

FORUM

Additive partitioning of diversity including species differences: a comment on Hardy & Senterre (2007)

Sébastien Villéger* and David Mouillot

UMR CNRS-IFREMER-UM2 5119 Ecosystèmes Lagunaires, Université Montpellier 2, CC 093,
34 095 Montpellier Cedex 5, France

Summary

1. The distribution of biodiversity at multiple spatial scales has been traditionally investigated through the additive partitioning of γ -biodiversity (regional) into α -(within-site) and β -(among-site) components.
2. However, this decomposition is almost exclusively applied using species turnover among communities while two communities with no species in common can be very similar because they share some 'biological' similarity.
3. To overcome this limitation, Hardy & Senterre (2007) (*J. Ecol.*, **95**, 493–506) presented a new statistical framework partitioning the phylogenetic diversity into α - and β -components using the Rao's quadratic entropy.
4. We show that their decomposition is correct only when sites have the same total abundance, otherwise it may lead to negative β -diversity values. As an alternative, we provided a general decomposition of the quadratic entropy into α -, β - and γ -diversities.
5. *Synthesis.* We suggest that the 'biological' turnover quantified by the β -component of the regional quadratic entropy may help, at least, to disentangle dispersal vs. niche influences on biodiversity patterns.

Key-words: biodiversity pattern, biogeography, β -diversity, community structure, functional diversity, phylogenetic diversity, Rao's quadratic entropy

Introduction

Partitioning biological diversity within and among sites has been the focus of ecological research for several years because the pioneering works of Whittaker (1960, 1972). Indeed, the amount of species turnover among assemblages is a key component for the conservation of biodiversity and for understanding the processes underlying species distribution at multiple spatial scales. This diversity turnover was termed β -diversity while local and regional diversities were called α - and γ -diversities, respectively. In contrast to α - and γ -diversities, there is still no consensus about the measure of β -diversity because the perception of this component differs extensively among authors (Whittaker *et al.* 2001; Legendre *et al.* 2005).

Whittaker (1960, 1972) expressed β -diversity using a multiplicative model where, $\gamma = \beta \times \bar{\alpha}$, $\bar{\alpha}$ being the average α -diversity across sites. As the breakthrough paper of Lande (1996), the additive partition of diversity has rapidly emerged as a consensual framework to quantify the three components of diversity following the formula: $\gamma = \beta + \bar{\alpha}$. Hence, through additive

partitioning, β -diversity has the same unit as α - and γ -diversities and is nothing more than a partial amount of the regional diversity and, as such, deserves to be properly quantified.

In most studies dealing with additive partitioning of diversity, the term β -diversity is almost exclusively applied to species turnover among assemblages. In other words, species are assumed to have the same level of 'biological' dissimilarity between each other, that is all the 'biological' distances between species pairs are the same. However, two communities can be very dissimilar in terms of species composition, with few species in common, but very similar from a phylogenetic or functional point of view. This would be the case for two assemblages with no species in common but containing the same functional groups. It is thus surprising that current measures of β -diversity ignore what makes species different in an assemblage: their relative abundances and their biological traits. This kind of species-based measure is a narrow view of β -diversity and we trust that we can obtain a completely different degree of β -diversity among assemblages if, instead of relying solely on species composition, we also take into account relationships among species, be they phylogenetic or functional.

*Correspondence author. E-mail: villegier@supagro.inra.fr

Analogous to the intra vs. interspecific genetic diversity distinction, Hardy & Senterre (2007) presented a new statistical framework partitioning phylogenetic diversity into α -(within-site) and a β -(among-site) components. Their measures of the diversity components are based on the Rao's quadratic entropy (Rao 1982) in the sense that they include species relative abundances and distances among species (phylogenetic distances). Here we argue that their additive partitioning of the Rao's quadratic entropy is correct only when the different sampled communities hold exactly the same total abundance, which is rarely the case in ecology. Otherwise, we show that their proposal may lead to negative β -diversity values which are not interpretable. Keeping the relevant framework of Hardy & Senterre (2007) and using the seminal contribution of Ricotta (2005), we formalize a flexible additive partitioning of diversity taking into account species differences.

The framework of Hardy & Senterre (2007)

In their paper, Hardy & Senterre (2007) proposed the following framework for diversity decomposition: let A_{ik} being the abundance (number of individuals or biomass) of species i within site k ; its relative abundance within site k , f_{ik} , is thus $f_{ik} = A_{ik} / \sum_i A_{ik}$, and its relative abundance over all sites, f_i , is $f_i = \sum_k A_{ik} / \sum_k \sum_i A_{ik}$.

The distance between species i and j was noted δ_{ij} .

The total or regional diversity (γ), according to the quadratic entropy (Rao 1982), was defined as

$$D_T = \sum_i \sum_j \delta_{ij} f_i f_j \quad \text{eqn 1}$$

The diversity within site k (α) was

$$D_k = \sum_i \sum_j \delta_{ij} f_{ik} f_{jk} \quad \text{eqn 2}$$

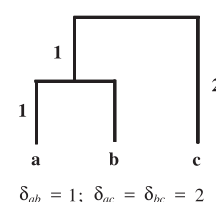
The average within-site ($\bar{\alpha}$) diversity, D_S , was presented as the expectation of D_k over all sites. The β -diversity is then the difference $D_T - D_S$ according to the additive partitioning of diversity: $\beta = \gamma - \bar{\alpha}$.

In Fig. 1 we present a case study where these formula lead to $D_S > D_T$, so to a negative β -diversity. Actually if the formulae proposed by Hardy & Senterre (2007) are correct to compute D_k and D_T based on the quadratic entropy, it is not the case for D_S . Indeed if we consider N sites, Hardy & Senterre (2007) calculate D_S as $D_S = (1/N) \sum_{k=1}^N D_k$, therefore, giving the same weight ($1/N$) to all sites without taking into account the total abundance at each site.

Consider a case with three sites, all having the same three species but where total abundances are unbalanced between sites (Fig. 1). In this example we consider that the samples from the three sites are not biased, that is, they reflect the community structure and composition of each site. The three species are placed on a theoretical phylogenetic tree and an ultrametric distance can be derived between all species pairs (Fig. 1, top right). The sites with low total abundances (sites 1 and 3) have a high α -diversity value ($D_1 = \alpha_1 = D_3 = \alpha_3 = 1.125$) whereas site 2, with the highest total abundance, has a rela-

Number of individuals (A_{ik})

	Sp a	Sp b	Sp c	Total
Site 1	1	1	2	4
Site 2	28	1	1	30
Site 3	1	1	2	4
Total	30	3	5	40



$$\delta_{ab} = 1; \delta_{ac} = \delta_{bc} = 2$$

Site relative abundances (f_{ik})

	Sp a	Sp b	Sp c	Total
Site 1	0.25	0.25	0.5	1
Site 2	0.934	0.033	0.033	1
Site 3	0.25	0.25	0.5	1

$$\begin{aligned} \alpha_1 &= 1.125 \\ \alpha_2 &= 0.191 \\ \alpha_3 &= 1.125 \end{aligned} \quad \begin{aligned} D_S &= 0.814 \\ D'_S &= 0.388 \end{aligned}$$

Regional relative abundances (p_{ik})

	Sp a	Sp b	Sp c	f_i
Site 1	0.026	0.026	0.053	0.105
Site 2	0.737	0.026	0.026	0.789
Site 3	0.026	0.026	0.053	0.105
f_i	0.789	0.079	0.132	1

$$\begin{aligned} D_T &= 0.582 \\ D_T - D_S &= -0.232 \\ D_T - D'_S &= 0.194 \end{aligned}$$

Fig. 1. Hypothetical abundances (number of individuals) for a three species regional pool and three local communities. The distances (phylogenetical or functional) between species are ultrametric and are represented on the dendrogram on the right of the table containing abundances. From local abundances (A_{ik}) we computed species relative abundances by site (f_{ik}) and then the three local diversities (α) based on the Rao's quadratic entropy index. Similarly, using regional relative abundances (p_{ik}), we computed regional diversity (D_T). When the mean local diversity is not weighted (D_S), as suggested by Hardy & Senterre (2007), it leads to a negative value of β -diversity ($D_T - D_S$). By contrast, if we use the weighted mean local diversity (D'_S) by the contribution of each site (f_i), then $\beta = (D_T - D'_S)$ is positive, as intuitively expected under this scenario.

tively low α -diversity ($D_2 = \alpha_2 = 0.191$). Then the average α -diversity (D_S) is high (0.814) because Hardy & Senterre (2007) give the same weight to all sites regardless of their total abundance. In parallel, γ -diversity (D_T) computed using species relative abundances over all sites (f_i) is mainly determined by the most abundant species and hence its value is lower than D_S ($D_T = 0.582 < D_S$). Finally, we obtain a negative β -diversity value $D_T - D_S = -0.232$ – an unwanted result.

Fortunately, in their application, Hardy & Senterre (2007) analysed a data set where 100 individuals were sampled in each site to estimate relative abundances. Therefore, the bias we demonstrate did not affect their conclusions. However, their estimator of D_S , $\hat{D}_S = (1/N) \sum_{k \leq N} \hat{D}_k$ is neither correct for a general use, nor their estimator of D_T , $\hat{D}_T = (1/(N(N-1))) \sum_k \sum_{l \neq k} \sum_i \sum_j \delta_{ij} f_{ik} f_{jl}$.

Indeed, these two last estimators are correct only when (i) total abundances are equivalent among sites and samples are not biased; (ii) samples from the different sites contain exactly the same total abundance (the sampling effort would force the samples to contain the same total abundance); or (iii) total abundances across communities are purely artificial results of sampling effort and can be normalized to a predefined number of individuals.

These three cases are not the norm in ecology and if total abundances reflect real species abundances in different sites we have to use the formula provided by Ricotta (2005) and then the proposed corrections would be appropriate to estimate the correct average α -quadratic entropy and accordingly, the correct β -quadratic entropy. Indeed, Ricotta (2005) gives a weight to each site that is proportional to its relative abundance over sites ($f_{.k}$), to estimate the average α -diversity and his formula is equivalent to: $D'_S = \sum_{k=1}^N f_{.k} D_k$.

$$\text{With: } f_{.k} = \frac{\sum_i A_{ik}}{\sum_k \sum_i A_{ik}}$$

Using this last formula we obtained with our example an average within-site (α) diversity of $D'_S = 0.388$ and then $D_\beta = 0.194$ instead of -0.232 (Fig. 1, bottom). It means that the phylogenetic structure is dissimilar among the three assemblages, which is the case.

Additive partitioning of the Rao's quadratic entropy

Ricotta (2005) paved the way to the correct estimation of the α -, β - and γ -diversity components of Rao's quadratic entropy but Hardy & Senterre (2007) propose a complete framework with parameter estimations. Here we show that the Rao's γ -diversity can be decomposed into an intra-site and inter-site components but that those two components do not correspond to α - and β -diversities.

Our new decomposition is based on p_{ik} which is the relative abundance of species i in site k over all sites (hereafter called regional relative abundance) with:

$$p_{ik} = \frac{A_{ik}}{\sum_k \sum_i A_{ik}} = f_{ik} \times f_{.k}$$

$$f_{.i} = \sum_k p_{ik}$$

The γ -quadratic entropy (D_T) is nothing more than the sum of the distances between all pairs of species i and j within and between all pairs of sites k and l weighed by their regional relative abundances:

$$D_T = \sum_k \sum_l \left(\sum_i \sum_j \delta_{ij} p_{ik} p_{jl} \right) = \sum_i \sum_j \delta_{ij} \sum_k p_{ik} \sum_l p_{jl} \quad \text{eqn 3}$$

$$= \sum_i \sum_j \delta_{ij} f_{.i} f_{.j}$$

This is equivalent to the Formula (1) following Ricotta (2005) and Hardy & Senterre (2007).

Hence, the γ -quadratic entropy (D_T) can be divided into (i) the sum of the distances between all pairs of species i and j within sites weighed by their regional relative abundances; and (ii) the sum of the distances between all pairs of species i and j between all pairs of sites k and l ($l \neq k$) weighed by their regional relative abundances. This is equivalent to an intra-

vs. inter-site decomposition of the γ -quadratic entropy. More formally we obtain the additive partitioning:

$$D_T = \sum_k \sum_l \left(\sum_i \sum_j \delta_{ij} p_{ik} p_{jl} \right) \quad \text{eqn 4}$$

$$= \sum_k \sum_i \sum_j \delta_{ij} p_{ik} p_{jk} + \sum_k \sum_{l \neq k} \left(\sum_i \sum_j \delta_{ij} p_{ik} p_{jl} \right)$$

Which is equivalent to: $D_T = D_{\text{Intra}} + D_{\text{Inter}}$

In addition we have:

$$D_{\text{Intra}} = \sum_k \sum_i \sum_j \delta_{ij} p_{ik} p_{jk} = \sum_k \sum_i \sum_j \delta_{ij} f_{ik} f_{.k} f_{jk} f_{.k} \quad \text{eqn 5}$$

$$= \sum_k f_{.k}^2 \sum_i \sum_j \delta_{ij} f_{ik} f_{jk} = \sum_k f_{.k}^2 D_k$$

Then, using eqn 4, β -diversity can be expressed as:

$$D_\beta = D_T - D_S = \sum_k \sum_i \sum_j \delta_{ij} p_{ik} p_{jk} + \sum_k \sum_{l \neq k} \left(\sum_i \sum_j \delta_{ij} p_{ik} p_{jl} \right) - \sum_k f_{.k} D_k$$

And using eqn 5 we obtained:

$$D_\beta = \sum_{k=1} f_{.k}^2 D_k + \sum_k \sum_{l \neq k} \left(\sum_i \sum_j \delta_{ij} p_{ik} p_{jl} \right) - \sum_{k=1} f_{.k} D_k$$

Finally, we can express the β -diversity as:

$$D_\beta = \sum_k (f_{.k}^2 - f_{.k}) D_k + \sum_k \sum_{l \neq k} \left(\sum_i \sum_j \delta_{ij} p_{ik} p_{jl} \right) \quad \text{eqn 6}$$

So, as $0 \leq f_{.k} \leq 1$ and $f_{.k}^2 - f_{.k} \leq 0$, D_β is less than the inter-site component of the quadratic entropy.

D_β is also:

$$D_\beta = \sum_k f_{.k} (f_{.k} - 1) D_k + \sum_k \sum_{l \neq k} \left(\sum_i \sum_j \delta_{ij} p_{ik} p_{jl} \right)$$

With $\sum_{l \neq k} f_{.l} = 1 - f_{.k}$ we obtained:

$$D_\beta = - \sum_k f_{.k} \sum_{l \neq k} (f_{.l}) D_k + \sum_k \sum_{l \neq k} \left(\sum_i \sum_j \delta_{ij} p_{ik} p_{jl} \right) \quad \text{eqn 7}$$

$$D_\beta = \sum_k \sum_{l \neq k} \left(\sum_i \sum_j \delta_{ij} p_{ik} p_{jl} \right) - \sum_k \sum_{l \neq k} f_{.k} f_{.l} D_k$$

$$D_\beta = \sum_k \sum_{l \neq k} f_{.k} f_{.l} \left(\sum_i \sum_j \delta_{ij} f_{ik} f_{jl} \right) - \sum_k \sum_{l \neq k} f_{.k} f_{.l} D_k$$

$$D_\beta = \sum_k \sum_{l \neq k} f_{.k} f_{.l} \left[\sum_i \sum_j \delta_{ij} (f_{ik} f_{jl} - f_{ik} f_{jk}) \right]$$

$$D_\beta = \sum_k \sum_{l \neq k} f_{.k} f_{.l} \left[\sum_i \sum_j \delta_{ij} f_{ik} (f_{jl} - f_{jk}) \right]$$

Using this new expression of β -diversity, it appears that D_β equals zero when all the $f_{jl} - f_{jk} = 0$, so $f_{jl} = f_{jk}$, for each species j and for each couple of sites (k, l). In other words, species must have the same relative abundances in each site to obtain $D_\beta = 0$.

Conclusion

We show that, despite the importance of the framework proposed by Hardy & Senterre (2007), their decomposition of the quadratic entropy into α -, β - and γ -diversities may lead to negative β -diversity values because the average α -diversity may exceed γ -diversity. Here we propose a correct decomposition of the quadratic entropy into α -, β - and γ -diversities based on the initial proposal of Ricotta (2005). In addition, we show that α - and β -diversity components do not correspond to the partition of γ -diversity into the intra-site and the inter-site components. Indeed, β -diversity is not simply the sum of the distances between all pairs of species i and j between all pairs of sites k and l ($k \neq l$) which is the inter-site component of the quadratic entropy. In a similar vein, α -diversity is not the sum of the distances between all pairs of species i and j in each site which is the intra-site component of the quadratic entropy. Instead, the α -component of the quadratic entropy is the average within-site diversity while the β -component is a 'biological' turnover across sites, which expresses the differentiation among sites using pairwise differences between species, be they phylogenetic or functional.

We suggest that β -diversity based on species 'biological' dissimilarities and species relative abundances will not react as classical β -indices based solely on species composition. Moreover, as demonstrated by Champely & Chessel (2002) and underlined by Ricotta (2005), the additive decomposition of Rao's entropy is not limited to ultrametric distances, such as, those extracted from phylogenetic trees. Rather, Euclidean distances between species pairs satisfy the same properties. As such, our additive partitioning of diversity allows the use of functional distances and a decomposition of functional diversity across scales. Because functional traits may reveal mechanisms underlying community structure (Mason *et al.* 2007; Mouillot *et al.* 2007; Petchey *et al.* 2007) we trust that the relative values of α -, β - and γ -functional diversity components will be investigated in the near future. In this perspective, we agree with Hardy & Senterre (2007) who suggest expressing β -diversity as a percentage of γ -diversity using $\beta_{st} = \beta/\gamma$ in order to compare β -diversity values between studies.

More generally, the decomposition of the quadratic entropy into α -, β - and γ -diversities has the potential to shed light on many macro-ecological patterns and may contribute to disentangling the effects of dispersal, niche filtering and competitive interactions in community establishment. For instance, even with no environmental gradient, we can expect an increase in species turnover with geographic distance because of dispersal limitation (e.g. Thompson & Townsend 2006). However, 'biological' turnover based on species' 'biological' differences may stay at a low-level if a niche process selects species composition according to their biological traits. Indeed we may expect very similar communities even between distant sites if environmental conditions are similar. Therefore, β -diversity based on the quadratic entropy quantifies a new facet of community similarity that can be useful to discriminate between niche vs. dispersal influences on biodiversity patterns.

References

- Champely, S. & Chessel, D. (2002) Measuring biological diversity using Euclidean metrics. *Environmental and Ecological Statistics*, **9**, 167–177.
- 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.
- Lande, R. (1996) Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos*, **76**, 5–13.
- Legendre, P., Borcard, D. & Peres-Neto, P.R. (2005) Analyzing β -diversity: partitioning the spatial variation of community composition data. *Ecological Monographs*, **75**, 435–450.
- Mason, N.W.H., Lanoiselée, C., Mouillot, D., Irz, P. & Argillier, C. (2007) Functional characters combined with null models reveal inconsistency in mechanisms of species turnover in lacustrine fish communities. *Oecologia*, **153**, 441–452.
- Mouillot, D., Dumay, O. & Tomasini, J.A. (2007) Limiting similarity, niche filtering and functional diversity in brackish lagoon fish communities. *Estuarine, Coastal and Shelf Science*, **71**, 443–456.
- Petchey, O.L., Evans, K.L., Fishburn, I.S. & Gaston, K.J. (2007) Low functional diversity and no redundancy in British avian assemblages. *Journal of Animal Ecology*, **76**, 977–985.
- Rao, C.R. (1982) Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology*, **21**, 24–43.
- Ricotta, C. (2005) Additive partitioning of Rao's quadratic diversity: a hierarchical approach. *Ecological Modelling*, **183**, 365–371.
- Thompson, R. & Townsend, C. (2006) A truce with neutral theory: local deterministic factors, species traits and dispersal limitation together determine patterns of diversity in stream invertebrates. *Journal of Animal Ecology*, **75**, 476–484.
- Whittaker, E.H. (1960) Vegetation of the Siskiyou Mountains Oregon and California. *Ecological Monograph*, **30**, 279–338.
- Whittaker, E.H. (1972) Evolution and measurement of species diversity. *Taxon*, **21**, 213–251.
- Whittaker, R.J., Willis, K.J. & Field, R. (2001) Scale and species richness: towards a general, hierarchical theory of species diversity. *Journal of Biogeography*, **28**, 453–470.

Received 26 September 2007; revision accepted 11 December 2007
Handling Editor: David Burslem