Decomposing phylodiversity

Eric Marcon[1]* and Bruno Hérault[2]

[1]*AgroParisTech, UMR EcoFoG, BP 709, F-97310 Kourou, French Guiana; and* [2]*Cirad, UMR EcoFoG, BP 709, F-97310 Kourou, French Guiana*

Summary

1. Measuring functional or phylogenetic diversity is the object of an active literature. The main issues to address are relating measures to a clear conceptual framework, allowing unavoidable estimation-bias correction and decomposing diversity along spatial scales.

2. We provide a general mathematical framework to decompose measures of species-neutral, phylogenetic or functional diversity into α and β components. We first unify the definitions of phylogenetic and functional entropy and diversity as a generalization of HCDT entropy and Hill numbers when an ultrametric tree is considered. We then derive the decomposition of diversity. We propose a bias correction of the estimates allowing meaningful computation from real, often undersampled communities. Entropy can be transformed into true diversity, that is an effective number of species or communities.

3. Estimators of α- and β-entropy, phylogenetic and functional entropy are provided.

4. Proper definition and estimation of diversity is the first step towards better understanding its underlying ecological and evolutionary mechanisms.

Key-words: entropy, biodiversity, phylogenetic diversity, functional diversity

Introduction

The species-neutral approach of diversity measurement is based on Hill numbers, that is the effective number of species (Jost 2006). It is now being completed by far more interesting conceptual frameworks taking into account the species relatedness, that is either their functional or their phylogenetic proximity. This is what has been called, in the first case, 'functional diversity' (Tilman *et al.* 1997) and, in the second one, 'phylogenetic diversity' or 'phylodiversity' (Webb, Losos & Agrawal 2006). When both relative abundance and degree of relatedness between species (or individuals) are quantified, Pielou (1975) suggested that diversity measures should be generalized, integrating taxonomic differences between species. A little later, Rao (1982) proposed that the average of the species differences can be used as a measure of biodiversity. Despite some attempts to take into account taxonomic distinctness into a taxic diversity measure (Vane-Wright, Humphries & Williams 1991), this 'avant-garde' idea has been hardly applied in ecology (e.g. Warwick & Clarke 1995; Crozier 1997). During the last decade, increasing interests into the evolutionary history of communities (Webb 2000) as well as the need for conservation strategies taking phylogenetic risks into account (Faith 2008) revived the interest in phylodiversity partitioning.

Phylogenetic trees are built upon the genetic similarities among various biological individuals or other superior taxa. In a given local assemblage, phylogenetic diversity aims to quantify the evolutionary history shared among individuals since the time of the most recent common ancestor (Faith 1992;

Chao, Chiu & Jost 2010). All else being equal, an assemblage of phylogenetically divergent species is often seen as more diverse than a local assemblage of closely related species (Vellend *et al.* 2010). There is an increasing interest to partition this phylogenetic diversity not only between local communities but also between time periods in order to elucidate community assembly rules (Pavoine, Love & Bonsall 2009) and investigate what is commonly called the phylogenetic structure of communities (e.g. Cavender-Bares *et al.* 2004). For instance, Hardy & Senterre (2007) argued that a proper partitioning of phylodiversity is a necessary step prior to deciphering phylogenetic clustering (either due to local speciation of allopatric clades or habitat filtering of phylogenetically conserved traits) from phylogenetic overdispersion (allopatric speciation of two ancestral sympatric species, habitat filtering of phylogenetically convergent traits, competitive exclusion of related species).

Functional diversity was often defined as the extent of functional differences among individuals or species in a local community (Tilman 2001), an important determinant of ecosystem processes (Loreau *et al.* 2001). Functional diversity based on functional trees is a great tool to estimate the complementarity among individuals' or species' trait values by estimating their dispersion in trait space at all hierarchical scales simultaneously, avoiding discretization of continuous trait variation into functional groups (Petchey & Gaston 2002). Functional trees differ from phylogenetic trees as phylogenetic trees reflect evolutionary constraints whilst functional trees also take into account functional convergence (Hérault 2007). Each time a 'proper' functional tree can be constructed from a functional trait-based distance matrix (Podani & Schmera 2007), it should be possible to estimate and partition functional diversity in a manner similar to phylogenetic diversity (Petchey & Gaston

*Correspondence author. E-mail: Eric.Marcon@ecofog.gf

2002). However, functional differences among species or individuals and, *in fine*, the functional diversity value itself will depend strongly on the *a priori* choice of important functional traits (Weiher *et al.* 1999).

In this paper, we consider that all individuals or species of a given local community are placed in an ultrametric phylogenetic or functional tree. The distance between two species is measured as the length of the branches between them and their first common node. Our methods apply regardless from which biological information and how the tree is constructed, but phylogenetic diversity is the main target, as we will discuss it. We will write *phylodiversity* and *phyloentropy* for short when presenting the methods, and *phylogenetic* or *functional diversity* when we are more specific. The last two terms are also existing measures of diversity, PD (Faith 1992) and FD (Petchey & Gaston 2002). We will show that they are special cases of our measures (Table 1) and we will write PD and FD explicitly when considering them.

Chao, Chiu & Jost (2010) generalized Hill numbers to measure phylodiversity. Pavoine, Love & Bonsall (2009) generalized HCDT entropy to measure phyloentropy (Shimatani 2001; Ricotta 2005 had already done it, but for Rao's quadratic entropy only). We first show here their equivalence: phyloentropy is transformed into phylodiversity the same way HCDT entropy is transformed into diversity *sensu stricto*. Then, we derive phylodiversity partitioning as a straightforward generalization of that of HCDT diversity. We discuss the difference between our approach and that of Chiu, Jost & Chao (2014). Finally, we provide estimation-bias corrections for phyloentropy in order to obtain bias-corrected measures of phylodiversity.

Partitioning phylodiversity

TSALLIS ENTROPY

Tsallis entropy, also known as HCDT entropy (Havrda & Charvát 1967; Daróczy 1970; Tsallis 1988), has proven to be a powerful tool to measure diversity, generalizing the classical indices of diversity, including the number of species, Shannon and Simpson indices (Jost 2006). The order of diversity q gives more or less importance to rare species. Entropy can be converted into diversity *sensu stricto* (Hill 1973; Jost 2006), which is easy to interpret and compare. Statistical estimators of

diversity measures are intrinsically biased because of unseen species and also because they are not linear functions of probabilities (Marcon *et al.* 2014a). This is a serious issue (Dauby & Hardy 2012; Beck, Holloway & Schwanghart 2013), even if some bias corrections are available for HCDT entropy estimators (Grassberger 1988; Chao & Shen 2003; Marcon *et al.* 2014a).

SPECIES-NEUTRAL DIVERSITY

We first recall some features of HCDT diversity partitioning (Marcon *et al.* 2014a). Consider a metacommunity made of several local communities. Abundances of species in each local community are denoted $n_{s,i}$ ($s = 1, 2, \ldots, S$ is the index of species, i the index of communities). n_s is the number of individuals of species s in the metacommunity, n_i the number of individuals sampled in local community i and n the total number. The same notations are used for probabilities of occurrence $p_{s,i}$ whose population values are unknown but estimated with $\hat{p}_{s,i} = n_{s,i}/n_i$. Community weights are w_i: they may be equal to n_i/n, but any positive values summing to 1 are allowed. Probabilities in the metacommunity depend on these weights: $p_s = \sum_i w_i p_{s,i}$. Diversity of the metacommunity is γ-diversity. Diversity of local communities is α-diversity. The formalism of deformed logarithms is appropriated: it allows elegant and intuitive algebra. The logarithm of order q is defined as follows:

$$\ln_q x = \frac{x^{1-q} - 1}{1 - q} \qquad \text{eqn 1}$$

Its inverse function is the deformed exponential given as follows:

$$e_q^x = [1 + (1 - q)x]^{\frac{1}{1-q}} \qquad \text{eqn 2}$$

Note that

$$e_q^{x+y} = e_q^x e_q^{\frac{y}{1+(1-q)x}} \qquad \text{eqn 3}$$

Tsallis entropy of the metacommunity, $^qH_\gamma$, can be written as follows:

$$^qH_\gamma = \frac{1 - \sum_s p_s^q}{q - 1} = -\sum_s p_s^q \ln_q p_s \qquad \text{eqn 4}$$

Last, diversity is the deformed exponential of entropy, $^qD_\gamma = e_q^{^qH_\gamma}$, and entropy is the deformed logarithm of diversity: $^qH_\gamma = \ln_q {}^qD_\gamma$.

Table 1. Many usual measures of diversity are special cases of phyloentropy, either reducing it to species-neutral diversity or limiting it to values of q equal to 0, 1 or 2

	Diversity of order q	Special values of q
Phylogenetic or functional entropy/diversity	Entropy: $^q\bar{H}(T)$ Diversity: $^q\bar{D}(T)$	$T[^0\bar{H}(T) + 1]$ equals PD (Faith 1992) and FD (Petchey & Gaston 2002) $T[^1\bar{H}(T)]$ equals Hp, the phylogenetic generalization of Shannon's index (Allen, Kon & Bar-Yam 2009) $T[^2\bar{H}(T)]$ equals Rao's quadratic entropy
Species-neutral diversity	Entropy: ^qH Diversity: ^qD	$^0H + 1$ is species richness ^1H is Shannon entropy ^2H is Simpson entropy

PHYLOENTROPY AND PHYLODIVERSITY

Consider a phylogenetic or functional ultrametric tree (Fig. 1) partitioned into depth intervals delimited by slices passing through the internal nodes. Following Chao, Chiu & Jost (2010), the first slice starts at the bottom of the tree and ends at the lowest node. In slice k, L_k, leaves are found. The probabilities of occurrence of the species belonging to the branches that were below leaf l in the original tree are summed to give the grouped probability $u_{k,l}$.

We follow Pavoine, Love & Bonsall (2009) to define phyloentropy as the sum of the entropies in each tree slice, weighted by the slice height. However, we normalize it by the total tree height, $T = \sum_{k=1}^{K} T_k$. We denote it as $^q\bar{H}(T)$:

$$^q\bar{H}(T) = \sum_{k=1}^{K} \frac{T_k}{T}{}^q_k H \qquad \text{eqn 5}$$

$^q_k H$ is HCDT entropy in slice k. It is calculated as $^q_k H = -\sum_s u^q_{k,l}\ln_q u_{k,l}$.

Chao, Chiu & Jost (2010) generalized Hill numbers to phylogenetic diversity, defined as follows:

$$^q\bar{D}(T) = \left(\sum_{k=1}^{K} \frac{T_k}{T} \sum_{l=1}^{L_k} u^q_{k,l} \right)^{\frac{1}{1-q}} \qquad \text{eqn 6}$$

Simple algebra shows that

$$^q\bar{D}(T) = e_q^{{}^q\bar{H}(T)} \qquad \text{eqn 7}$$

This relation is exactly the same as the relation between HCDT entropy and diversity. In other words, phyloentropy is the weighted average of entropy along the tree, and phylodiversity is the corresponding Hill number. Entropy is linear, it can be summed over slices, but diversity is not: phylodiversity is not the weighted average of diversity along the tree.
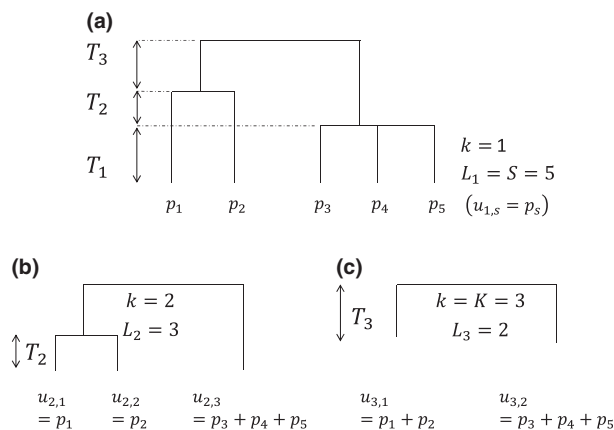

Fig. 1. Hypothetical ultrametric tree. (a) The whole tree contains three slices, delimited by two nodes. The length of slices is T_k. (b) Focus on slice 2. The tree without slice 1 is reduced to three leaves. Frequencies of collapsed species are $u_{k,l}$. (c) Slice 3 only.

DECOMPOSITION

Marcon *et al.* (2014a) derived the decomposition of HCDT entropy, generalizing Shannon entropy partitioning (Rao & Nayak 1985; Marcon *et al.* 2012), based on Patil and Taillie's concept of diversity of a mixture (Patil & Taillie 1982). Note that it differs from Jost's (2007) non-additive partitioning when community weights are unequal; see Marcon *et al.* (2014a) for a full discussion.

$$^qH_\gamma = {}^qH_\alpha + {}^qH_\beta = \sum_i w_i {}^q_i H_\alpha + \sum_i w_i {}^q_i H_\beta \qquad \text{eqn 8}$$

α- and β-entropies are the weighted sums of local community entropies $^q_i H_\alpha$ and $^q_i H_\beta$:

$$^q_i H_\alpha = -\sum_s p^q_{s,i}\ln_q p_{s,i}$$
$$^q_i H_\beta = \sum_s p^q_{s,i}\ln_q \frac{p_{s,i}}{p_s} \qquad \text{eqn 9}$$

Since phyloentropy is a linear transformation of generalized entropy, its decomposition is identical and follows equation (8). In slice k, HCDT γ-entropy is denoted $^q_k H_\gamma$, and the contributions of local community i to α- and β-entropy are $^q_{k,i} H_\alpha$ and $^q_{k,i} H_\beta$. This can be summed over slices and rearranged to obtain the decomposition of γ-phyloentropy:

$$\sum_k \frac{T_k}{T}{}^q H_\gamma = \sum_k \frac{T_k}{T} \sum_i w_{i\,k,i}\,{}^q H_\alpha + \sum_k \frac{T_k}{T} \sum_i w_{i\,k,i}\,{}^q H_\beta$$

$$\Leftrightarrow {}^q\bar{H}_\gamma(T) = \sum_i w_i \sum_k \frac{T_k}{T}{}^q_{k,i} H_\alpha + \sum_i w_i \sum_k \frac{T_k}{T}{}^q_{k,i} H_\beta$$

$$\Leftrightarrow {}^q\bar{H}_\gamma(T) = \sum_i w_i {}^q_i \bar{H}_\alpha(T) + \sum_i w_i {}^q_i \bar{H}_\beta(T) = {}^q\bar{H}_\alpha(T) + {}^q\bar{H}_\beta(T)$$

$$\text{eqn 10}$$

The deformed exponential of equation (8) is the decomposition of phylodiversity given as follows:

$$^q_\gamma\bar{D}(T) = {}^q_\alpha\bar{D}(T){}^q_\beta\bar{D}(T)$$

$$^q_\gamma\bar{D}(T) = e_q^{{}^q\bar{H}_\gamma(T)}; \; {}^q_\alpha\bar{D}(T) = e_q^{{}^q\bar{H}\alpha(T)}; \; {}^q_\beta\bar{D}(T) = e_q^{\frac{{}^q\bar{H}_\beta(T)}{1+(1-q)^q\bar{H}_\alpha(T)}}$$

$$\text{eqn 11}$$

α- and γ-phylodiversities can be interpreted as an equivalent number of species, that is to say the number of species equally different from each other (i.e. in an ultrametric tree made of a single slice), with the same probability of occurrence, that would give the same measure of diversity. β-phylodiversity is an equivalent number of communities, that is to say the number of completely distinct, equally weighted communities that would yield the same β-diversity as the actual metacommunity.

BIAS CORRECTION

α- and γ-HCDT entropies can be corrected following Marcon *et al.* (2014a). When q is low, unobserved species are the main issue that can be corrected according to Chao & Shen (2003). When q is high, the contribution of rare species to entropy is small, so the bias they cause is little, but entropy is less linear with respect to probabilities, requiring the correction of Grass-

berger (1988). The limit between low and high values of q is reached when both estimators are equal, empirically above $q = 1$ (Marcon *et al.* 2014a). Bias correction relies on the number of sampled individuals (probabilities are not enough) and can be computed for positive values of q. The unbiased estimators are denoted $^q\tilde{H}$ instead of $^q\hat{H}$. Their formulas are in Marcon *et al.* (2014a) and are not repeated here.

Phyloentropy can be corrected by summing the bias-corrected estimators of HCDT entropy in each slice of the tree. Bias-corrected α-entropy, $^q\tilde{H}_\alpha(T)$, relies on values of $^q_{k,i}\tilde{H}_\alpha$, the bias-corrected estimators of HCDT α-entropy in slice k in local community i.

$$^q\tilde{H}_\alpha(T) = \sum_i w_i \sum_k \frac{T_k}{T}{}^q_{k,i}\tilde{H}_\alpha \qquad \text{eqn 12}$$

Since the number of individuals in some leaves $u_{k,l}$ increases in slices close to the root of the tree, the bias decreases with k.

$^q\tilde{H}_\gamma(T) = \sum_k \frac{T_k}{T}{}^q_{k,i}\tilde{H}_\gamma$ is calculated in the same way. β-phyloentropy is obtained as the difference between $^q\tilde{H}_\gamma(T)$ and $^q\tilde{H}_\alpha(T)$ because Grassberger's correction is not available to allow direct calculation.

EXAMPLE

We used the tropical forest data set already investigated by Marcon *et al.* (2012, 2014a). Two 1-ha plots were fully inventoried in the Paracou field station in French Guiana. 1124 individual trees (diameter at breast height over 10 cm) have been sampled among 229 species. The phylogenetic tree was built introducing a rough taxonomy of the 229 species in the analysis: distance between species of the same genus is set to 1, 2 for different genera of the same family and 3 for different families. The functional tree was based on species relatedness using four key functional traits, each of them related to one axis of the leaf–height–seed–stem economic spectra of tropical trees (Baraloto *et al.* 2010b): seed mass and tree maximum height (Hérault *et al.* 2011) plus specific leaf area and wood specific gravity (Baraloto *et al.* 2010a). The functional tree was built from a Gower's similarity matrix agglomerated using Ward's method (full details in Hérault & Honnay 2007). Diversity was calculated with the *entropart* package (Marcon & Hérault 2014) under ʀ (R Development Core Team 2014): bias-corrected entropy was calculated first, summed and finally transformed into diversity. Necessary R codes are in the supporting information, Appendix S1.

We first calculated the species-neutral, phylogenetic and functional diversity of order 1 of the metacommunity (the two plots) and partitioned it (each plot is considered as a local community, weights are proportional to the numbers of individuals). The γ-species-neutral diversity (Hill number of Shannon entropy) is 134 effective species, partitioned into α-diversity equal to 92 effective species (82 and 107 in each plot) and β-diversity equal to 1·46 equivalent communities. Phylogenetic and functional diversity values, respectively, are: $^1_\gamma\bar{D}(T) = 55$ and 5·9, $^1_\alpha\bar{D}(T) = 42$ and 5·5 with $^1_\beta\bar{D}(T) = 1·29$ and 1·06. Considering the taxonomy of Paracou species, γ-phylodiversity is around 2·5 times smaller than species-neutral diversity.

Functional diversity is only six equivalent species, showing an extreme redundancy according to the functional tree: FD (Petchey & Gaston 2002), that is functional diversity of order 0, is estimated equal to 18 whilst the number of estimated species is 297.

Since γ-diversity is the product of α by β, they can be represented as nested rectangles (Fig. 2). The rectangle of size $^q_\alpha\bar{D}(T)$ by $^q_\beta\bar{D}(T)$ has the same area as that of size $^q_\gamma\bar{D}(T)$ by 1. Plotting species-neutral and phylodiversity together summarizes the essential information: the reduction of diversity due to the consideration of species phylogenetic or functional proximity.

Profiles (Fig. 3) can be drawn for species-neutral, phylogenetic and functional diversities.
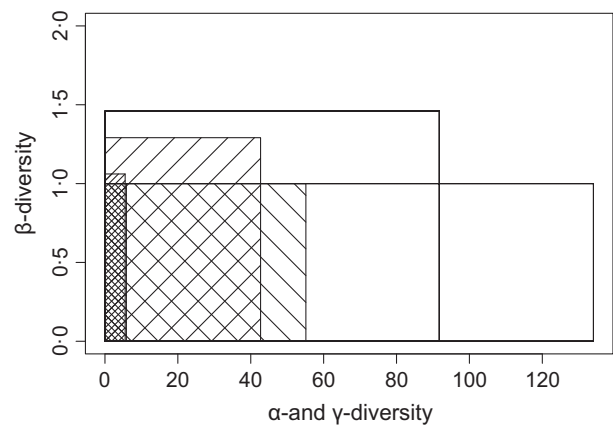

Fig. 2. Graphical representation of the diversity of order 1 in Paracou plots. Transparent rectangles represent species-neutral diversity, hatched rectangles phylogenetic diversity and shaded rectangles functional diversity. In each case, the horizontal rectangle of height 1 represents γ-diversity (respectively, 134, 55 and 6 effective species). The other rectangle has the same area, but its size is α-diversity by β-diversity.
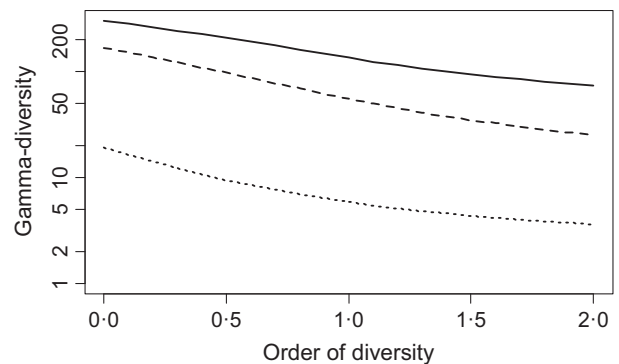

Fig. 3. γ-diversity profile of Paracou plots. Species-neutral diversity (solid line), phylogenetic diversity (dashed line) and functional diversity (dotted line) are plotted against the order of diversity, between 0 (number of species, PD and FD) and 2 (Simpson diversity and Rao's quadratic entropy transformed into diversity), with estimation-bias correction. Diversity scale is logarithmic for readability.

Discussion

UNIFICATION OF MEASURES OF DIVERSITY

Phyloentropy generalizes many previous indices of diversity. Rao's (1982) quadratic entropy is phyloentropy of order 2 multiplied by T, the tree height. It has been explored in depth and several results obtained here were already known in this special case. It has been partitioned early by Rao himself, weighting communities according to their number of individuals, as Villeger & Mouillot (2008) whilst Hardy & Senterre (2007) or Pavoine *et al.* (2013) used equal weights. Hardy & Jost (2008) validated both weightings but a general framework allowing the additive partitioning of Rao's entropy was missing (Guiasu & Guiasu 2011). We showed that arbitrary weights are acceptable.

Other indices of diversity can be considered as special cases of phyloentropy (Table 1).

ALTERNATIVE PARTITIONING

Chiu, Jost & Chao (2014) propose a different partitioning of phylodiversity (Chao, Chiu & Jost 2010) focusing on the independence between α and β components, following Jost (2007). It requires a particular definition of α-diversity (in Chiu, Jost & Chao 2014; equation 6 for neutral diversity and (8) for phylodiversity), whilst we adopt Routledge's (1979) definition: α-entropy is the weighted average entropy of communities, see equation (8). Chiu *et al.*'s approach is completely different from ours, as we will show it with a simple example. Consider N communities containing a single species, no species is shared between communities. Whatever q, the entropy of each community is 0, its diversity is 1 effective species. In our framework, α-entropy equals 0 and α-diversity is 1. More generally, whatever the weights and whatever q, if all communities have the same diversity, α-diversity equals it.

In Chiu *et al.*'s framework, β-diversity must be N since no species are shared, so α-diversity is γ-diversity divided by N. Species-neutral α-diversity of our example is not 1 but $\frac{1}{N}\left(\sum_{i=1}^{N} w_i^q\right)^{\frac{1}{1-q}}$. Community weights and species frequencies play a similar role: low-weighted communities, as rare species, have a lower influence when q increases, and inversely, α-diversity is driven by rare species of low-weighted communities when q decreases. We consider in this paper that community weights are arbitrary, such as sampling unit sizes, so Chiu *et al.*'s α-diversity is not suitable here.

We believe that Routledge's definition of α-diversity is more appropriate. Entropy is the average information in each community so it can meaningfully be averaged between communities according to their weight to define α-entropy. Adding an infinitesimal community (with weight close to 0) does not change the metacommunity's diversity, whilst it changes discontinuously in Chiu *et al.*'s framework (β-diversity jumps from N to $N + 1$, for example).

The price to pay is α- and β-diversities are not independent, as discussed more thoroughly in Marcon *et al.* (2014a). The real consequences of this dependence will have to be studied in depth.

NON-ULTRAMETRIC DISTANCES BETWEEN SPECIES

Our framework relies on ultrametric trees, since entropy must be calculated slice by slice. Phylogenetic data are usually organized as a tree, but not necessarily ultrametric. Chao, Chiu & Jost (2010) calculate $q\bar{D}(T)$ as a sum over the branches rather than other slices of the trees, allowing them to address non-ultrametric trees. Although it is defined mathematically, such a value of phylodiversity faces several issues. Pavoine & Bonsall (2009) discuss its inconsistency in the special case of $q = 2$, for example, the fact that the species distribution maximizing diversity is not unique then. Leinster & Cobbold (2012) show that the distance between species used to calculate $q\bar{D}(T)$ depends on species frequencies, questioning the very sense of what is measured. For these two reasons, we conclude that non-ultrametric trees are not appropriate to measure phylodiversity in our framework, not only for technical issues (only ultrametric trees can be sliced to allow estimation-bias correction) but for conceptual ones.

Functional diversity is more frequently calculated as a non-ultrametric matrix of distances between species, whose transformation into a dendrogram causes deformations (Pavoine, Ollier & Dufour 2005). The choice of the clustering method influences the shape of the tree and may lead to inconsistent results (Podani & Schmera 2006), although appropriate methods, applied to the example above, reduce these issues (Podani & Schmera 2007). A more appropriate way to address functional diversity is probably using directly the distance matrix between species or its transformation into a similarity matrix. Similarity-based diversity (Leinster & Cobbold 2012) may be preferred to evaluate functional diversity. We derive its decomposition and propose reduced-bias estimators elsewhere (Marcon, Zhang & Hérault 2014b).

Conclusion

In this paper, we provide a general, consistent and operational framework to decompose measures of species-neutral, phylogenetic or even functional diversity into α (within local communities) and β (between local communities) components. We show that entropy can be calculated and its estimation bias corrected in each slice of the phylogenetic or functional tree, summed over slices and finally transformed into diversity. In fact, phylodiversity can be analysed without using any species concept (i.e. diversity of individuals without categorizing them into a set of species) provided that phylogenetic or functional distance between individuals can be assessed, for example using molecular data or functional trait measured for each individual member of a metacommunity (Paine *et al.* 2011). Being able to properly partition phylodiversity is a necessary step towards deciphering the ecological and evolutionary mechanisms that underlie the structure and assembly of communities. Moreover, diversity partitioning will improve our assessment of human-driven modifications of ecosystem functioning in conservation studies.

Acknowledgements

Data accessibility

The R scripts used to work the examples are available in the online supplement of the paper. They rely on the entropart package (Marcon & Hérault 2014) for R, which contains the data.

References

Allen, B., Kon, M. & Bar-Yam, Y. (2009) A new phylogenetic diversity measure generalizing the Shannon index and its application to phyllostomid bats. *American Naturalist*, **174**, 236–243.

Baraloto, C., Paine, C.E.T.P., Patiño, S., Bonal, D., Hérault, B. & Chave, J. (2010a) Functional trait variation and sampling strategies in species rich plant communities. *Functional Ecology*, **24**, 208–216.

Baraloto, C., Paine, C.E.T., Poorter, L., Beauchêne, J., Bonal, D., Domenach, A.M. *et al.* (2010b) Decoupled leaf and stem economics in rain forest trees. *Ecology Letters*, **13**, 1338–1347.

Beck, J., Holloway, J.D. & Schwanghart, W. (2013) Undersampling and the measurement of beta diversity. *Methods in Ecology and Evolution*, **4**, 370–382.

Cavender-Bares, J., Ackerly, D.D., Baum, D.A. & Bazzaz, F.A. (2004) Phylogenetic overdispersion in Floridian oak communities. *The American Naturalist*, **163**, 823–843.

Chao, A., Chiu, C.-H. & Jost, L. (2010) Phylogenetic diversity measures based on Hill numbers. *Philosophical Transactions of the Royal Society B*, **365**, 3599–3609.

Chao, A. & Shen, T.J. (2003) Nonparametric estimation of Shannon's index of diversity when there are unseen species in sample. *Environmental and Ecological Statistics*, **10**, 429–443.

Chiu, C.H., Jost, L. & Chao, A. (2014) Phylogenetic beta diversity, similarity, and differentiation measures based on Hill numbers. *Ecological Monographs*, **84**, 21–44.

Crozier, R.H. (1997) Preserving the information content of species: genetic diversity, phylogeny, and conservation worth interaction with reasons justifying the preservation. *Annual Review of Ecology and Systematics*, **28**, 243–268.

Daróczy, Z. (1970) Generalized information functions. *Information and Control*, **16**, 36–51.

Dauby, G. & Hardy, O.J. (2012) Sampled-based estimation of diversity sensu stricto by transforming Hurlbert diversities into effective number of species. *Ecography*, **35**, 661–672.

Faith, D.P. (1992) Conservation evaluation and phylogenetic diversity. *Biological Conservation*, **61**, 1–10.

Faith, D.P. (2008) Threatened species and the potential loss of phylogenetic diversity: conservation scenarios based on estimated extinction probabilities and phylogenetic risk analysis. *Conservation Biology : The Journal of the Society for Conservation Biology*, **22**, 1461–1470.

Grassberger, P. (1988) Finite sample corrections to entropy and dimension estimates. *Physics Letters A*, **128**, 369–373.

Guiasu, R.C. & Guiasu, S. (2011) The weighted quadratic index of biodiversity for pairs of species: a generalization of Rao's index. *Natural Science*, **3**, 795–801.

Hardy, O.J. & Jost, L. (2008) Interpreting and estimating measures of community phylogenetic structuring. *Journal of Ecology*, **96**, 849–852.

Hardy, O.J. & Senterre, B. (2007) Characterizing the phylogenetic structure of communities by an additive partitioning of phylogenetic diversity. *Journal of Ecology*, **95**, 493–506.

Havrda, J. & Charvát, F. (1967) Quantification method of classification processes. Concept of structural a-entropy. *Kybernetika*, **3**, 30–35.

Hérault, B. (2007) Reconciling niche and neutrality through the Emergent Group approach. *Perspectives in Plant Ecology, Evolution and Systematics*, **9**, 71–78.

Hérault, B. & Honnay, O. (2007) Using life-history traits to achieve a functional classification of habitats. *Applied Vegetation Science*, **10**, 73–80.

Hérault, B., Bachelot, B., Poorter, L., Rossi, V., Bongers, F., Chave, J., Paine, C.E.T., Wagner, F. & Baraloto, C. (2011) Functional traits shape ontogenetic growth trajectories of rain forest tree species. *Journal of Ecology*, **99**, 1431–1440.

Hill, M.O. (1973) Diversity and evenness: a unifying notation and its consequences. *Ecology*, **54**, 427–432.

Jost, L. (2006) Entropy and diversity. *Oikos*, **113**, 363–375.

Jost, L. (2007) Partitioning diversity into independent alpha and beta components. *Ecology*, **88**, 2427–2439.

Leinster, T. & Cobbold, C. (2012) Measuring diversity: the importance of species similarity. *Ecology*, **93**, 477–489.

Loreau, M., Naeem, S., Inchausti, P., Bengtsson, J., Grime, J.P., Hector, A. *et al.* (2001) Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science*, **294**, 804–808.

Marcon, E. & Hérault, B. (2014). entropart, an R package to partition diversity. *Journal of Statistical Software*, in press.

Marcon, E., Zhang, Z. & Hérault, B. (2014b) The decomposition of similarity-based diversity and its bias correction. *HAL*, **00989454**, 1–12.

Marcon, E., Hérault, B., Baraloto, C. & Lang, G. (2012) The decomposition of Shannon's entropy and a confidence interval for beta diversity. *Oikos*, **121**, 516–522.

Marcon, E., Scotti, I., Hérault, B., Rossi, V. & Lang, G. (2014a) Generalization of the partitioning of Shannon diversity. *PLoS One*, **9**, e90289.

Paine, C.E.T., Baraloto, C., Chave, J. & Hérault, B. (2011) Functional traits of individual trees reveal ecological constraints on community assembly in tropical rain forests. *Oikos*, **120**, 720–727.

Patil, G.P. & Taillie, C. (1982) Diversity as a concept and its measurement. *Journal of the American Statistical Association*, **77**, 548–561.

Pavoine, S. & Bonsall, M.B. (2009) Biological diversity: distinct distributions can lead to the maximization of Rao's quadratic entropy. *Theoretical Population Biology*, **75**, 153–163.

Pavoine, S., Love, M.S. & 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.

Pavoine, S., Ollier, S. & Dufour, A.-B. (2005) Is the originality of a species measurable? *Ecology Letters*, **8**, 579–586.

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.

Petchey, O.L. & Gaston, K.J. (2002) Functional diversity (FD), species richness and community composition. *Ecology Letters*, **5**, 402–411.

Pielou, E.C. (1975) *Ecological Diversity*. Wiley, New York.

Podani, J. & Schmera, D. (2006) On dendrogram-based measures of functional diversity. *Oikos*, **115**, 179–185.

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

R Development Core Team. (2014). *R: A Language and Environment for Statistical Computing*. R Development Core Team, R Foundation for Statistical Computing, Vienna, Austria.

Rao, C.R. (1982) Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology*, **21**, 24–43.

Rao, C. & Nayak, T. (1985) Cross entropy, dissimilarity measures, and characterizations of quadratic entropy. *Information Theory, IEEE Transactions on*, **31**, 589–593.

Ricotta, C. (2005) On hierarchical diversity decomposition. *Journal of Vegetation Science*, **16**, 223–226.

Routledge, R.D. (1979) Diversity indices: which ones are admissible? *Journal of Theoretical Biology*, **76**, 503–515.

Shimatani, K. (2001) Multivariate point processes and spatial variation of species diversity. *Forest Ecology and Management*, **142**, 215–229.

Tilman, D. (2001). Functional diversity. *Encyclopedia of Biodiversity* (ed. S. Levin), pp. 109–121. Academic Press, San Diego.

Tilman, D., Knops, J., Wedin, D., Reich, P., Ritchie, M. & Siemann, E. (1997) The influence of functional diversity and composition on ecosystem processes. *Science*, **277**, 1300–1302.

Tsallis, C. (1988) Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*, **52**, 479–487.

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.

Vellend, M., Cornwell, W.K., Magnuson-Ford, K. & Mooers, A.Ø. (2010). Measuring phylogenetic biodiversity. *Biological Diversity: Frontiers in Measurement and Assessment* (eds A.E. Magurran & B.J. McGill), pp. 194–207. Oxford University Press, Oxford.

Villeger, S. & Mouillot, D. (2008) Additive partitioning of diversity including species differences: a comment on Hardy & Senterre (2007). *Journal of Ecology*, **96**, 845–848.

Warwick, R.M. & Clarke, K.R. (1995) New "biodiversity" measures reveal a decrease in taxonomic distinctness with increasing stress. *Marine Ecology Progress Series*, **129**, 301–305.

Webb, C.O. (2000) Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *American Naturalist*, **156**, 145–155.

Webb, C.O., Losos, J.B. & Agrawal, A.A. (2006) Integrating phylogenies into community ecology. *Ecology*, **87**, S1–S2.

Weiher, E., van der Werf, A., Hompson, K., Roderick, M., Garnier, E. & Eriksson, O. (1999) Challenging Theophrastus: a common core list of plant traits for functional ecology. *Journal of Vegetation Science*, **10**, 609–620.

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Appendix S1. R code of the example.