Research

Oikos 117: 794–800, 2008
doi: 10.1111/j.2008.0030-1299.16594.x,

Subject Editor: Nicholas Gotelli, Accepted 1 February 2008

Towards a consensus for calculating dendrogram-based functional diversity indices

Maud Mouchet, François Guilhaumon, Sébastien Villéger, Norman W. H. Mason, Jean-Antoine Tomasini and David Mouillot

M. Mouchet (maud.mouchet@univ-montp2.fr), F. Guilhaumon, S. Villéger, J.-A. Tomasini and D. Mouillot, UMR CNRS-IFREMER-UM2 5119 Écosystèmes Lagunaires, Univ. Montpellier 2 cc 093, FR-34 095 Montpellier Cedex 5, France. – N. W. H. Mason, Landcare Research, Private Bag 3127, Hamilton 3240, New Zealand.

The widely used FD index of functional diversity is based on the construction of a dendrogram. This index has been the subject of a strong debate concerning the choice of the distance and the clustering method to be used, since the method chosen may greatly affect the FD values obtained. Much of this debate has been centred around which method of dendrogram construction gives a faithful representation of species distribution in multidimensional functional trait space. From artificially generated datasets varying in species richness and correlations between traits, we test whether any single combination of clustering method(s) and distance consistently produces a dendrogram that most closely corresponds to the matrix of functional distances between pairs of species studied. We also test the ability of consensus trees, which incorporate features common to a range of different dendrograms, to summarize distance matrices. Our results show that no combination of clustering method(s) and distance constantly outperforms the others due to the complexity of interactions between correlations of traits, species richness, distance measures and clustering methods. Furthermore, the construction of a consensus tree from a range of dendrograms is often the best solution. Consequently, we recommend testing all combinations of distances and clustering methods (including consensus trees), then selecting the most reliable tree (with the lowest dissimilarity) to estimate FD value. Furthermore we suggest that any index that requires the construction of functional dendrograms potentially benefits from this new approach.

Functional diversity has been identified as a key component of biodiversity for ecosystem functioning and sustainability (Ives et al. 1999, Hooper and Dukes 2004, Petchey et al. 2004). In this way, biodiversity loss can be responsible for a decrease in resilience or productivity (Loreau et al. 2001, Bellwood et al. 2004, Petchey 2004). Despite the emerging importance of the subject, no agreement exists about quantifying functional diversity of a community of species, though many indices have been recently proposed (reviewed by Petchey and Gaston 2006). Indeed, there has been a great deal of controversy over the statistical validity of these indices (Petchey and Gaston 2007, Podani and Schmera 2007) as well as their properties and what they actually measure (Ricotta 2005, Petchey and Gaston 2006). The FD index suggested by Petchey and Gaston (2002) provides an illustration of such controversy. FD is the total branch length of a functional dendrogram (hierarchical classification of the species according to their functional features). It measures functional diversity at all ecological scales simultaneously (Petchey and Gaston 2002), and is neither affected by units of characters nor by the splitting of a species into two identical functional species (Mason et al. 2003). Moreover, FD has been applied to a variety of taxa: mammals (Blackburn et al. 2005), plants (Thompson et al. 2005), birds (Petchey et al. 2007), fish (Mouillot et al. 2007) and zooplankton (Barnett et al. 2007), and the frequency of its use is likely to increase.

A dendrogram-based functional classification of species is carried out in three steps: (1) building the trait matrix (containing the value of each species for each functional trait); (2) calculating a matrix of distances between pairs of species in functional trait space; and (3) constructing a dendrogram to classify species according to the distance matrix (Petchey and Gaston 2002). The choices of the distance and of the classification method – steps (2) and (3) respectively – are of crucial importance since they may lead to different results.

Several distances are available in ecology (Bray–Curtis, Manhattan, etc). But Euclidean and Gower distances are the only two ones recommended to measure interspecific functional dissimilarity based on trait values. In fact, Petchey and Gaston (2002) employed Euclidean distance, while Podani and Schmera (2006) advised the use of Gower's distance (which allows missing data and inclusion of qualitative traits) even when all traits are quantitative.

There are also several clustering methods for constructing a hierarchical classification (Legendre and Legendre 1998). Each one of these methods is based on particular criteria to aggregate species into hierarchical clusters. The topology of the resulting tree varies according to the procedure employed for the same initial distance matrix. Consequently, numerical values of the total branch length of the dendrogram (i.e. FD) also depend on the distance and/or clustering algorithm used. However, there is a lack of agreement over which method of constructing the functional dendrogram gives the best representation of species distribution in functional trait space (Podani and Schmera 2006).

Thus, despite the broad use of the FD index, several questions remain to be answered. To what extent does the hierarchical classification represent the distribution of species in functional space? Which is the most faithful method of clustering? How can we improve the representation of the distance matrix by a hierarchical classification? Is there any factor influencing the performance of cluster algorithms? Does the same combination of distance and clustering method(s) always perform best? These questions are all crucial because the choice of a clustering method may strongly influence the functional diversity value obtained for a community. This problem was recently revealed by Podani and Schmera (2006). Indeed, no previous studies have quantitatively tested whether a single clustering method consistently outperforms all others. The choice of a clustering method remains arbitrary in functional ecology and no consensual solution has been yet proposed despite the popularity of FD.

In this paper, we present a novel procedure based on an objective criterion to choose the most appropriate method of constructing hierarchical dendrograms from functional trait data. This method aims to find the combination of distance and clustering algorithm(s) that best represents species distribution in functional trait space. Because each clustering algorithm has its flaws, we suggest also testing consensus trees resulting from a combination of several clustering algorithms to enhance the reliability of the classification and, ultimately, the estimation of functional diversity.

Material and methods

Hierarchical classifications synthesise the multidimensional distribution of objects (e.g. species, communities, assemblages) in a one-dimensional diagram. This procedure results in a loss of information about the dispersion of objects along each dimension summarized, resulting in distortion of data representation (Legendre and Legendre 1998). Our purpose is to find the least altered representation of species in a functional multidimensional space through a new method validated with simulated datasets.

Functional distance based on traits

To construct the distance matrix from artificial communities, we considered two distances. The Euclidean distance between species a and b for N quantitative traits was defined as:

$$ED_{ab} = \sqrt{\sum_{i=1}^{N} (x_{ia} - x_{ib})^2}$$

with x_{ia} and x_{ib} being the values of trait i (variable) for species a and b, respectively.

The Gower's distance can cope with mixed scale types of data (quantitative, interval, nominal or ordinal data, ratios, missing values). The general Gower's formula is given by:

$$GD_{ab} = \frac{\sum_{i=1}^{N} w_{iab} d_{iab}}{\sum_{i=1}^{N} w_{iab}}$$

where d_{iab} measures the dissimilarity between species a and b for the variable i:

$$d_{iab} = \frac{|x_{ia} - x_{ib}|}{\max(x_i) - \min(x_i)}$$

The weight $w_{iab} = 0$ when x_{ia} and/or x_{ib} is missing. In our framework, there was no missing value, trait values were all quantitative and each trait was equally weighted ($w_{iab} = 1$). Consequently, the formula was reduced to:

$$GD_{ab} = \frac{1}{N} \sum_{i=1}^{N} d_{iab}$$

Construction of a consensus tree

There are several clustering methods available for hierarchically clustering species into a dendrogram. In the present study, all methods tested belong to the family of hierarchical agglomerative (or ascending) classifications: single linkage, complete linkage, UPGMA, WPGMA, UPGMC, WPGMC and Ward's method (Table 1). Because clustering procedures fail to perfectly fit data distribution in multidimensional space, we tested whether using a consensus algorithm would enhance the faithfulness of dendrograms. Consensus algorithms make a synthesis from different dendrograms into one classification which highlights the concordant parts between the various methods relative to the parts that disagree (Darlu and Tassy 1993). In short, the algorithm builds the consensus tree that optimally represents several clustering methods. Several methods are available to form a consensus tree. The algorithm adopted here minimized the Euclidean distance between matrices (according to Gordon and Vichi 2001):

$$D_c = \sum_{j=1}^{n} d_{jc}$$

where d_{jc} is the Euclidean distance between the cophenetic distance matrix of dendrogram j and that of the consensus dendrogram c, and n is the number of individual dendrograms being considered.

Although the selected algorithm produces the best projection of the data according to its criteria, the quality

Table 1. Clustering methods tested (according to Legendre and Legendre 1998).

Clustering method	Acronym	Signification
Single linkage	SL	Fusion of the closest objects
Complete linkage	CL	Fusion of the most distant objects
Unweighted pair group method using arithmetic averages	UPGMA	Fusion of clusters when the similarity reaches the mean inter cluster similarity value
Weighted pair group method using arithmetic averages	WPGMA	Mean dissimilarity between two items is the sum of the weighted dissimilarities
Unweighted pair group centroid method	UPGMC	Fusion of the clusters with closest centroids
Weighted pair group centroid method	WPGMC	Fusion of the clusters with closest centroids after adjustment for group sizes
Ward's method	Ward	Minimization of the within-group sum of squares

of this representation can vary (contraction or dilation of space) according to the topology and the branch length of the resulting dendrogram. To assess the reliability of the dendrogram, we evaluated the dissimilarity between initial distance and cophenetic distance matrices, D_M, using the cophenetic correlation (Legendre and Legendre 1998). Cophenetic correlation has already been used by Blackburn et al. (2005) and Petchey et al. (2007) or discussed by Petchey and Gaston (2006, 2007) to evaluate the quality of the dendrogram which provided FD values. In this study, the dissimilarity value, D_M, is given by $1-c^2$; where c stands for the cophenetic correlation coefficient (Pearson correlation coefficient in our case). The objective was thus to identify, among all combinations of distance and clustering algorithm(s), the one that minimized the dissimilarity between the initial distance matrix and the cophenetic distance matrix ($D_M = 0$ corresponds to no distortion between distance matrices).

Theoretical datasets

To test this new flexible method in many different situations, we created artificial communities in order to control two parameters: species richness and correlation between functional traits. Because FD generally measures the functional diversity of a community using a dendrogram including all species present in the regional pool (hereafter called species assemblage) (Petchey and Gaston 2002), we defined two reasonable species richness levels for our simulated assemblages: 20 and 40 species. All artificial species assemblages were characterized by five simulated standardised functional traits sampled among a normal distribution with a mean of 5 and a standard deviation of 1. Traits were either forced to be all independent, or four of them were constrained to be highly correlated (Pearson correlation coefficient higher than 0.9) with the fifth being entirely independent of the others. We simulated one hundred assemblages for every combination of species richness levels and correlation levels (40 species with 5 independent traits; 40 species with 1 independent + 4 correlated traits; 20 species with 5 independent traits; 20 species with 1 independent + 4 correlated traits), giving a total of 400 artificial species assemblages.

Data analyses

The 7 'one-method' clustering algorithms and the 120 potential consensus trees (from 2 to 7 combined trees) multiplied by 2 distances made a total of 254 combinations of distance-clustering (consensus and 'one-method') algorithms that have been tested without any selection of methods, on each artificial assemblage. The D_M value of the 254 combinations was thus calculated for every species assemblage. Among all the 254 possible combinations of distances and clustering algorithms, the one with the lowest D_M value gave the most faithful picture of data distribution in space and thus was chosen as the most appropriate to calculate FD index. Then we retained the D_M value of the most reliable dendrogram (i.e. corresponding to the combination with the lowest dissimilarity) of every assemblage of each dataset to calculate the mean D_M of the sets. The significance of all main effects, and all possible interactions between factors (distances, clustering methods, correlations between traits and species richness) on DM were explored using a four factor ANOVA.

The effects of distances, clustering methods, correlations between traits and species richness on D_M were explored using a four factor ANOVA.

All calculations, including analyses of variances, datasets creation and tree constructions (consensus or not), were implemented using the R statistical environment (R Development Core Team 2006). Construction of consensus trees and dissimilarity tests were performed with, respectively, functions 'cl_consensus' and 'cl_dissimilarity' from the Clue package (Hornik 2005, Cluster ensembles, R package ver. 0.3–13). An R script is available from the authors to compute the different trees and to choose the most appropriate method (with the lowest dissimilarity) to built a functional tree.

Results

The effects of species richness, trait correlations, clustering methods and distances were estimated on the level of dissimilarity between initial distance and cophenetic distance matrices (D_M). The four factor ANOVA revealed that D_M varied significantly with the clustering method, the level of correlation among functional traits and species richness but the effect of distances was the less important (Table 2). Because most interactions were significant, the influence of single factors was, in general, dependent on the others. In other words, D_M was influenced by all the factors and their interactions. For example, when the traits were independent, the distance which provided the best tree (lowest D_M) was consistently Euclidean distance (with

Table 2. Effects of clustering methods, distances, levels of correlations between traits and species richness on dissimilarity D_M tested by a four factor ANOVA. F values with associated levels of significances (ns: non significant, *: $p<0.05$, **: $p<0.01$). (Cor: correlation; Met: clustering method; Dist: distance; SR: species richness; ×: interaction).

Factors	DF	F
Met	6	379.06**
Cor	1	3803.32**
SR	1	651.71**
Dist	1	12.25**
Met × Cor	6	32.67**
Met × SR	6	6.76**
Met × Dist	6	5.02**
Cor × SR	1	9.68**
Cor × Dist	1	247.72**
SR × Dist	1	5.29*
Met × Cor × SR	6	2.65*
Met × Cor × Dist	6	2.11*
Met × SR × Dist	6	0.91ns
Cor × SR × Dist	1	0.11ns
Met × Cor × SR × Dist	6	0.31ns

UPGMA) whereas, when they were correlated, Gower distance consistently provided the best result (still with UPGMA) (Fig. 1, Table 3a). Furthermore, higher species richness tended to decrease D_M (Fig. 1). However the number of clustering algorithms giving the lowest D_M value differed from one case to the other. For instance, the assemblage with the lowest D_M among the dataset 40 Indep (i.e. 100 assemblages of 40 species characterized by five independent traits) involved five clustering algorithms but only two algorithms in the other datasets 40 Cor, 20 Cor and 20 Indep (Table 3b). However, the complexity of interactions between factors made identification of the direct effects of each factor (distance measures, clustering algorithms, species levels, and correlations between traits) on D_M, difficult. Consequently, no combination of distance and clustering methods gave systematically the best representation of species from the multidimensional functional space. Each clustering algorithm and distance was, at least once, used in the best representation of data. Nevertheless, among all combinations, UPGMA/UPGMC consensus trees seemed to be most often used (Table 3a).

In addition, the comparison among dendrograms built with different combinations from the same sample (a 20 species assemblage characterized by five independent traits) illustrated the differences in species clustering provided by different combinations (Fig. 2). In our particular case, the lowest dissimilarity was obtained using the combination of Euclidean distance and consensus tree resulting from UPGMA/UPGMC/single linkage cluster algorithms ($D_M = 0.27$, Fig. 2a). The combination Euclidean distance – UPGMA (Fig. 2b) obtained an intermediate dissimilarity level ($D_M = 0.52$). The highest dissimilarity was given by the consensus tree WPGMC/Ward presented in Fig. 2c. In this case, the initial distance matrix (made with Gower's distance) had a dissimilarity value of 0.9 with the

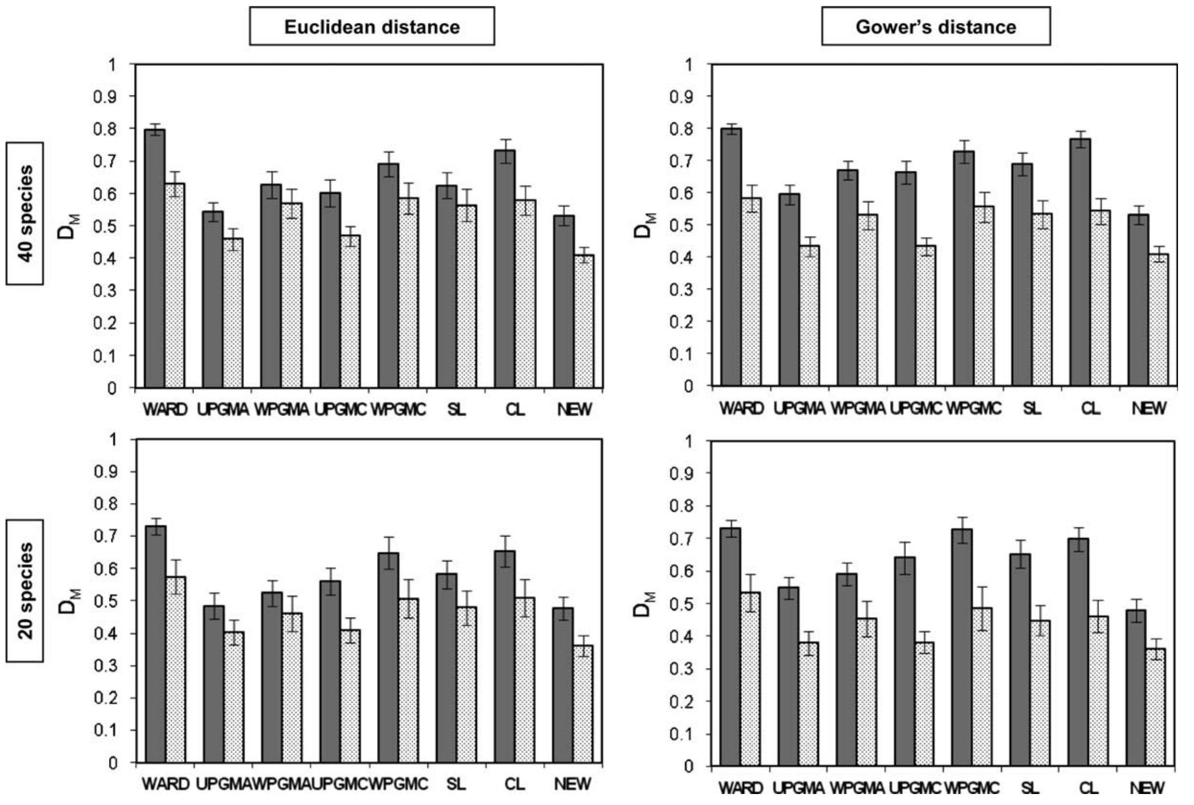

Fig. 1. Impact of clustering algorithms on dissimilarity. D_M values (y-axis) from 4 virtual datasets of 100 assemblages each. (WARD: Ward's method; SL: single linkage; CL: complete linkage; UPGMA: unweighted arithmetic average clustering; WPGMA: weighted arithmetic average clustering; UPGMC: unweighted centroid clustering; WPGMC: weighted centroid clustering; NEW: the new method proposed in this article). ■ Assemblages characterized by 5 independent traits. ▨ Assemblages characterized by 4 correlated traits and 1 independent trait.

Table 3. Dissimilarity D_M tested on the best combination of distance – clustering algorithm(s) of each of the one hundred species assemblages of the four datasets. (a) Mean, D_{Mmean}, and standard deviation, D_{Msd}, of all datasets calculated with the dissimilarity value of the best combination of each assemblage, proportion of Euclidean and Gower distances involved in the best combinations and of the clustering algorithm(s) that gave the lowest dissimilarity in the greatest percentage of assemblages; (b) combination of distance – clustering method giving the lowest D_M of each species richness and trait correlation levels. (ED: Euclidean distance; GD: Gower's distance; 40 Indep: 40 species with 5 independent traits; 20 Indep: 20 species with 5 independent traits; 40 Cor: 40 species with 1 independent+4 correlated traits; 20 Cor: 20 species with 1 independent+4 correlated traits).

(a)

Datasets	D_{Mmean}	D_{Msd}	Distances	Clustering algorithms
40 Indep	0.533	0.059	ED: 94.1% GD: 5.9%	UPGMA–UPGMC: 34%
20 Indep	0.478	0.07	ED: 90.9% GD: 9.1%	UPGMA–WPGMC: 22%
40 Cor	0.411	0.05	ED: 11.4% GD: 88.6%	UPGMA–UPGMC: 35%
20 Cor	0.361	0.066	ED: 20.8% GD: 79.2%	UPGMA–UPGMC: 24%

(b)

Datasets	D_M	Distances	Clustering algorithms
40 Indep	0.386	ED	UPGMA – UPGMC – CL – WPGMA – SL
20 Indep	0.295	ED	UPGMA – WPGMC
40 Cor	0.275	GD	UPGMA – SL
20 Cor	0.195	GD	UPGMA – UPGMC

cophenetic distance matrix of the tree. Some strong inconsistencies appeared between these three dendrograms (Fig. 2). For instance, the position of species 1 changed drastically among dendrograms. Species 1 was joined together with species 3 using Euclidian distance and UPGMA cluster algorithm while this species was associated to the couple of species 2–5 according to the consensus tree combining UPGMA/UPGMC/single linkage methods. However, the results obtained from the initial distance matrix (here, Euclidean) showed that distance $ED_{1-3} = 5$ was higher than the distances $ED_{2-1} = 4.51$, and $ED_{5-1} = 3.61$. Thus, species 1 was more distant in functional trait space from species 3 than from species 2 and 5, indicating that some methods may incorrectly classify species 1 and 3 together.

Discussion

Indices based on a hierarchical classification are highly dependent on the capacity of the clustering method to accurately represent the distribution of species in a functional space (Podani and Schmera 2007). This critical point deserves to be quantitatively studied. From artificially generated data, we have tested whether any single of seven clustering methods and two distances always performs better than the others, or if it is necessary to choose a "consensus" dendrogram. Taken together (Fig. 1, Table 2, 3), our results show (1) that the characteristics of the species assemblages inconsistently influence the performance of the methods, (2) that there is no general rule for the choice of a distance or a clustering method and (3) that the best representation is often provided by a consensus tree.

Rao and Srinivas (2006) defined a threshold for validity of a dendrogram: the dissimilarity between initial distance and cophenetic distance matrices, D_M, must be less than 0.36 (corresponding to the 0.8 correlation coefficient limit stated by Rao and Srinivas). Some algorithms are well known to cause a distortion of space (Legendre and Legendre 1998). Single linkage tends to contract space by compressing edges of the dendrograms whereas complete linkage tends to dilate data space to produce compact classes (Podani and Schmera 2006). Despite their drawbacks, single linkage and complete linkage must not be ignored because they are regularly included in the consensus trees yielding the lowest D_M (Table 3b). Dendrograms built using the UPGMA method give the lowest D_M values (Fig. 1), that is a better representation of data. Among the algorithms selected in the study, UPGMA is a good compromise between single linkage and complete linkage because it preserves most of the initial distances in dimensionally reduced space. In contrast, the Ward's method, though widely used, produces the least reliable classifications (Fig. 1). Nevertheless, hierarchical classifications are one-dimensional pictures of data dispersed in an n-dimension space. So, none of the clustering algorithms (even UPGMA) perfectly corresponds to the distribution of objects in a multidimensional space. As demonstrated here, the quality of representation offered by clustering methods is also modulated by other factors: level of correlations between traits and species richness (Table 2). Therefore, it is worth gathering information from different clustering methods into consensus trees.

To evaluate the faithfulness of a hierarchical classification, the method presented here relies on cophenetic correlation through a dissimilarity measure (D_M). Cophenetic correlation quantifies the agreement between the initial distance matrix and the cophenetic distance matrix. In other words, cophenetic correlation enables one to choose the best combination of clustering method according to the distance employed. However cophenetic correlation does not take into account the transformation of the trait matrix into a distance matrix. More studies are needed to assess the impact of that transformation and of the use of other distances (Manhattan, Bray–Curtis, etc.). Besides, the faithfulness of a classification may decline with a decrease in the number of functional traits and/or species richness (Sokal et al. 1992).

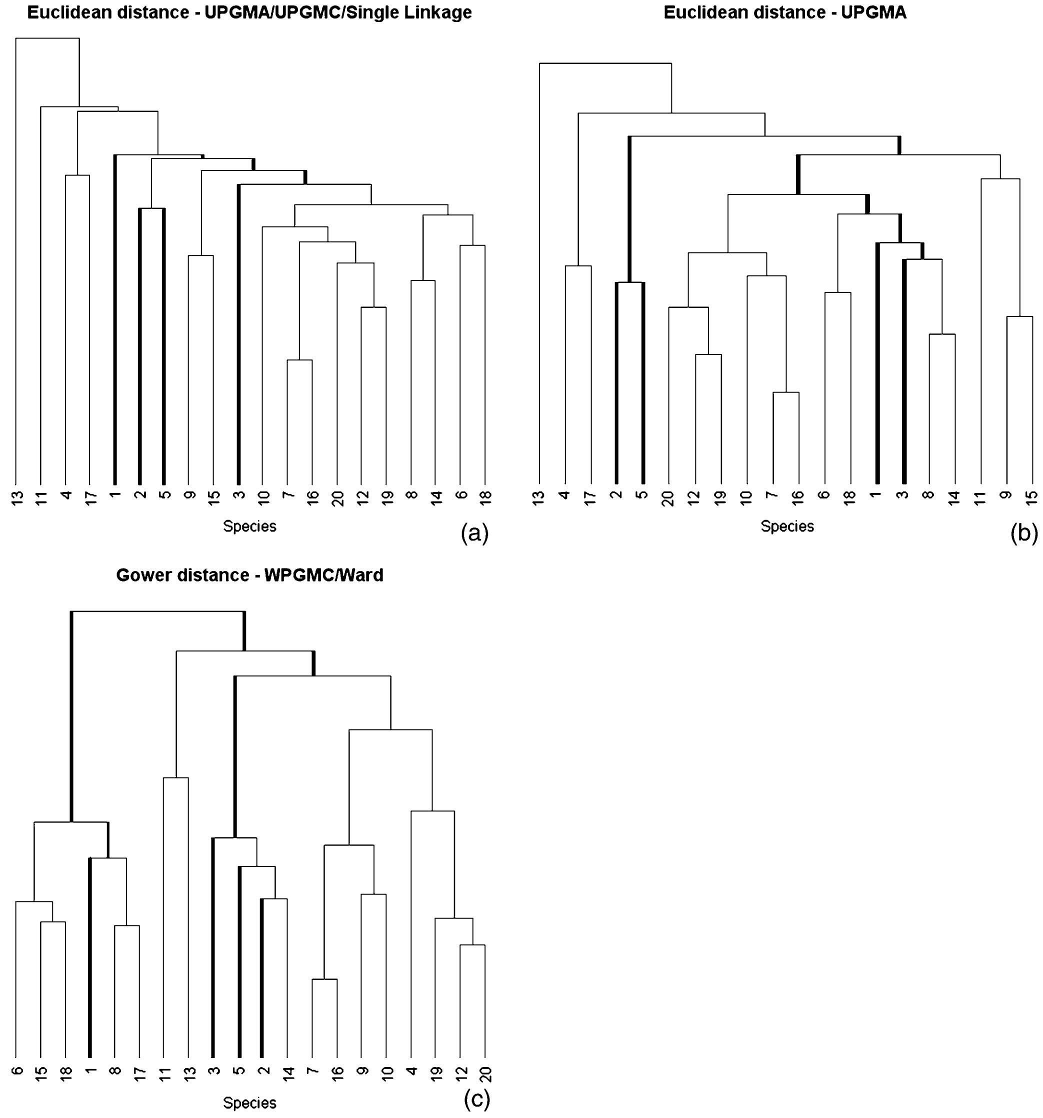

Fig. 2. Hierarchical classifications of an artificial species assemblage composed of 20 species characterized by 5 independent functional traits. (a) with the lowest D_M ($D_M = 0.27$); (b) with an intermediate D_M ($D_M = 0.52$); (c) with the highest D_M ($D_M = 0.9$).

The debate on the calculation of index FD originates from the difficulty of choosing the right combination of distance-clustering method (Podani and Schmera 2006). Our results show that there is not a unique answer, each case being different because of the complexity of interactions between factors such as correlations of traits, species richness, distances and clustering methods. For instance, the choice of the distance is not only related to the type of data. Even if Gower's formula is the only way to cope with mixed data, Euclidean distance must not be excluded when all data are quantitative (Fig. 1). Overall, in our results, the best solution is often given by a consensus tree combining dendrograms produced using several different clustering algorithms. Consensus trees increase the quality of representation of the species on the dendrogram (Fig. 1). They minimize the dissimilarity with the initial distances and preserve the parts common to several trees. As a consequence, FD values obtained will be more accurate. Nevertheless, if the less distorted representation is still very dissimilar from the original distance matrix, FD should not be measured to avoid erroneous interpretations about the functional structure of a community.

Alternatively, we advise against making any a priori choice of a distance or a clustering algorithm but, instead, we propose testing all the possible combinations (including consensus trees) and selecting the one with the lowest dissimilarity value to estimate FD. The choice of the method is therefore objective. Searching for the best combination of distance and clustering algorithm(s) is a good answer to the debate around FD's statistical validity. In addition, our method potentially has a broader 'field' of application than the calculation of index FD. Indeed, the simple classification of species into functional groups is still largely subjective since many alternatives exist for the clustering algorithm. This lack of agreement may weaken the conclusions about the role of functional diversity on ecosystem functioning (Wright et al. 2006). The investigation of the relationship between the shape of functional dendrograms and the patterns of species relative abundances would also benefit from such an objective tool (Sugihara et al. 2003). Finally, the calculation of indices that require the construction of dendrograms to obtain ultrametric distances among species pairs, such as the originality of a species (Pavoine et al. 2005) or the 'turn-over' of communities along environmental gradients (Hardy and Senterre 2007), could also benefit from this method.

Acknowledgements – This study was partially funded by a LITEAU III (PAMPA) and two ANR (GAIUS and AMPHORE) financial supports to study ecological indicators.

References

Barnett, A. J. et al. 2007. Functional diversity of crustacean zooplankton communities: towards a trait-based classification. – Freshwater Biol. 52: 796–813.

Bellwood, D. R. et al. 2004. Confronting the coral reef crisis. – Nature 429: 827–833.

Blackburn, T. M. et al. 2005. Functional diversity of Mammalian predators and extinction in island birds. – Ecology 86: 2916–2923.

Darlu, P. and Tassy, P. 1993. La Reconstruction phylogénétique. Concepts et méthodes. – Masson.

Gordon, A. D. and Vichi, M. 2001. Fuzzy partition for fitting a set of partitions. – Psychometrika 66: 229–248.

Hardy, O. J. and Senterre, B. 2007. Characterizing the phylogenetic structure of communities by an additive partitioning of phylogenetic diversity. – J. Ecol. 95: 493–506.

Hooper, D. U. and Dukes, J. S. 2004. Overyielding among plant functional groups in a long-term experiement. – Ecol. Lett. 7: 95–105.

Hornik, K. 2005. A clue for cluster ensembles. – J. Stat. Software 14. URL: http://www.jstatsoft.org/v14/i12/.

Ives, A. R. et al. 1999. Stability and variability in competitive communities. – Science 286: 542–544.

Legendre, P. and Legendre, L. 1998. Numerical ecology (2nd ed.). – Elsevier.

Loreau, M. et al. 2001. Biodiversity and ecosystem functioning: current knowledge and future challenges. – Science 294: 804–808.

Mason, N. W. H. et al. 2003. An index of functional diversity. – J. Veg. Sci. 14: 571–578.

Mouillot, D. et al. 2007. Limiting similarity, niche filtering and functional diversity in brackish lagoon fish communities. – Estuar. Coast. Shelf Sci. 71: 443–456.

Pavoine, S. et al. 2005. Is the originality of a species measurable? – Ecol. Lett. 8: 579–586.

Petchey, O. L. 2004. On the statistical significance of functional diversity effects. – Funct. Ecol. 18: 297–303.

Petchey, O. L. and Gaston, K. J. 2002. Functional diversity (FD), species richness and community composition. – Ecol. Lett. 5: 402–411.

Petchey, O. L. and Gaston, K. J. 2006. Functional diversity: back to basics and looking forward. – Ecol. Lett. 9: 741–758.

Petchey, O. L. and Gaston, K. J. 2007. Dendrograms and measuring functional diversity. – Oikos 116: 1422–1426.

Petchey, O. L. et al. 2004. How do different measures of functional diversity perform? – Ecology 85: 847–857.

Petchey, O.L. et al. 2007. Low functional diversity and no redundancy in British avian assemblages. – J. Anim. Ecol. 76: 977–985.

Podani, J. and Schmera, D. 2006. On dendogram-based measures of functional diversity. – Oikos 115: 179–185.

Podani, J. and Schmera, D. 2007. How should a dendrograms-based measure of functional diversity function? A rejoinder to Petchey and Gaston. – Oikos 116: 1427–1430.

Rao, A. R. and Srinivas, V. V. 2006. Regionalization of watersheds by hybrid-cluster analysis. – J. Hydrol. 318: 37–56.

Ricotta, C. 2005. A note on functional diversity measures. – Basic Appl. Ecol. 6: 479–486.

Sokal, R. R. et al. 1992. Character and OTU stability in five taxonomic groups. – J. Classif. 9: 117–140.

Sugihara, G. et al. 2003. Predicted correspondence between species abundances and dendrograms of niche similarities. – Proc. Natl Acad. Sci. USA 100: 5246–5251.

Thompson, K. et al. 2005. Biodiversity, ecosystem function and plant traits in mature and immature plant communities. – Funct. Ecol. 19: 355–358.

Wright, J. P. et al. 2006. Conventional functional classification schemes underestimate the relationship with ecosystem functioning. – Ecol. Lett. 9: 111–120.