Nipperess, D.A., Faith, D.P., Barton, K. (2010). Resemblance in phylogenetic diversity among ecological assemblages. *Journal of Vegetation Science*, 21, 809–820. <https://doi.org/10.1111/j.1654-1103.2010.01192.x>
- Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P. McGlinn, D., ..., Wagner, H. (2019). *vegan: Community Ecology Package*. R package version 2.5-5. <https://CRAN.R-project.org/package=vegan>
- Orloci, L. (1978). *Multivariate Analysis in Vegetation Research*. The Hague: Junk.
- Paradis, E., Claude, J., Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*, 20, 289–290. <https://doi.org/10.1093/bioinformatics/btg412>
- Pavoine, S. (2016) A guide through a family of phylogenetic dissimilarity measures among sites. *Oikos*, 125, 1719–1732. <https://doi.org/10.1111/oik.03262>
- Pavoine, S. (2019). An ordination approach to explore similarities among communities. *Journal of Theoretical Biology*, 462, 85–96. <https://doi.org/10.1016/j.jtbi.2018.11.002>
- Pavoine, S., Baguette, M., Bonsall, M.B. (2010). Decomposition of trait diversity among the nodes of a phylogenetic tree. *Ecological Monographs*, 80, 485–507.
- Pavoine, S., Blondel, J., Dufour, A.-B., Gasc, A., Bonsall, M. B. (2013). A new technique for analysing interacting factors affecting biodiversity patterns: crossed-DPCoA. *PloS One*, 8, e54530. <https://doi.org/10.1371/journal.pone.0054530>
- Pavoine, S., Bonsall, M.B., Dupaix, A., Jacob, U., Ricotta, C. (2017). From phylogenetic to functional originality: guide through indices and new developments. *Ecological Indicators*, 82, 196–205. <https://doi.org/10.1016/j.ecolind.2017.06.056>
- Pavoine, S., Izsak, J. (2014). New biodiversity measure that includes consistent interspecific and intraspecific components. *Methods in Ecology and Evolution*, 5, 165–172. <https://doi.org/10.1111/2041-210X.12142>
- Pavoine, S., Love, M., Bonsall, M. B. (2009). Hierarchical partitioning of evolutionary and ecological patterns in the organization of phylogenetically-structured species assemblages: application to rockfish (genus: *Sebastes*) in the Southern California Bight. *Ecology Letters*, 12, 898–908. <https://doi.org/10.1111/j.1461-0248.2009.01344.x>
- Pavoine, S., Marcon, E., Ricotta, C. (2016). "Equivalent numbers" for species, phylogenetic, or functional diversity in a nested hierarchy of multiple scales. *Methods in Ecology and Evolution*, 7, 1152–1163. <https://doi.org/10.1111/2041-210X.12591>
- Pavoine, S., Ollier, S., Dufour, A. B. (2005). Is the originality of a species measurable? *Ecology Letters*, 8, 579–586. <https://doi.org/10.1111/j.1461-0248.2005.00752.x>
- Pavoine, S., Ollier, S., Pontier, D., Chessel, D. (2008). Testing for phylogenetic signal in phenotypic traits: new matrices of phylogenetic proximities. *Theoretical Population Biology*, 73, 79–91. <https://doi.org/10.1016/j.tpb.2007.10.001>
- Pavoine, S., Ricotta, C. (2013). Testing for phylogenetic signal in biological traits: the ubiquity of cross-product statistics. *Evolution*, 67, 828–840. <https://doi.org/10.1111/j.1558-5646.2012.01823.x>

- 
- Pavoine, S., Ricotta, C. (2014). Functional and phylogenetic similarity among communities. *Methods in Ecology and Evolution*, 5, 666–675. <https://doi.org/10.1111/2041-210X.12193>
- Pavoine, S., Ricotta, C. (2019a). A simple translation from indices of species diversity to indices of phylogenetic diversity. *Ecological Indicators*, 101, 552–561. <https://doi.org/10.1016/j.ecolind.2019.01.052>
- Pavoine, S., Ricotta, C. (2019b). Measuring functional dissimilarity among plots: adapting old methods to new questions. *Ecological Indicators*, 97, 67–72. <https://doi.org/10.1016/j.ecolind.2018.09.048>
- Pavoine, S., Ricotta, C. (2021) On the relationships between rarity, uniqueness, distinctiveness, originality and functional/phylogenetic diversity. *BiorXiv*. <https://doi.org/10.1101/2021.08.09.455640>
- Pavoine, S., Vela, E., Gachet, S., de Belair, G., Bonsall, M.B. (2011). Linking patterns in phylogeny, traits, abiotic variables and space: a novel approach to linking environmental filtering and plant community assembly. *Journal of Ecology*, 99, 165–175. <https://doi.org/10.1111/j.1365-2745.2010.01743.x>
- Pielou, E.C. (1975). *Ecological diversity*. New York: Wiley InterScience.
- R Hackathon et al. (2019). *phylobase: Base Package for Phylogenetic Structures and Comparative Data*. R package version 0.8.6. <https://CRAN.R-project.org/package=phylobase>
- Rao, C.R. (1982). Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology*, 21, 24–43. [https://doi.org/10.1016/0040-5809\(82\)90004-1](https://doi.org/10.1016/0040-5809(82)90004-1)
- Rao, C.R. (1986). Rao’s axiomatization of diversity measures. In: S. Kotz, N.L. Johnson (Eds), *Encyclopedia of Statistical Sciences* (pp. 614–617). New York: Wiley and Sons.
- Redding, D.W. (2003). *Incorporating genetic distinctness and reserve occupancy into a conservation prioritisation approach*. Master thesis: University of East Anglia, Norwich.
- Redding, D.W., Mooers, A.O. (2006). Incorporating evolutionary measures into conservation prioritization. *Conservation Biology*, 20, 1670–1678. <https://doi.org/10.1111/j.1523-1739.2006.00555.x>
- Renyi, A. (1960). On measures of entropy and information. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 1, 547–561.
- Ricotta, C. (2004). A parametric diversity measure combining the relative abundances and taxonomic distinctiveness of species. *Diversity and Distributions*, 10, 143–146. <https://doi.org/10.1111/j.1366-9516.2004.00069.x>
- Ricotta, C., Bacaro, G., Caccianiga, M., Cerabolini, B.E.L., Pavoine, S. (2018). A new method for quantifying the phylogenetic redundancy of biological communities. *Oecologia*, 186, 339–346. <https://doi.org/10.1007/s00442-017-4026-x>
- Ricotta, C., Bacaro, G., Pavoine, S. (2015). A cautionary note on some phylogenetic dissimilarity measures. *Journal of Plant Ecology*, 8, 12–16. <https://doi.org/10.1093/jpe/rtu008>
- Ricotta, C., de Bello, F., Moretti, M., Caccianiga, M., Cerabolini, B.E., Pavoine, S. (2016). Measuring the functional redundancy of biological communities: A quantitative guide. *Methods in Ecology and Evolution*, 7, 1386–1395. <https://doi.org/10.1111/2041-210X.12604>
- Ricotta, C., Laroche, F., Szeidl, L., Pavoine, S. (2020) From alpha to beta functional and phylogenetic redundancy. *Methods in Ecology and Evolution*, 11, 487–493. <https://doi.org/10.1111/2041-210X.13353>
- Ricotta, C., Kosman, E., Laroche, F., Pavoine, S. (2021) Beta redundancy for functional ecology. *Methods in Ecology and Evolution*, 12, 1062–1069. <https://doi.org/10.1111/2041-210X.13587>
- Ricotta, C., Pavoine, S. (2015a). A multiple-site dissimilarity measure for species presence/absence data and its relationship with nestedness and turnover. *Ecological Indicators*, 54, 203–206. <https://doi.org/10.1016/j.ecolind.2015.02.026>



- 
- Ricotta, C., Pavoine, S. (2015b) Measuring similarity among plots including similarity among species: an extension of traditional approaches. *Journal of Vegetation Science*, 26, 1061–1067. <https://doi.org/10.1111/jvs.12329>
- Ricotta, C., Podani, J., Pavoine, S. (2016). A family of functional dissimilarity measures for presence and absence data. *Ecology and Evolution*, 6, 5383–5389. <https://doi.org/10.1002/ece3.2214>
- Schmera, D., Podani, J., Eros, T. (2009). Measuring the contribution of community members to functional diversity. *Oikos*, 118, 961–971. <https://doi.org/10.1111/j.1600-0706.2009.17076.x>
- Siberchicot, A., Julien-Laferrrière, A., Dufour, A.-B., Thioulouse, J., Dray, S. (2017). adegraphics: an S4 lattice-based package for the representation of multivariate data. *The R Journal*, 9, 198–212.
- Shannon, C.E. (1948). A mathematical theory of communication. *Bell System technical journal*, 27, 379–423, 623–656.
- Simpson, E.H. (1949). Measurement of diversity. *Nature*, 163, 688. <https://doi.org/10.1038/163688a0>
- Smith, B., Wilson, J.B. (1996). A consumer’s guide to evenness measures. *Oikos*, 76, 70–82. <https://www.jstor.org/stable/3545749>
- Thioulouse, J., Dray, S., Dufour, A.-B., Siberchicot, A., Jombart, J., Pavoine, S. (2018). *Multivariate Analysis of Ecological Data with ade4*. Springer.
- Tsallis, C. (1988). Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*, 52, 479–487. <https://doi.org/10.1007/BF01016429>
- Vane-Wright, R.I., Humphries, C.J., Williams, P.H. (1991). What to protect? Systematics and the agony of choice. *Biological Conservation*, 55, 235–254. [https://doi.org/10.1016/0006-3207\(91\)90030-D](https://doi.org/10.1016/0006-3207(91)90030-D)