INTRODUCTION

The maintenance of ecosystem processes under ongoing perturbations causing local species extinctions requires the presence of functionally similar species. This property is usually known as functional redundancy (Naeem, 1998; Walker, 1992). When several species perform similar functions but differ in their responses to disturbances, functional redundancy may enhance community stability, while providing insurance against the loss of ecosystem functions due to a disturbance (Yachi & Loreau, 1999). Therefore, if functional redundancy is not maintained, species extinctions may jeopardize ecosystem stability, suggesting the need for new conservation strategies to address species extinctions (Naeem et al., 2004).
redundancy is high, it is more likely that the loss of a given species will have relatively little effect on ecosystem processes (Petchey, Evans, Fishburn, & Gaston, 2007; Pillar et al., 2013).

Ricotta et al. (2016; see also de Bello, Carmona, Lepš, Szava-Kovats, & Pärtel, 2016) proposed to summarize functional uniqueness within plots (the complement of redundancy) as the ratio of two diversity measures: the Rao quadratic entropy $Q$, which is calculated using the species abundances and their trait dissimilarities and the Simpson index $S$, which considers all species maximally dissimilar from each other such that $U_1 = Q/S$. Functional uniqueness thus measures the diversity decrease that is obtained by including trait dissimilarities in the calculation of plot-level diversity or alpha diversity. Similarly, since $Q \leq S$, alpha-redundancy can be interpreted as the fraction of species diversity not expressed by functional diversity. According to this definition, a necessary condition to obtain a normalized index of beta redundancy in the range $[0,1]$ is that for a given pair of plots, the functional (or phylogenetic) dissimilarity $D_F$ (or $D_S$) is always lower or equal to their species dissimilarity $D_S$, where $D_S$ is the compositional dissimilarity of a hypothetical pair of plots with the same species abundances of the actual pair of plots in which all species are considered equally and maximally dissimilar from each other. In this case, beta redundancy can be summarized as:

$$ R_B = 1 - U_B = (D_S - D_F)/D_S. \tag{1} $$

However, under particular circumstances, many of the existing indices of functional or phylogenetic dissimilarity can lead to values greater than for species dissimilarity, thus violating the basic requirement that $D_F, D_P \leq D_S$ (see Appendix S1). The aim of this paper is to introduce a new family of tree-based measures of phylogenetic and functional dissimilarity that conform to the requirement $D_F, D_P \leq D_S$.

A worked example with published data on Alpine plant communities along a primary succession on a glacier foreland in northern Italy is used to illustrate our approach.

### 2 | NEW TREE-BASED MEASURES OF PHYLOGENETIC BETA REDUNDANCY

Phylogenetic trees have long been used to represent the evolutionary history and relationships among groups of organisms, such as species (Podani, 2017). In ecology, phylogenetic trees have been widely adopted to analyse the phylogenetic component of biodiversity (Tucker et al., 2017). Likewise, most measures of phylogenetic dissimilarity take into account the branching pattern of the species phylogeny (Chiu et al., 2014; Pavoine, 2016). For a given assemblage, the terminal nodes of the phylogenetic tree usually represent the species in the assemblage, while the interior nodes are ‘hypothetical taxonomic units’ (i.e. the hypothetically most recent common ancestor of all descendent taxa in the phylogenetic tree; Farris, 1970). If the branch lengths are proportional to the time of divergence among species, all species are at the same distance from the root node and the tree is called ‘ultrametric’. A general method for incorporating the branching pattern of ultrametric trees like that in Figure 1 in the calculation of a wide range of phylogenetic diversity metrics has been developed by Pavoine, Love, and Bonsall (2009) and Chao, Chiu, and Jost (2010).

For an ultrametric phylogeny composed of $N$ species, let $A$ and $B$ be two plots and $x_i$ be the abundance of species $i$ ($i=1,2,...,N$) in plot $A$. The phylogenetic tree is first divided into $K$ evolutionary periods as follows: the tree is sliced at each node and the node ages are
The age of the root node is labelled as \( t_0 \) (the age of the root node) and \( t_0 = T \) (the present-day time). The root node can either be the common ancestor of the species in \( A \) and \( B \) or the common ancestor of a larger group of species defined, for example at regional scale, depending on the question at hand (see Ricotta et al., 2018).

Two adjacent slices delimit an evolutionary period \( \Delta_k = t_{k-1} - t_k \) with \( \sum_{k=1}^N \Delta_k = T \). Any given evolutionary period \( k \) contains \( M_k \) lineages with abundances \( x_{km} \) \((m = 1, 2, ..., M_k)\). The abundance of each lineage along the phylogeny is calculated by summing the abundances of all descendant species in the phylogeny. For example \( x_{11} = x_{11} + x_{21}, x_{12} = x_{21} + x_{22} + x_{23} + x_{24} + x_{25} \) and \( x_{k1} = x_1 \) labelled in temporal order as \( t_k \) with \( t_0 = T \) (the age of the root node) and \( t_k = 0 \) (the present-day time).

Given two plots \( A \) and \( B \), the corresponding lineage abundances \( x_{Akm} \) and \( x_{Bkm} \) can be then used for calculating a lineage dissimilarity \( D_k \) between \( A \) and \( B \) for each node age \( t_k \) of the tree. The lineage dissimilarity at time \( t_k \) is the usual present-day dissimilarity calculated from the species abundances \( x_{Akm} \) and \( x_{Bkm} \). The overall phylogenetic dissimilarity \( D_p \) is then obtained by averaging the measures \( D_k \) over the time intervals of the corresponding evolutionary periods (Ricotta et al., 2018):

\[
D_p = \sum_{k=1}^K D_k \times r_k,
\]

where \( r_k = \Delta_k / T \) is the duration of the \( k \)-th evolutionary period after rescaling the tip-to-root length to unit, and \( \Delta_k \) (and hence \( r_k \)) = 0 by definition. \( D_p \) thus measures the phylogenetic dissimilarity between plots \( A \) and \( B \) incorporating information about the branching pattern of the species phylogeny and its lineage abundances.

If phylogenetic dissimilarity at time \( t_k \) is calculated with the Bray–Curtis index (Bray & Curtis, 1957), the present-day species dissimilarity \( D_3 = D_k = \sum_{i=1}^{N} |x_{Ai} - x_{Bi}| / \sum_{i=1}^{N} (x_{Ai} + x_{Bi}) \) is always greater than or equal to any other lineage dissimilarity \( D_{k} \) (proof in Appendix S2). Therefore, \( D_p \leq D_3 \) and we can get a normalized index of phylogenetic beta redundancy \( R_p \) or beta uniqueness \( U_p \) as \( R_p = 1 - U_p = (D_3 - D_p) / D_3 \).

This measure represents the fraction of species dissimilarity between a pair of plots not expressed by phylogenetic dissimilarity. Note that for a pair of plots \( A \) and \( B \) with given species abundances \( x_{Akm} \) and \( x_{Bkm} \), the uniqueness component \( U_p = D_p / D_3 \) can be interpreted as a measure of phylogenetic dissimilarity between both plots normalized by its maximum value. This maximum is the phylogenetic dissimilarity of a pair of plots with identical abundance vectors but in which all species had independent evolution. In phylogenetic terms, this means that all terminal branches originate from a single polytomy at the root node, thus forming a star phylogeny (Ricotta et al., 2018). Likewise, the complement of uniqueness, beta redundancy \( R_p = 1 - U_p \), can be interpreted as a normalized measure of phylogenetic similarity in the range [0, 1].

### 3 | NEW TREE-BASED MEASURES OF FUNCTIONAL BETA REDUNDANCY

In principle, the same tree-based approach can be used to assess functional dissimilarity. This requires to organize functional differences between species to obtain a hierarchical structure as first suggested by Petchey and Gaston (2002). Functional dendrograms can be constructed by first calculating functional dissimilarities between species and then using any hierarchical clustering approach to obtain a rooted tree from these dissimilarities with species as tips (Pavoine & Ricotta, 2019). There are several clustering methods for constructing a tree-like classification in which species are aggregated into hierarchical clusters (Legendre & Legendre, 2012; Podani, 2000). The structure of the resulting functional dendrogram is thus sensitive to the methods used to define a tree hierarchy from functional dissimilarities between species.

Much of the discussion on tree-based functional measures has been focused around the selection of appropriate clustering methods (Mouchet et al., 2008; Petchey & Gaston, 2007; Podani & Schmera, 2006). Nonetheless, there is a lack of agreement over which method produces the best representation of species distribution in hierarchical trait space. For example Weitzman (1992) developed an algorithmic measure of (functional) diversity, which can be geometrically represented as the length of a rooted dendrogram, whereas Blackburn, Petchey, Cassey, and Gaston (2005) and Petchey and Gaston (2007) suggested to adopt case by case the clustering algorithm that maximizes the cophenetic correlation between pairwise species dissimilarities in trait space and the corresponding pairwise dissimilarities across the functional dendrogram.
Podani and Schmera (2006, 2007) argued that the ideal requirement of maximizing the cophenetic correlation is hardly feasible in practice as the virtually infinite number of available clustering procedures (see Lance & Williams, 1966) would unnecessarily overcomplicate the calculations. They suggested instead to use group average (i.e. unweighted pair group method with arithmetic mean, UPGMA) as a standard clustering procedure. Mouchet et al. (2008) proposed to use consensus trees, which incorporate features common to a number of different clustering methods. Additional hierarchical clustering methods may rely on the maximization of within-cluster homogeneity based on some a-priori defined criteria, such as variance, sum of squares or within-cluster average similarity of objects (see Podani, 1989).

In this paper, we used the UPGMA clustering method because it represents a good compromise between the methodological extremes of a continuous series of space-contracting and space-dilating algorithms (Gordon, 1999). Space-contracting algorithms, such as the single linkage method, tend to contract the input dissimilarity structure significantly. Therefore, they may leave some clusters undetected even if the input data space shows a clear group structure. Space-dilating algorithms, such as the complete linkage method, tend to produce a clustered structure even if it is not apparent in the input data space. Within this continuous series of clustering procedures, UPGMA generally preserves a large proportion of the input dissimilarity structure, thus tending to minimize the distortion induced by transforming the input dissimilarity matrix into a functional dendrogram (Podani & Schmera, 2006).

As for the phylogenetic case, the functional dendrogram obtained from the UPGMA algorithm can then be used for calculating a measure of pairwise functional dissimilarity among plots $D_F$ together with the corresponding measures of functional redundancy and uniqueness $R_F = 1 - U_F = (D_s - D_F) / D_s$.

Note that, irrespective of the clustering method, the functional dendrogram needs to be rooted and scaled to unit height from the tips to the root, so that two species sharing no branch in the tree would have maximum possible functional differences between them (for details, see Appendix S3).

## 4 | WORKED EXAMPLE

### 4.1 | Data

We analysed the functional and phylogenetic beta redundancy of plant communities collected by Caccianiga, Luzzaro, Pierce, Ceriani, and Cerabolini (2006) along a primary succession on glacial deposits. The dataset is composed of 59 plots, each of about 25 m² in size sampled above the tree line at the foreland of the Rutor Glacier (Italy). See Caccianiga et al. (2006) for additional information on the study site. Based on the age of the moraine ridges, the plots were assigned to three successional stages: early-succession (17 plots), mid-succession (32 plots) and late-succession (10 plots). For each plot, species abundances were measured with a five-point ordinal scale transformed to ranks.

The 45 species sampled were classified in terms of Grime’s (1974) plant strategy theory, as competitors (C), stress tolerators (S) and ruderals (R) with fuzzy-coded values in the range 0–100, such that the sum of $C + S + R$ was equal to 100 (all data are available in Ricotta et al., 2016, Appendix S2). A functional dissimilarity matrix between species was then obtained by applying the Bray–Curtis coefficient to the CSR species classification. Finally, we used the phylogeny of the 45 species available in appendix A of Ricotta, Bacaro, Caccianiga, Cerabolini, and Moretti (2015) after rescaling its tip-to-root length to unit.

### 4.2 | Methods

We calculated functional and phylogenetic beta uniqueness $U_F$ and $U_P$ for all pairs of plots in each successional stage. All calculations were done with a new R function ‘DP’ available in Appendix S3. We next tested for differences in the dispersion of $U_F$ and $U_P$ among the three successional stages. Since the beta uniqueness of a pair of plots $A$ and $B$ is basically a normalized measure of dissimilarity between $A$ and $B$, we used the PERMDISP test (Permutational Analysis of Multivariate Dispersions) of Anderson (2006) to test for differences in the average dissimilarity of individual plots from their group centroid among successional stages. This test has been extensively used in ecology for comparing the beta diversity among groups of plots.

First, starting from the functional and phylogenetic beta uniqueness between pairs of plots, we calculated for each successional stage the normalized dissimilarity (uniqueness) of each individual plot from its group centroid. Next, a permutational $t$ test for pairwise differences in average dissimilarity from the group centroids is performed. $p$-values were obtained using 9,999 pairwise permutations of the dissimilarities of individual plots from the corresponding group centroids.

### 4.3 | Results

The phylogenetic and functional dendrograms of the 45 species used in this study are shown in Appendix S4 and the behaviour of functional and phylogenetic beta uniqueness along the successional stages is summarized in Table 1. As shown by Caccianiga et al. (2006) and Ricotta et al. (2016, 2018) the colonization of the glacial deposits by the first pioneer species in the early-successional stage is mainly driven by random dispersal mechanisms, while in the mid- and late-successional stages, a general tendency towards an increased structural homogeneity is observed. At the plot level, the increased uniformity of vegetation structure over time is associated with a functional and phylogenetic shift from early-successional ruderal forbs to late-successional stress-tolerator graminoids that goes together with an increase in functional and phylogenetic alpha redundancy.
TABLE 1 Mean (SD) dissimilarity values of each individual plot from its group centroid for the three successional stages on the moraine deposits of the Rutor Glacier. Pairwise comparisons of index differences between the successional stages were performed with permutation-based tests for pairwise differences in average dissimilarity from the group centroids (9,999 permutations). For each dissimilarity coefficient, numbers followed by the same letter do not differ significantly at p < .05. Ds = species dissimilarity (Bray–Curtis dissimilarity); Df = functional dissimilarity; Dp = phylogenetic dissimilarity; UP/UP = functional/phylogenetic uniqueness.

<table>
<thead>
<tr>
<th></th>
<th>Early-succ. stage</th>
<th>Mid-succ. stage</th>
<th>Late-succ. stage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Functional data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ds</td>
<td>0.399 (0.071)a</td>
<td>0.357 (0.084)a</td>
<td>0.354 (0.101)a</td>
</tr>
<tr>
<td>Df</td>
<td>0.118 (0.034)a</td>
<td>0.085 (0.024)b</td>
<td>0.073 (0.033)b</td>
</tr>
<tr>
<td>UP = Df/Ds</td>
<td>0.203 (0.045)a</td>
<td>0.168 (0.035)b</td>
<td>0.144 (0.069)b</td>
</tr>
<tr>
<td><strong>Phylogenetic data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ds</td>
<td>0.399 (0.071)a</td>
<td>0.357 (0.084)a</td>
<td>0.354 (0.101)a</td>
</tr>
<tr>
<td>Dp</td>
<td>0.291 (0.070)a</td>
<td>0.214 (0.053)b</td>
<td>0.222 (0.063)ab</td>
</tr>
<tr>
<td>UP = Dp/Ds</td>
<td>0.498 (0.076)a</td>
<td>0.422 (0.072)b</td>
<td>0.427 (0.101)ab</td>
</tr>
</tbody>
</table>

Looking at the beta component of uniqueness, Table 1 shows that, due to the low structural homogeneity of the pioneer vegetation, the species, functional and phylogenetic turnover among the early-successional plots is generally higher than that of the mid- and late successional plots. More specifically, while traditional species turnover does not differ significantly among the three successional stages, the functional turnover and uniqueness of the 17 plots in the early successional stage are both significantly higher than those in the mid- and late successional stages. This means that in the mid- and late-successional stages the species in one plot tend to be replaced by functionally related species in the other plot. As a result, functional uniqueness UP tends to be low. In contrast, due to the more random nature of the dispersal mechanisms, in the early-successional stage, the species in one plot tend to be replaced by functionally unrelated (or, at least, less related) species. Therefore, functional uniqueness tends to be higher. The same pattern is shown by the phylogenetic uniqueness, although in this case the statistical differences among the three successional stages are not always significant at p < .05.

5 | DISCUSSION

In this paper, we introduced the notion of beta redundancy. In analogy to plot-level redundancy or alpha redundancy, beta redundancy can be defined as the fraction of species dissimilarity between pairs of plots not expressed by functional or phylogenetic dissimilarity. In its very essence, beta redundancy tells us to what degree the species-based similarity between two plots is associated to the functional or phylogenetic similarities among the species in both plots. While functional and phylogenetic dissimilarity Df and Dp are absolute measures that take into account the abundances and the functional or phylogenetic differences among the species in both plots, beta redundancy Rp and its complement beta uniqueness Up = 1 − Rp are standardized coefficients that relate Df and Dp to plot-to-plot species dissimilarity Ds. A fundamental requirement to get a meaningful index of beta redundancy is thus that Df, Dp ≤ Ds. However, contrary to common belief, many of the existing indices of functional or phylogenetic dissimilarity do not conform to this requirement. Therefore, they cannot be used to calculate an appropriate measure of beta redundancy. To solve this problem, we introduced a family of tree-based indices of functional and phylogenetic dissimilarity that conform to this requirement. Regardless of whether beta redundancy is calculated from functional or phylogenetic data, Df and Dp have a great potential for future research on community organization, as both measures are variations on the theme of the classical index of Bray and Curtis (1957), the properties of which have been studied by ecologists for decades (e.g. Legendre & De Cáceres, 2013; Ricotta & Podani, 2017).

Being based on a hierarchical structure of species differences, the proposed phylogenetic measures take into account the branching pattern of the phylogenetic tree, which is the more natural way for describing the evolutionary relationships among the species. In contrast, a tree-based representation of the species functional relationships may not be the most appropriate solution. However, if we accept the idea of Petchey and Gaston (2007, p. 1422) that from a statistical viewpoint ‘the functional dendrogram can be thought of as a description of the functional relationships shared by the species it includes, in the same way as a phylogenetic tree describes phylogenetic relationships (although without any associated inference about evolutionary relationships)’, we obtain a powerful and flexible framework for the calculation of different facets of functional diversity, richness and redundancy.

From a more technical viewpoint, our approach requires that the tip-to-root length T of the phylogenetic or functional tree be normalized to the unit interval. This can be done either by a-priori normalizing the (functional) dissimilarities among species in the range [0–1], or by a-posteriori normalizing the length of the resulting dendrogram. Since there is no univocal way to normalize tree length, flexibility is important to adapt the way the data are transformed to the specific objectives of our work. A short guide on how to normalize the length of a hierarchical dendrogram according to the specific question at hand can be found in Pavoine, Marcon, and Ricotta (2016, Appendix S3).

In conclusion, we see this work more as a starting point for the summarization of beta redundancy rather than as an ultimate solution, and many questions remain still open. For example how do the proposed beta redundancy metrics respond to large differences in species richness and evenness? (Here, it will also be important to explore which among rare and common species contribute most to redundancy, as this is related to the metrics sensitivity to sampling effort). How does species splitting affect the index values or, more generally, how sensitive is phylogenetic redundancy to taxonomic changes (see Robuchon et al., 2019)? Also, functional and phylogenetic dissimilarity measures need not necessarily be based on a tree. For example we could also
consider evolutionary distances calculated from molecular data before any tree is derived from these distances. Therefore, apart from tree-based indices, are there other classes of measures of functional or phylogenetic dissimilarity which conform to the requirement $D_F, D_P \leq D_S$? Which additional properties should these measures possess? In Appendix S1 we show that eight existing indices of plot-to-plot dissimilarity fulfil the requirement $D_F, D_P \leq D_S$. However, for the same indices, we also show some unexpected and counter-intuitive behaviours: part of these measures attribute positive dissimilarity values to two identical plots, whereas, by definition, the dissimilarity between a plot and itself should be zero. Other indices retain only nearest-neighbour distances between species for the calculation of functional or phylogenetic dissimilarity (see Appendix S1; Ricotta, Bacaro, & Pavoine, 2015). Further studies should thus consider a range of properties fulfilled by any potential index of plot-to-plot functional or phylogenetic dissimilarity, rather than focusing exclusively on the requirement $D_F, D_P \leq D_S$.

Last but not least, in diversity theory alpha and beta diversity have been usually related to gamma or regional diversity either by an additive model $\alpha + \beta = \gamma$ (McArthur, Recher, & Cody, 1966) or by a multiplicative model $\alpha \times \beta = \gamma$ (Whittaker, 1960). Will it be possible to develop a similar model for alpha and beta redundancy? These are critical questions, and their answers may provide new tools for analysing the complex processes that drive community assembly and species co-occurrence.

ACKNOWLEDGEMENTS

We thank two anonymous reviewers for their comments. They both contributed to improve our paper.

AUTHORS’ CONTRIBUTIONS

C.R. and S.P. conceived the ideas and formulated the research problem; F.L. and L.S. provided important feedback which helped shape the research; C.R. and S.P. analysed the data; C.R. took the lead in writing the main text of the manuscript and S.P. in writing the appendices. All authors revised the manuscript critically and approved the final version.

DATA AVAILABILITY STATEMENT

The data and the R code used in this paper are deposited in the Dryad Digital Repository https://doi.org/10.5061/dryad.ttdz08ktg (Ricotta, Larocche, Szeidl, Pavoine, & Isaac, 2020).

ORCID

Carlo Ricotta https://orcid.org/0000-0003-0818-3959

Sandra Pavoine https://orcid.org/0000-0003-2767-6484

REFERENCES


**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section.
Appendix S1. Examples of some undesired properties of traditional functional and phylogenetic dissimilarity measures and perspectives for further developments.

Most indices of functional and phylogenetic alpha diversity have been defined so that they are never higher than species diversity. Their maximum value over functional or phylogenetic data is obtained when all species are maximally dissimilar from each other with a distance of 1 between any two species (1 being the highest possible distance). This does not hold for many indices of functional or phylogenetic dissimilarity between two plots (a component of beta diversity). Under particular circumstances, many of these indices can lead to functional or phylogenetic dissimilarity greater than species-based dissimilarity, as illustrated below.

We designate here by ‘redundancy property’ the requirement for a plot-to-plot functional or phylogenetic dissimilarity index to always be lower than or equal to the associated plot-to-plot species-based dissimilarity (where all species are considered maximally dissimilar from each other).

To analyze index behavior, we used the R packages adiv (functions dissABC, evodiss, sDQ, Pavoine, 2018), picante (function phylosor, Kembel et al., 2010) and phyloseq (function UniFrac, McMurdie & Holmes 2013). We also used the function FD_Beta made available by Chao et al. (2019) and our own scripts that we will all make available in a new version of package adiv under the names Dnn, Dpw, DH and generalized_TradiDiss.

Part S1-1. Examples of indices that do not fulfill the redundancy property

Case study 1. As an illustration, we first considered the following case study.

4 species (named a, b, c, d) are connected by their ultrametric phylogenetic (or functional) relationships. Three of these species are closely related (or functionally close) (species b, c and d) and the fourth species (species a) is the sole member of its lineage (or functional group) (Fig. S1.1). Plot A contains all species with even relative abundances, while plot B contains species a only (Table S1.1).

Fig. S1.1. Theoretical phylogenetic or functional ultrametric tree with species a, b, c and d as tips (root on the left of the tree); the length of the branch that connects species a to the root is equal to 1.

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>plotA</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>plotB</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table S1.1. Relative abundances of species a, b, c and d in two plots (plot A and plot B)
The same data set can be also visualized in terms of a matrix of phylogenetic or functional dissimilarities between species together with a matrix of species relative abundances. The matrix of pairwise dissimilarities between species contains for each species pair the branch length between both species and their most recent common node in the tree of Fig. S1.1. These pairwise dissimilarities are shown in Table S1.2.

\[
\begin{array}{cccc}
  a & b & c & d \\
  a & 0.00 & 1.00 & 1.00 & 1.00 \\
  b & 1.00 & 0.00 & 0.02 & 0.02 \\
  c & 1.00 & 0.02 & 0.00 & 0.01 \\
  d & 1.00 & 0.02 & 0.01 & 0.00 \\
\end{array}
\]

Table S1.2. Dissimilarities between species, calculated from Fig. S1.1

All dissimilarities between species are bounded between 0 (the species are functionally or phylogenetically equivalent) and 1 (the species are maximally dissimilar). We thus compared the values assumed by selected plot-to-plot functional and phylogenetic dissimilarity indices (i.e. considering the phylo/functional differences between species; Fig. S1.1 and Tables S1.1 and S1.2) with the corresponding species-based dissimilarities. The species-based dissimilarities are calculated from the same species abundances (Table S1.1) but considering all species as maximally dissimilar from each other. Such maximum distances between species are described by the star tree in Fig. S1.2 and by the species dissimilarity matrix in Table S1.3.

![Star tree](image)

\[
\begin{array}{cccc}
  a & b & c & d \\
  a & 0 & 1 & 1 & 1 \\
  b & 1 & 0 & 0 & 1 \\
  c & 1 & 1 & 0 & 1 \\
  d & 1 & 1 & 1 & 0 \\
\end{array}
\]

Table S1.3. Dissimilarities between species, calculated from Fig. S1.2
In Table S1.4 we provide examples of indices that do not satisfy the redundancy property according to this case study. That is, for all indices in Table S1.4, the phylogenetic or functional plot-to-plot dissimilarity (the phylo/functional case) is higher than the corresponding species-based dissimilarity (the species case). [Note that the conclusions raised from this first case study are unchanged if the relative abundances of all four species in plot A are modified].

<table>
<thead>
<tr>
<th>Index</th>
<th>Reference</th>
<th>Phylo/functional case</th>
<th>Species case</th>
</tr>
</thead>
<tbody>
<tr>
<td>DISC</td>
<td>Rao (1982)</td>
<td>0.559</td>
<td>0.375</td>
</tr>
<tr>
<td>PhyloSor</td>
<td>Bryant et al. (2008)</td>
<td>0.660</td>
<td>0.400</td>
</tr>
<tr>
<td>$1-U_{22}=V_{22}$ ($q=2$)</td>
<td>Chiu, Jost, &amp; Chao (2014)</td>
<td>0.528</td>
<td>0.429</td>
</tr>
<tr>
<td>$1-C_{22}=S_{22}$</td>
<td>Chiu, Jost, &amp; Chao (2014)</td>
<td>0.691</td>
<td>0.600</td>
</tr>
<tr>
<td>$1-U_{22}=V_{22}$ ($q=2$)</td>
<td>Chao et al. (2019)*</td>
<td>0.528</td>
<td>0.429</td>
</tr>
<tr>
<td>$1-C_{22}=S_{22}$</td>
<td>Chao et al. (2019)*</td>
<td>0.691</td>
<td>0.600</td>
</tr>
<tr>
<td>$1-S_{Sokal-Sneath}$</td>
<td>Pavoine &amp; Ricotta (2014)</td>
<td>0.899</td>
<td>0.857</td>
</tr>
<tr>
<td>$1-S_{Jaccard}$</td>
<td>Pavoine &amp; Ricotta (2014)</td>
<td>0.817</td>
<td>0.750</td>
</tr>
<tr>
<td>$1-S_{Sørensen}$</td>
<td>Pavoine &amp; Ricotta (2014)</td>
<td>0.691</td>
<td>0.600</td>
</tr>
<tr>
<td>$1-S_{Ochiai}$</td>
<td>Pavoine &amp; Ricotta (2014)</td>
<td>0.682</td>
<td>0.500</td>
</tr>
<tr>
<td>$1-S_{S}$</td>
<td>Pavoine &amp; Ricotta (2014)</td>
<td>0.528</td>
<td>0.429</td>
</tr>
<tr>
<td>$1-S_{Jaccard}$</td>
<td>Ricotta &amp; Pavoine (2015)†</td>
<td>0.922</td>
<td>0.857</td>
</tr>
<tr>
<td>$1-S_{Sørensen}$</td>
<td>Ricotta &amp; Pavoine (2015)†</td>
<td>0.856</td>
<td>0.750</td>
</tr>
<tr>
<td>$1-S_{Ochiai}$</td>
<td>Ricotta &amp; Pavoine (2015)†</td>
<td>0.841</td>
<td>0.750</td>
</tr>
<tr>
<td>$1-S_{Kulczynski}$</td>
<td>Ricotta &amp; Pavoine (2015)†</td>
<td>0.824</td>
<td>0.750</td>
</tr>
<tr>
<td>$1-S_{Sokal-Sneath}$</td>
<td>Ricotta &amp; Pavoine (2015)†</td>
<td>0.960</td>
<td>0.923</td>
</tr>
<tr>
<td>$evoD_{Minkowski}$</td>
<td>Pavoine (2016)</td>
<td>1.058</td>
<td>0.866</td>
</tr>
<tr>
<td>$evoD_{chord}$</td>
<td>Pavoine (2016)</td>
<td>1.168</td>
<td>1.000</td>
</tr>
<tr>
<td>$evoD_{Morisita-Horn}$</td>
<td>Pavoine (2016)</td>
<td>0.691</td>
<td>0.600</td>
</tr>
<tr>
<td>$evoD_{profile}$</td>
<td>Pavoine (2016)</td>
<td>1.058</td>
<td>0.866</td>
</tr>
<tr>
<td>$evoD_{Anyonrel}$</td>
<td>Pavoine (2016)</td>
<td>0.612</td>
<td>0.515</td>
</tr>
</tbody>
</table>

* Here we used $\tau = \text{maximum observed dissimilarity between any two species}$ and $f(d_{ij}) = d_{ij} / \tau$.
† Here we used equations 6-8 of the main text of Ricotta & Pavoine (2015).

**Table S1.4.** Example of indices for which the phylo/functional dissimilarity between plot A and plot B is higher than the species-based dissimilarity with our first case study.

**Case study 2.** In the second case study, the phylogenetic or functional dissimilarity between any two species $i$ and $j$ is equal to 0.1 (Table S1.5). We consider two species only. Their abundances are given in Table S1.6.

```
<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>0.90</td>
</tr>
<tr>
<td>0.90</td>
<td>0.00</td>
</tr>
</tbody>
</table>
```

**Table S1.5.** Phylo/functional dissimilarities between two species
Table S1.6. Abundances of species $a$, and $b$ in two plots (plot A and plot B)

We applied the indices $GC$, $MS$, and $PE$ of Pavoine & Ricotta (2019) to Tables S1.5 and S1.6 ($GC =$ equation 6 in Pavoine & Ricotta, 2019, $MS =$ equation 8 and $PE =$ equations 9 and 10 with $\pi_j =$ equation 5). Similarities were calculated as $1 -$ dissimilarities. The results are given in Table S1.7.

<table>
<thead>
<tr>
<th>Index</th>
<th>Reference</th>
<th>Phylo/functional case</th>
<th>Species case</th>
</tr>
</thead>
<tbody>
<tr>
<td>Functional GC</td>
<td>Pavoine &amp; Ricotta (2019)</td>
<td>0.784</td>
<td>0.773</td>
</tr>
<tr>
<td>Functional PE</td>
<td>Pavoine &amp; Ricotta (2019)</td>
<td>0.509</td>
<td>0.500</td>
</tr>
<tr>
<td>Functional MS</td>
<td>Pavoine &amp; Ricotta (2019)</td>
<td>0.878</td>
<td>0.869</td>
</tr>
</tbody>
</table>

Table S1.7. Dissimilarity between plot A and plot B for the second case study. In the phylo/functional case, we used the dissimilarities between species given in Table S1.5; in the species case the dissimilarity between any two species equals 1.

This second case study shows that, dealing with absolute abundances, the indices $GC$, $MS$ and $PE$ of Pavoine & Ricotta (2019) do not fulfill the redundancy property: the species-based dissimilarities between plots can be lower than the functional dissimilarities between plots.

Case study 3. Here, we consider the same trees as in case study 1 (Fig. S1.1 for a phylo/functional tree and Fig. S1.2 for a star tree representing maximum dissimilarity between species). However, in case study 3 we consider that plot $A$ contains all species, while plot $B$ contains species $b$, c and d only. In that case, the index UniFrac (Lozupone & Knight, 2005) of phylogenetic (or functional) dissimilarity between both plots (obtained using the phylogenetic or functional relationships in Fig. S1.1) would be higher than the corresponding species-based dissimilarity between the two plots (obtained using the star tree in Fig. S1.2; see results in Table S1.8).

<table>
<thead>
<tr>
<th>Index</th>
<th>Reference</th>
<th>Phylo/functional case</th>
<th>Species case</th>
</tr>
</thead>
<tbody>
<tr>
<td>UniFrac</td>
<td>Lozupone &amp; Knight (2005)</td>
<td>0.493</td>
<td>0.250</td>
</tr>
</tbody>
</table>

Table S1.8. UniFrac index applied to the third case study.

This third case study shows that the UniFrac index does not fulfill the redundancy property.
Part S1-2: Directions for further research - alternative indices that fulfill the redundancy property

We provide below the proof that an additional tree-based index of plot-to-plot dissimilarity studied by Pavoine (2016) fulfills the redundancy property:

Using Nipperess et al. (2010) phylogenetic components \((A, B, C)\) in Sørensen index of plot-to-plot dissimilarity (Sørensen 1948) corresponds to applying the Bray-Curtis index of plot-to-plot dissimilarity to evolutionary units in a phylogenetic tree (for details, see Pavoine, 2016). Both approaches lead to the following formula of plot-to-plot dissimilarity, with the notations taken from Nipperess et al. (2010):

\[
evoD_{\text{Bray-Curtis}} = 1 - \frac{2A}{2A + B + C} = \frac{B + C}{2A + B + C} = \sum_{t \in T} w_t |n_{ij} - n_{ik}| \sum_{t \in T} w_t (n_{ij} + n_{ik}) \tag{S1.1}
\]

\[
evoD_{\text{Bray-Curtis}} = \frac{\sum_{t \in T} w_t \sum_{i \in S_t} (x_{ij} - x_{ik})}{\sum_{t \in T} w_t \sum_{i \in S_t} (x_{ij} + x_{ik})} = \frac{\sum_{t \in T} w_t \sum_{i \in S_t} (x_{ij} - x_{ik})}{\sum_{i \in N} h_i (x_{ij} + x_{ik})} \tag{S1.2}
\]

\(T\) is a rooted phylogenetic tree with species as tips. \(t\) is a branch of the tree. \(w_t\) is the length of branch \(t\). \(x_{ij}\) is the abundance of species \(i\) in community \(j\). \(S_t\) is the set of species that descend from branch \(t\). \(n_{ij}\) is the summed abundance of all species descending from branch \(t\). \(h_i\) is the sum of branch lengths between tip (species) \(i\) and the root of the phylogenetic tree. \(N\) is the total number of species.

Thm: Index \(evoD_{\text{Bray-Curtis}}\) applied to an ultrametric tree fulfills the redundancy property

Proof:

\[
\sum_{t \in T} w_t \sum_{i \in S_t} (x_{ij} - x_{ik}) \leq \sum_{t \in T} w_t \sum_{i \in S_t} (x_{ij} - x_{ik}) \tag{S1.3}
\]

which yields

\[
evoD_{\text{Bray-Curtis}} \leq \frac{\sum_{i \in N} h_i (x_{ij} - x_{ik})}{\sum_{i \in N} h_i (x_{ij} + x_{ik})} \tag{S1.4}
\]

By definition, for ultrametric trees, \(h_i\) is a constant over \(i\) (for all \(i, h_i = h\), where \(h > 0\)). This implies that

\[
evoD_{\text{Bray-Curtis}} \leq \frac{\sum_{i \in N} (x_{ij} - x_{ik})}{\sum_{i \in N} (x_{ij} + x_{ik})} \tag{S1.5}
\]

and thus that this phylogenetic version of the Bray-Curtis index \((evoD_{\text{Bray-Curtis}})\) is never higher than the original version of the Bray-Curtis index applied to species abundances.
As discussed in the main text, $evoD_{Bray–Curtis}$, as a tree-based index, could also be applied to an ultrametric functional tree, also fulfilling the redundancy property. In addition, equation (S1.4) also implies that, for non-ultrametric trees, the value of this phylogenetic version of the Bray-Curtis index ($evoD_{Bray–Curtis}$) applied to a real tree is never higher than its value when it is applied to a star-shaped phylogenetic tree with the same distance ($h_i$) between a species ($i$) and the root of the tree as the original tree.

This index belongs to a family of indices of phylogenetic dissimilarity between plots where species are replaced by evolutionary units in traditional species-based dissimilarity indices (see Pavoine 2016 for a list of such indices). The indices of this family could thus be systematically tested to verify whether they fulfill the redundancy property.

**An important point to be aware of is that some functional or phylogenetic dissimilarity indices which fulfill the redundancy property may have other unexpected properties.**

As an example, we also analyzed the dissimilarity indices suggested by Swenson (2011) ($D_{nn}$, $D_{nn}'$, $D_{pw}$, $D_{pw}'$, Rao’s $D$) and the related indices ($D_{CW}$, $D_{IP}$). These indices are defined as follows:

$$D_{nn} = \frac{\sum_{i=1}^{nA} \min d_{ib} + \sum_{j=1}^{nB} \min d_{jA}}{2}$$

where $n_A$ and $n_B$ are the number of species in plot $A$ and $B$, respectively, $\min d_{ib}$ is the dissimilarity from the nearest phylogenetic (or functional) neighbor in plot $B$ to species $i$ that occurs in plot $A$ and $\min d_{jA}$ is the dissimilarity from the nearest phylogenetic (or functional) neighbor in plot $A$ to species $j$ that occurs in plot $B$.

$D_{nn}$ is closely related to the following two other plot-to-plot dissimilarity indices:

$$D_{CW} = \frac{1}{2} \left( \frac{\sum_{i=1}^{nA} \min d_{ib}}{n_A} + \frac{\sum_{j=1}^{nB} \min d_{jA}}{n_B} \right)$$

by Clarke & Warwick (1998) and

$$D_{IP} = \frac{1}{2} \left( \frac{\sum_{i=1}^{nA} \min d_{ib} + \sum_{j=1}^{nB} \min d_{jA}}{n_A + n_B} \right)$$

by Izsak & Price (2001). A weighted version of $D_{CW}$ is (Ricotta & Burrascano 2008):

$$D'_{nn} = \frac{\sum_{i=1}^{nA} p_{iA} \min d_{ib} + \sum_{j=1}^{nB} p_{jB} \min d_{jA}}{2}$$
where \( p_{iA} \) is the relative abundance of species \( i \) in plot \( A \) and \( p_{jB} \) is the relative abundance of species \( j \) in plot \( B \). Using average dissimilarities instead of minimum dissimilarities leads to:

\[
D_{pw} = \frac{\sum_{i=1}^{nA} \bar{d}_{iB} + \sum_{j=1}^{nB} \bar{d}_{jA}}{2}
\]

where \( \bar{d}_{iB} \) is the average phylogenetic (or functional) dissimilarity between species \( i \) that occurs in plot \( A \) and all species that occur in plot \( B \) and \( \bar{d}_{jA} \) is the average phylogenetic (or functional) dissimilarity between species \( j \) that occurs in plot \( B \) and all species that occur in plot \( A \). A modified version of \( D_{pw} \) was proposed by Ricotta et al. (2015):

\[
D_{pw}^* = \frac{1}{2} \left( \sum_{i=1}^{nA} \frac{1}{n_A} \bar{d}_{iB} + \sum_{j=1}^{nB} \frac{1}{n_B} \bar{d}_{jA} \right)
\]

A weighted version of this index is given by:

\[
D_{pw}' = \frac{\sum_{i=1}^{nB} p_{iB} \bar{d}_{iA} + \sum_{j=1}^{nA} p_{jA} \bar{d}_{jB}}{2}
\]

Swenson (2011) also suggested the use of an additional index called Rao's \( D \)

\[
Rao's \ D = \sum_{i=1}^{nA} \sum_{j=1}^{nB} p_{iA} p_{jB} d_{ij}
\]

where \( d_{ij} \) is the dissimilarity between species \( i \) in plot \( A \) and \( j \) in plot \( B \).

One can easily demonstrate that all these indices (\( D_{nn}, D_{nn}', D_{pw}, D_{pw}', Rao's \ D, D_{CW}, D_{IP} \) and \( D_{pw}' \)) fulfill the redundancy property (as \( 0 \leq d_{ij} \leq 1 \)). However, they all have limitations as measures of plot-to-plot dissimilarity. First, \( D_{pw}, D_{pw}', D_{pw}^* \) and Rao's \( D \) can be positive when comparing two identical plots, whereas, by definition, the dissimilarity between a plot and itself should be zero (Ricotta et al., 2015).

As regards, \( D_{nn}, D_{nn}', D_{CW}, \) and \( D_{IP} \), in Ricotta et al. (2015) we have already shown that, as nearest-neighbor indices, they may lead to unexpected results. For example, consider two plots with 10 species each for a total of 20 species (the plots do not share any species). Consider that the species are characterized by their body mass. Plot \( A \) contains the species represented by black circles in Fig. S1.3, while plot \( B \) contains the species represented by red squares.

Assuming that all species have even abundances in each plot, for the indices \( D_{nn}, D_{nn}', D_{CW}, \) and \( D_{IP} \) scenario#1 in Fig. S1.3 would lead to lower plot-to-plot dissimilarity than scenario#2, although we would intuitively expect the two plots in scenario#1 to be more dissimilar than the plots in scenario#2 (see results in Table S1.9). Indeed, in scenario#1 plot \( A \) mainly contains species with low body mass, while plot \( B \) mainly contains species with high body mass. In contrast, in scenario#2, the species in both plots have a balanced distribution of body mass with half species having low body mass and half species having high body mass. The unexpected
behavior of these indices is inherent to the fact that they only retain nearest-neighbor distances between species in their calculation.

As mentioned in the main text, further studies should thus consider a range of properties fulfilled by any potential index of functional or phylogenetic dissimilarity between plots, rather than focusing solely on the redundancy property.

Fig. S1.3. Two scenarios of functional differences between plots. In Scenario#1, plot A (black circles) contains 9 species with low body mass and 1 species with high body mass, while plot B (red squares) contains 1 species with low body mass and 9 species with high body mass. In Scenario#2, both plots A (black circles) and B (red squares) contain 5 species with low body mass and 5 species with high body mass, though plot B tends to have, on average, species with higher body mass than species from plot A. For both scenarios, we calculated the indices $D_{nn}$, $D'_{nn}$, $D_{CW}$, and $D_{IP}$, assuming that all species have even abundances in each plot (so that $D'_{nn} = D_{CW}$). The functional distances between species were calculated as the Euclidean distances between the body mass of all species pairs. These distances were then bounded between 0 and 1 by dividing them by the highest observed distance. Results are given in Table S1.9.

<table>
<thead>
<tr>
<th>Index</th>
<th>Scenario#1</th>
<th>Scenario#2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{nn}$</td>
<td>0.544</td>
<td>0.665</td>
</tr>
<tr>
<td>$D'<em>{nn} = D</em>{CW}$</td>
<td>0.054</td>
<td>0.067</td>
</tr>
<tr>
<td>$D_{IP}$</td>
<td>0.027</td>
<td>0.033</td>
</tr>
</tbody>
</table>

Table S1.9. Plot-to-plot dissimilarity indices applied to the two scenarios described in Fig. S1.3.

References


Appendix S2. Redundancy property of $P_D$

Here we show that the plot-to-plot Bray-Curtis phylogenetic dissimilarity $D_P$ as defined in Equation (2) of main text is always lower than the plot-to-plot Bray-Curtis species dissimilarity $D_K = \sum_{i=1}^{N} |x_{Ai} - x_{Bi}|/\left(\sum_{i=1}^{N} (x_{Ai} + x_{Bi})\right)$, where $x_{Ai}$ and $x_{Bi}$ denote the abundances of species $i$ $(i = 1, 2, ..., N)$ in plots $A$ and $B$, respectively.

Recalling that lineage dissimilarity $D_k$ between $A$ and $B$ for each node age $t_k$ is defined as $D_k = \sum_{m=1}^{M_k} |x_{Akm} - x_{Bkm}|/\left(\sum_{m=1}^{M_k} (x_{Akm} + x_{Bkm})\right)$, where $x_{Akm}$ and $x_{Bkm}$ denote the abundances in plot $A$ and $B$ of lineage $m$ $(m = 1, 2, ..., M_k)$ for the $k$-th evolutionary period, we first prove the following lemma:

**Lemma:** The dissimilarities $D_k$, $k = 1,...,K$ form a monotonically increasing sequence, i.e. $D_1 \leq D_2 \leq ... \leq D_K$.

**Proof of lemma.**
Let us consider an evolutionary period $k$, $1 \leq k \leq K - 1$. The abundances of each lineage along the phylogeny are calculated by summing the abundances of all descendent species in the phylogeny. Therefore, for every $k$-th evolutionary period, we have:

$$\sum_{m=1}^{M_k} (x_{Akm} + x_{Bkm}) = \sum_{i=1}^{N} (x_{Ai} + x_{Bi}) \quad (S2.1)$$

For $1 \leq m \leq M_k$, denote $H_{k+1,m}$ the set of all species originated from the species $(k,m)$ at period $k + 1$. By definition, we have $\sum_{j \in H_{k+1,m}} x_{A,k+1,j} = x_{A,k,m}$ and $\sum_{j \in H_{k+1,m}} x_{B,k+1,j} = x_{B,k,m}$. Then, using the triangular inequality:

$$\sum_{m=1}^{M_k} |x_{Akm} - x_{Bkm}| = \sum_{m=1}^{M_k} \left| \sum_{j \in H_{k+1,m}} (x_{A,k+1,j} - x_{B,k+1,j}) \right| \leq \sum_{m=1}^{M_k} \sum_{j \in H_{k+1,m}} |x_{A,k+1,j} - x_{B,k+1,j}| \quad (S2.2)$$

from (S2.1) and (S2.2) immediately follows that
\[ D_k = \frac{\sum_{m=1}^{M_k} |\chi_{A km} - \chi_{B km}|}{\sum_{m=1}^{M_k} (\chi_{A km} + \chi_{B km})} \leq \frac{\sum_{m=1}^{M_{k+1}} |\chi_{A, k+1, m} - \chi_{B, k+1, m}|}{\sum_{m=1}^{M_{k+1}} (\chi_{A, k+1, m} + \chi_{B, k+1, m})} = D_{k+1} \]  

(S2.3)

Since (S2.3) is true for all evolutionary periods \( k, 1 \leq k \leq K - 1 \), we have: \( D_1 \leq D_2 \leq \ldots \leq D_K \).

\[ \square \]

Recalling that \( D_p = \sum_{k=1}^{K} D_k \times \tau_k \), and \( \sum_{k=1}^{K} \tau_k = 1 \), the lemma implies that:

\[ D_p \leq \sum_{k=1}^{K} D_k \tau_k = D_k \sum_{k=1}^{K} \tau_k = D_k \]  

(S2.4)

which proves that the plot-to-plot Bray-Curtis phylogenetic dissimilarity \( D_p \) as defined in Equation (2) of the main text is always lower than the plot-to-plot Bray-Curtis species dissimilarity.

Note that the proof is valid irrespective of the values of the evolutionary periods \( \Delta_k = |t_k - t_{k-1}| \) with \( \sum_{k=1}^{K} \Delta_k = T \), while the number of species originated from any species \((k, m)\) at period \( k + 1 \) may be greater than 2. The weighting factors \( \tau_k \) may be also arbitrary with \( \sum_{k=1}^{K} \tau_k = 1 \).
Appendix S3: R scripts and examples

The R function “DP” calculates tree-based dissimilarity between pairs of plots taking into account the functional or phylogenetic dissimilarities between species. With this function, tree-based dissimilarity is calculated according to Equation 2 in the main text with the Bray-Curtis index. This program is free software: you can redistribute it and/or modify it under the terms of the GNU General Public License [http://www.gnu.org/licenses/](http://www.gnu.org/licenses/). It will be integrated in version 1.4 of the adiv package of R. [https://cran.r-project.org/web/packages/adiv/index.html](https://cran.r-project.org/web/packages/adiv/index.html).

Disclaimer: users of this code are cautioned that, while due care has been taken and it is believed accurate, it has not been rigorously tested and its use and results are solely the responsibilities of the user.

Description: given a matrix of $N$ plots × $S$ species’ relative or absolute abundance value, together with an $S \times S$ dissimilarity matrix, the function “DP” calculates a semimatrix with the values of the dissimilarity index for each pair of plots, as proposed in the main text.

Dependencies: package adiv.

Usage: DP(mtree, comm, height = NULL, diag = FALSE, upper = FALSE, tol = 0.001)

Arguments:

mtree: an object inheriting the class phylo (see package ape), phylo4 (see package phylobase), or hclust. The tree must be ultrametric: equal distance from any tip to the root.

comm: a matrix of $N$ plots × $S$ species containing the relative or absolute abundance of all species. Columns are species and plots are rows. Column labels (species names) should be assigned as in mtree.

height: either NULL or a numeric. See details.

diag: a logical value indicating whether the diagonal of the distance matrix should be printed by function print.dist.

upper: a logical value indicating whether the upper triangle of the distance matrix should be printed by function print.dist.

 tol: a tolerance threshold. A value between $-\text{tol}$ and $\text{tol}$ is considered as zero. See details.

Value: The function returns a (semi-)matrix of class dist with the values of the proposed dissimilarities for each pair of plots.

Details: Object mtree defines a tree with species as tips. If argument height is NULL, then the root of the tree will be placed at the most recent common ancestor of all species occurring in the set of plots (given in object comm). An alternative position for the root can be given by specifying the height of the tree (argument height). In that case, height must be higher than the distance between tips and the most recent common ancestor of all species.

The tolerance threshold tol is particularly important if your tree is not exactly ultrametric due to approximation problems. In that case, the distance from tip to root varies according to the tip considered, although it should not (variations are due to approximation problems). A difference smaller than tol in the distance to root for two species will thus be considered as null.
Function Syntax:

DP <- function(mtree, comm, height = NULL, diag = FALSE, upper = FALSE, tol = 0.001){
  if(!(inherits(comm, "matrix") | inherits(comm, "data.frame")))
    stop("comm must be a matrix or a data frame")
  ncom <- nrow(comm)
  if(is.null(colnames(comm)))
    stop("comm must have names for column")
  if(ncom < 2)
    stop("At least two rows for comm are required")
  if(is.null(colnames(comm)))
    stop("comm must have names for column")
  TA <- tecAptree(mtree, tol = tol)
  if(any(!colnames(comm) %in% names(TA$list[[1]])))
    stop("comm contains tip names that are not available in mtree")
  if(any(comm < 0))
    stop("comm should contain nonnegative values")
  if(any(rowSums(comm) == 0))
    stop("empty communities should be discarded in comm")
  sp_names <- names(TA$list[[1]])
  comm <- comm[, sp_names]
  FUN_COM <- function(groups){
    COM <- apply(comm, 1, function(x) tapply(x, groups, sum))
    return(t(COM))
  }
  FUN_BC <- function(tab){
    d <- matrix(0, ncom, ncom)
    funBC <- function(x) {
      p <- tab[x[1], ]
      q <- tab[x[2], ]
      ps <- p[(p + q) > 0]
      qs <- q[(p + q) > 0]
      w <- sum(abs(ps - qs))/sum(ps + qs)
      return(w)
    }
    index <- cbind(col(d)[col(d) < row(d)],
                   row(d)[col(d) < row(d)])
    d <- unlist(apply(index, 1, funBC))
    return(d)
  }
  LISTCOM <- lapply(TA$list, FUN_COM)
  LISTd <- lapply(LISTCOM, FUN_BC)
  d <- LISTd[[1]] * TA$plength[1]
  for(i in 2:length(LISTd)) {
    d <- d + LISTd[[i]] * TA$plength[i]
  }
  if(!is.null(height) && is.numeric(height) && height > sum(TA$plength))
    d <- d / height
  else
    d <- d / sum(TA$plength)
  attr(d, "Size") <- ncom
  attr(d, "Labels") <- rownames(comm)
  attr(d, "Diag") <- diag
attr(d, "Upper") <- upper
attr(d, "method") <- "DP"
attr(d, "call") <- match.call()
class(d) <- "dist"
return(d)
}

Example:
Load in R the table contained in the Appendix S2 of Ricotta et al. (2016) with species names as row names and name the table FullTab. Take care to put an underscore between Genus and species names. The first six rows of this table will thus be:

<table>
<thead>
<tr>
<th>X1</th>
<th>X2</th>
<th>X3</th>
<th>X4</th>
<th>X5</th>
<th>X6</th>
<th>X7</th>
<th>X8</th>
<th>X9</th>
<th>X10</th>
<th>X11</th>
<th>X12</th>
<th>X13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Achillea_moschata</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Adenostyles_leucophylla</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Agrostis_rupestris</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Agrostis_schraderiana</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Antennaria_dioica</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Anthoxanthum_odoratum</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

X14 X15 X16 X17 X18 X19 X20 X21 X22 X23 X24

| Achillea_moschata | 0   | 0   | 0   | 0   | 2   | 1   | 2   | 0   | 1   | 2   | 2   | 2   |
| Adenostyles_leucophylla | 0   | 0   | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| Agrostis_rupestris | 0   | 0   | 2   | 0   | 0   | 1   | 0   | 3   | 2   | 0   | 0   | 0   |
| Agrostis_schraderiana | 0   | 0   | 0   | 0   | 0   | 0   | 2   | 0   | 0   | 0   | 0   | 0   |
| Antennaria_dioica | 0   | 0   | 0   | 0   | 1   | 1   | 0   | 0   | 0   | 0   | 0   | 0   |
| Anthoxanthum_odoratum | 0   | 0   | 0   | 0   | 3   | 4   | 1   | 1   | 1   | 1   | 0   | 0   |

X25 X26 X27 X28 X29 X30 X31 X32 X33 X34 X35

| Achillea_moschata | 0   | 4   | 3   | 0   | 1   | 2   | 2   | 3   | 2   | 2   | 2   | 2   |
| Adenostyles_leucophylla | 0   | 0   | 0   | 3   | 0   | 3   | 2   | 2   | 2   | 2   | 2   | 2   |
| Agrostis_rupestris | 2   | 3   | 1   | 2   | 4   | 1   | 1   | 2   | 2   | 1   | 3   | 3   |
| Agrostis_schraderiana | 0   | 0   | 0   | 0   | 0   | 1   | 0   | 0   | 0   | 2   | 0   | 0   |
| Antennaria_dioica | 0   | 0   | 0   | 0   | 2   | 1   | 1   | 0   | 1   | 1   | 0   | 0   |
| Anthoxanthum_odoratum | 2   | 1   | 3   | 1   | 2   | 1   | 2   | 1   | 2   | 1   | 2   | 2   |

X36 X37 X38 X39 X40 X41 X42 X43 X44 X45 X46

| Achillea_moschata | 2   | 2   | 2   | 1   | 2   | 2   | 2   | 1   | 3   | 3   | 1   | 2   |
| Adenostyles_leucophylla | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| Agrostis_rupestris | 2   | 2   | 2   | 2   | 2   | 1   | 0   | 0   | 0   | 0   | 0   | 0   |
| Agrostis_schraderiana | 0   | 0   | 0   | 0   | 0   | 0   | 2   | 0   | 0   | 0   | 0   | 0   |
| Antennaria_dioica | 1   | 0   | 0   | 1   | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   |
| Anthoxanthum_odoratum | 1   | 1   | 0   | 3   | 2   | 3   | 1   | 2   | 1   | 0   | 1   | 0   |

X47 X48 X49 X50 X51 X52 X53 X54 X55 X56 X57

| Achillea_moschata | 0   | 2   | 3   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| Adenostyles_leucophylla | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| Agrostis_rupestris | 1   | 0   | 1   | 0   | 0   | 0   | 1   | 2   | 0   | 0   | 2   | 2   |
| Agrostis_schraderiana | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| Antennaria_dioica | 1   | 0   | 0   | 0   | 0   | 1   | 2   | 1   | 0   | 1   | 0   | 0   |
| Anthoxanthum_odoratum | 1   | 0   | 4   | 2   | 2   | 2   | 2   | 2   | 2   | 2   | 1   | 2   |

X58 X59 C S R

| Achillea_moschata | 0   | 0   | 38.57 | 23.81 | 37.62 |
| Adenostyles_leucophylla | 0   | 0   | 49.59 | 0.00  | 50.41 |
| Agrostis_rupestris | 1   | 1   | 17.45 | 65.26 | 17.29 |
| Agrostis_schraderiana | 0   | 0   | 41.97 | 39.47 | 18.56 |
| Antennaria_dioica | 0   | 0   | 28.92 | 56.75 | 14.34 |
| Anthoxanthum_odoratum | 0   | 0   | 17.65 | 60.53 | 21.82 |

The species traits are in the columns 60, 61 and 62 of this table (columns named "C", "S", and "R"):

Traits <- FullTab[60:62]

The abundances of species in plots are in the remaining columns:

ab <- FullTab[(-60:62)]
ab <- t(ab)

In ab, plots are in rows and species in columns.
Then we need to perform the functional clustering with the unweighted pair group method with arithmetic mean (UPGMA) on the Bray-Curtis distance applied to species’ traits.

```r
library(vegan)
library(ape)
H <- hclust(vegdist(Traits), "average")  # By default, function vegdist in package vegan calculates the Bray Curtis coefficient.
```

With function `read.tree` of package `ape`, load in R the phylogenetic tree contained in the Appendix A of Ricotta et al. (2015) and name it `phy`.

The tree is as follows:

```
Load function DP in R.

The species-based dissimilarities between plots are:
Ds <- vegdist(ab)

The functional dissimilarities are:
library(adiv)
Df <- DP(H, ab, height=0.5, tol=0.00001)

and the phylogenetic dissimilarities are:
Dp <- DP(phy, ab, tol=0.00001)
```
In function DP, the parameter "height" allows to control the scaling procedure to unit height from the tips to the root. The default was used for the phylogenetic tree which means that the root of the tree was placed at the most recent common ancestor of all species occurring in the set of plots, and the distance from tips to root was set to unity. In the UPGMA tree obtained above, the functional dissimilarity between two species is estimated by twice the way between each of these species and their first shared interior node, which is equal to the sum of branch lengths in the shortest path that connects the two species on the tree. If \( d_{\text{max}} \) is the maximum possible functional dissimilarity between two species, then the root of the tree must thus be placed at a distance of \( d_{\text{max}}/2 \) from the tips. To normalize the tree height, the branch lengths on the tree must then be divided by \( d_{\text{max}}/2 \) so that the height of the tree is equal to 1. We used the Bray-Curtis index to calculate functional dissimilarities between species. This index is bounded between 0 and 1, so that \( d_{\text{max}}=1 \) and \( d_{\text{max}}/2=0.5 \) (height=0.5). The dissimilarity matrices DF and DP can then be included in further analyses, such as the analysis of variance using distance matrices (function adonis in package vegan). In that case, a factor defining in which successional stage each plot belongs to can be defined as follows:

```r
FAC <- rep(c("early", "mid", "late"), c(17, 32, 10))
```

In FAC, "early" means "Early-successional stage", "mid" means "Mid-successional stage" and "late" means "Late-successional stage".

The functional and phylogenetic uniqueness values can be computed as follows:

```r
Uf <- DF/DS
Up <- DP/DS
```

Also note that the square root of these matrices (DS, DF and DP) are Euclidean and can be analyzed via a principal coordinate analysis (e.g. function dudi.pco in package ade4).

All data will be included in version 1.4 of the adiv package. We will provide the following name to the data set in adiv: RutorGlacier.

**References**


Appendix S4
Phylogenetic tree of the 45 species used in this study. The age of the root node is 147.8 Million years (Ma)

UPGMA dendrogram of the same 45 species obtained by applying the Bray-Curtis (BC) dissimilarity to the CSR strategies. The root node of the classification is located at BC = 0.461