

## Can additive beta-diversity be reliably partitioned into nestedness and turnover components?

Matthews, Thomas; Aspin, Thomas; Ulrich, Werner; Baselga, Andrés; Kubota, Yasuhiro; Proios, Konstantinos ; Triantis, Kostas A. ; Whittaker, Robert J.; Strona, Giovanni

DOI:  
[10.1111/geb.12921](https://doi.org/10.1111/geb.12921)

License:  
Other (please specify with Rights Statement)

*Document Version*  
Peer reviewed version

*Citation for published version (Harvard):*  
Matthews, T, Aspin, T, Ulrich, W, Baselga, A, Kubota, Y, Proios, K, Triantis, KA, Whittaker, RJ & Strona, G 2019, 'Can additive beta-diversity be reliably partitioned into nestedness and turnover components?', *Global Ecology and Biogeography*, vol. 28, no. 8, pp. 1146-1154. <https://doi.org/10.1111/geb.12921>

[Link to publication on Research at Birmingham portal](#)

**Publisher Rights Statement:**  
Checked for eligibility: 19/03/2019

This is the peer reviewed version of the following article: Matthews TJ, Aspin TWH, Ulrich W, et al. Can additive beta diversity be reliably partitioned into nestedness and turnover components? *Global Ecol Biogeogr.* 2019., which has been published in final form at: <https://doi.org/10.1111/geb.12921>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

### General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

### Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.

1 Submission to: Global Ecology & Biogeography

2 Article Type: Research Paper

3 **Can additive beta-diversity be reliably partitioned into nestedness and turnover**  
4 **components?**

5

6 Thomas J. Matthews<sup>1,2,3</sup>, Thomas W.H. Aspin<sup>1</sup>, Werner Ulrich<sup>4</sup>, Andrés Baselga<sup>5</sup>, Yasuhiro  
7 Kubota<sup>6</sup>, Konstantinos Proios<sup>7</sup>, Kostas A. Triantis<sup>7</sup>, Robert J. Whittaker<sup>8,9</sup>, Giovanni Strona<sup>10</sup>

8

9 <sup>1</sup>GEES (School of Geography, Earth and Environmental Sciences), The University of  
10 Birmingham, Birmingham, B15 2TT, UK.

11 <sup>2</sup>CE3C – Centre for Ecology, Evolution and Environmental Changes/Azorean Biodiversity  
12 Group and Univ. dos Açores – Depto de Ciências e Engenharia do Ambiente, PT-9700-042,  
13 Angra do Heroísmo, Açores, Portugal.

14 <sup>3</sup>Birmingham Institute of Forest Research, University of Birmingham, Edgbaston,  
15 Birmingham, B15 2TT, UK

16 <sup>4</sup>Faculty of Biology and Environment Protection, Nicolaus, Copernicus University, Lwowska  
17 1, 87-100 Toruń, Poland

18

19 <sup>5</sup>Departamento de Zoología, Genética y Antropología Física, Facultad de Biología,  
20 Universidad de Santiago de Compostela, Rúa Lope Gómez de Marzoa, 15782, Santiago de  
21 Compostela, Spain

22

23 <sup>6</sup>Faculty of Science, University of the Ryukyus, 1 Senbaru, Nishihara, Okinawa, 903-0213,  
24 Japan.

25 <sup>7</sup>Department of Ecology and Taxonomy, Faculty of Biology, National and Kapodistrian  
26 University of Athens, Athens GR-15784, Greece.

27 <sup>8</sup>School of Geography and the Environment, University of Oxford, South Parks Road, OX1  
28 3QY, Oxford, UK.

29 <sup>9</sup>Center for Macroecology, Evolution and Climate, National Museum of Natural History,  
30 University of Copenhagen, Universitetsparken 15, 2100 Copenhagen, Denmark.

31 <sup>10</sup>European Commission, Joint Research Centre, Directorate D - Sustainable Resources –  
32 Bio-Economy Unit, Via Enrico Fermi 274 9, 21027 Ispra (VA), Italy

33

34

35 \*Correspondence: Thomas J. Matthews, School of Geography, Earth and Environmental  
36 Sciences, University of Birmingham, Birmingham, B15 2TT, UK

37 Email: t.j.matthews@bham.ac.uk

38 Running header: A warning on partitioning additive beta-diversity

39 **ACKNOWLEDGMENTS**

40 José María Fernández-Palacios and Juliano Cabral kindly reviewed an earlier version of the  
41 paper. Brian McGill, Petr Keil and two anonymous reviewers provided useful comments that  
42 greatly improved the paper.

43 **DATA ACCESSIBILITY**

44 The incidence matrices and the R code used to run the analyses are provided in a GitHub  
45 repository (<https://github.com/txm676/Partitioning-additive-beta>). The repository has been  
46 archived on the Zenodo research data repository (DOI: 10.5281/zenodo.2595322).

47

48 **BIOSKETCH**

49 **Tom Matthews** is a macroecologist and biogeographer at the University of Birmingham,  
50 UK. He is interested in the application of macroecological methods to global environmental  
51 change questions, and his previous work has focused on the impacts of habitat fragmentation  
52 and the form of the species–area relationship in fragmented landscapes.

53

54

55 **ABSTRACT**

56 **Aims:** Quantifying  $\beta$ -diversity (differences in the composition of communities) is central to  
57 many ecological studies. There are many  $\beta$ -diversity metrics, falling mostly into two  
58 approaches: variance-based (e.g. the Sørensen index), or diversity partitioning (e.g. additive  
59  $\beta$ -diversity). The former cannot be used when species–sites matrices are unavailable (which  
60 is often the case in island biogeography in particular) and only species richness data are  
61 provided. Recently, efforts have been made to partition additive  $\beta$ -diversity, a metric  
62 calculated using only  $\alpha$ -diversity and  $\gamma$ -diversity, into nestedness and turnover components  
63 (termed here ‘richness-only  $\beta$ -diversity partitioning’). We set out to test whether this form of  
64  $\beta$ -diversity partitioning generates interpretable results, comparable with metrics based on  
65 species incidence  $\beta$ -diversity partitioning.

66 **Location:** Global

67 **Time period:** Present day

68 **Major taxa studied:** Multiple taxa

69 **Methods:** We first provide a brief review of  $\beta$ -diversity partitioning methods, with a  
70 particular focus on the development of richness-only  $\beta$ -diversity partitioning. Second, we use  
71 254 empirical incidence matrices (provided with the paper) sourced from the literature to  
72 measure turnover and nestedness using incidence  $\beta$ -diversity partitioning, comparing the  
73 resulting values with those calculated using richness-only  $\beta$ -diversity.

74 **Results:** We provide an account of the emergence of  $\beta$ -diversity partitioning, with particular  
75 reference to the analysis of richness-only datasets and to the definition and usage of the  
76 relevant metrics. Analytically, we report weak correlations between turnover and nestedness  
77 calculated using the two different approaches. We show that this is because identical values

78 of  $\alpha$ -diversity and  $\gamma$ -diversity can correspond to incidence matrices with a range of different  
79 structures.

80 **Main conclusions:** Our results demonstrate that the use of richness-only  $\beta$ -diversity  
81 partitioning to measure turnover and nestedness is problematic and can produce patterns  
82 unrelated to conventional measures of turnover and nestedness. We therefore recommend that  
83 more accurate definitions are adopted for these terms in future studies.

84

## 85 INTRODUCTION

86 Quantifying differences in the composition of communities (i.e. measuring  $\beta$ -diversity) and  
87 testing the prevalence of nestedness in ecological communities are central to many ecological  
88 studies (Whittaker, 1960; Rosenzweig, 1995; Ulrich, Almeida-Neto, & Gotelli, 2009).

89 Various metrics have been proposed to measure  $\beta$ -diversity (Koleff, Gaston, & Lennon,  
90 2003; Tuomisto, 2010; Anderson et al., 2011; Chao, Chiu, & Hsieh, 2012) and many of these  
91 can be broadly divided into variance-based approaches (e.g. the Sørensen and Jaccard  
92 indices) and diversity partitioning-based approaches (Legendre & De Cáceres, 2013; Chao &  
93 Chiu, 2016). Recent work has bridged these two approaches to calculating  $\beta$ -diversity (Chao  
94 et al., 2012; Chao & Chiu, 2016). Nestedness, as originally conceived within island  
95 biogeography, refers to the ordered loss/gain of species along a richness gradient, whereby  
96 each larger assemblage tends to contain all the members of the previously considered site  
97 (Patterson & Atmar, 1986; Ulrich et al., 2009; Matthews, Cottee-Jones, & Whittaker, 2015).

98 The calculation of most nestedness and variance-based  $\beta$ -diversity metrics requires  
99 knowledge of species occurrences at each site, coded in a binary presence-absence matrix

100 (herein, incidence matrix), with species in rows and sites in columns (e.g. Legendre &  
101 Legendre, 1983; Koleff et al., 2003; Ulrich et al., 2009). Table 1 provides a glossary of the

102 (many) different metrics discussed in the present study, and Fig. 1 provides an illustration of  
103 the different metrics.

#### 104 **Diversity partitioning and richness-only $\beta$ -diversity partitioning: an overview**

105 Over time, as research questions and meta-analytical tools have developed, the published  
106 literature has become an increasingly important source of data to extend the power of  
107 analyses via data mining. This is particularly true in regard to studies that conduct meta-  
108 analyses of  $\beta$ -diversity and nestedness (e.g. Cabral, Weigelt, Kissling, & Kreft, 2014;  
109 Matthews et al., 2015). However, in many cases the full incidence matrix for a set of sites is  
110 not available in previously published studies, and only the number of species in each site ( $\alpha$ -  
111 diversity) and the regional diversity ( $\gamma$ -diversity) can be retrieved (e.g. Cowie, 1995). This is  
112 a particular problem in island biogeography, where several meta-analyses of ecological  
113 patterns on islands are based on these simple data (see, for example, various studies  
114 investigating species–area relationships (SAR), such as Triantis, Guilhaumon, & Whittaker,  
115 2012; Matthews, Guilhaumon, Triantis, Borregaard, & Whittaker, 2016). Authors have  
116 attempted to get around this problem by estimating  $\beta$ -diversity and nestedness through  
117 diversity partitioning approaches (e.g. Chiarucci et al., 2010; Sfenthourakis & Panitsa, 2012;  
118 Cabral et al., 2014; see Zhang et al., 2014, for a terrestrial example).

119 Estimating  $\beta$ -diversity using diversity partitioning has a long history in ecology (Whittaker,  
120 1960, 1965; MacArthur, Recher, & Cody, 1966; Lande, 1996; Jost, 2007; Tuomisto, 2010).

121 The two main diversity partitioning approaches used are additive ( $\beta_{\text{Total}} = \gamma - \alpha$ ), where  
122  $\beta_{\text{Total}}$  is the amount by which regional diversity ( $\gamma$ ) exceeds the mean diversity of a set of sites  
123 ( $\alpha$ ), and multiplicative ( $\beta_{\text{Mult}} = \gamma / \alpha$ ), where  $\beta_{\text{Mult}}$  is the regional-to-local diversity ratio  
124 (Whittaker, 1960; Veech, Summerville, Crist, & Gering, 2002; Crist, Veech, Gering, &  
125 Summerville, 2003; Tuomisto 2010). Their relative merits have been much debated (see

126 Lande, 1996; Crist et al., 2003; Jost, 2007; Baselga, 2010a; Tuomisto, 2010; Veech & Crist,  
127 2010), and recent work has shown how they (i.e. additive and multiplicative diversity  
128 partitioning) are mathematically linked (Chao et al., 2012). Additively partitioned  $\beta$ -diversity  
129 ( $\beta_{\text{Total}}$ ), which is the focus of this present study, has been shown to be dependent on both  
130  $\gamma$  and the number of sites ( $N$ ), and it has thus been argued that  $\beta_{\text{Total}}$  should be normalised (by  
131  $\gamma$  and  $N$ ;  $\beta_{\text{Stan}}$ ) in order to compare  $\beta$ -diversity values (Chao et al., 2012).

132 Using an additive partitioning approach,  $\beta_{\text{Total}}$  has been further partitioned into two sub-  
133 components, which have been argued to measure nestedness and turnover (we term this  
134 ‘richness-only  $\beta$ -diversity partitioning’). This is different from the partitioning of  
135 dissimilarity indices (such as Sørensen and Jaccard indices), through analysis of an incidence  
136 matrix (we term this ‘incidence  $\beta$ -diversity partitioning’), into turnover and nestedness-  
137 resultant dissimilarity / richness difference components (Baselga, 2010b, 2012; Carvalho,  
138 Cardoso, & Gomes, 2012). The use of richness-only  $\beta$ -diversity partitioning appears to have  
139 been based on the approach of Crist & Veech (2006), who used the power law SAR model to  
140 partition  $\beta_{\text{Total}}$  in isolated habitats into two components:  $\beta_{\text{Area}}$ , which is intended to describe  
141 how much of  $\beta_{\text{Total}}$  is due to the area effect, and  $\beta_{\text{Replace}}$ , which is intended to describe how  
142 much is explained by other factors (see Crist & Veech, 2006).  $\beta_{\text{Area}}$  is defined by Crist &  
143 Veech (2006, p.928) as “the mean deviation between the species richness of the largest  
144 habitat patch and the species richness of smaller patches”. Subsequent studies have used the  
145  $\beta_{\text{Area}}$  component as a measure of nestedness, and the  $\beta_{\text{Replace}}$  component as a measure of  
146 replacement/turnover (e.g. Chiarucci et al., 2010; Sfenthourakis & Panitsa, 2012; Cabral et  
147 al., 2014; Zhang et al., 2014). It is important to note that the interpretation of  $\beta_{\text{Area}}$  and  $\beta_{\text{Replace}}$   
148 as measures of nestedness and turnover, respectively, was not necessarily implied in the  
149 original study (Crist & Veech, 2006), a fact recognised by at least one of the subsequent  
150 studies (Cabral et al., 2014).

151 To take one study that used richness-only  $\beta$ -diversity partitioning as an example, Chiarucci et  
152 al. (2010, p.86), in their study of plants on a variety of Macaronesian islands, use the term  
153  $\beta_{\text{Nestedness}}$  (rather than  $\beta_{\text{Area}}$ ), stating that “the first component of [additive]  $\beta$ -diversity  
154 ( $\beta_{\text{Nestedness}}$ ) quantified the degree of nestedness of the flora,” whilst the “second  $\beta$  component  
155 ( $\beta_{\text{Replacement}}$ ) measure[s] the differences in species composition among the flora of the islands  
156 within an archipelago, and [is] a measure of the compositional differences across islands.”  
157 The authors then proceed to make inferences regarding the nestedness of their data; for  
158 example, “the higher importance of  $\beta_{\text{Nestedness}}$  for pteridophytes indicated that, for this taxon,  
159 the flora of each island is largely formed by a subset of species that make up the archipelago  
160 flora” (Chiarucci et al., 2010, p. 89). This example, and others (e.g. Sfenthourakis & Panitsa,  
161 2012; Cabral et al., 2014; Zhang et al., 2014), illustrates that colleagues have started to use  
162 richness-only  $\beta$ -diversity partitioning in their research; the use of the method and its  
163 implications is not confined to those four cited studies and a simple Google Scholar search  
164 indicates these four papers have been cited over 80 times. It is therefore timely to assess the  
165 implications of this approach and how well the richness-only  $\beta$ -diversity partitions  
166 correspond with conventional measures of nestedness and turnover, as this could constitute a  
167 useful analytical tool if it can be shown to be robust.

168 In this article, we use a dataset of 254 incidence matrices (details below) to assess to what  
169 extent nestedness and turnover calculated by richness-only  $\beta$ -diversity partitioning (i.e.  
170 partitioning additive  $\beta$ -diversity, i.e.  $\beta_{\text{Total}}$ ) are congruent with nestedness and compositional  
171 difference metrics calculated using the full incidence matrix. It is important to re-stress that  
172 we are focused on the issues surrounding the use of richness-only  $\beta$ -diversity partitioning (i.e.  
173 when there is no incidence matrix and thus no information on which species are present on  
174 which island; e.g. as employed by Chiarucci et al., 2010), and not incidence  $\beta$ -diversity  
175 partitioning (e.g. the partitioning of Sørensen dissimilarity; e.g. Baselga, 2012). As  $\beta_{\text{Total}}$  is



176 known to be problematic due to its dependency on both  $\gamma$  and the number of sites (N), we  
177 might expect the partitioned components of  $\beta_{\text{Total}}$  to also have issues. However, this  
178 possibility has not previously been explored, and it is important that any problems with the  
179 approach are highlighted to avoid the proliferation of incorrect metrics (or at least the  
180 incorrect interpretation of particular metrics) in the ecological literature.

## 181 **MATERIALS AND METHODS**

182 To compare nestedness and turnover calculated using richness-only  $\beta$ -diversity partitioning  
183 (i.e. additive partitioning of  $\beta_{\text{Total}}$ ) with nestedness and compositional difference metrics  
184 based on incidence  $\beta$ -diversity partitioning (i.e. analysis of the full incidence matrix), we  
185 used a collection of island incidence matrices that we sourced from the literature. Briefly, the  
186 database contains 254 incidence matrices of various taxa from different island systems (each  
187 representing a geographically coherent set of islands), including all major island types such as  
188 volcanic oceanic islands, continental-shelf islands, atolls, and habitat islands. Unlike other  
189 collections of incidence matrices (e.g. Atmar & Patterson, 1995; Strona, Ulrich, & Gotelli,  
190 2017), our database is comprised solely of island datasets and thus allows us to make general  
191 conclusions regarding patterns of interest in islands (Whittaker & Fernández-Palacios, 2007;  
192 Matthews, 2015). The full database will be published as part of a separate upcoming study,  
193 but the set of 254 incidence matrices are available from GitHub (txm676/ Partitioning-  
194 additive-beta). The repository has been archived on the Zenodo research data repository  
195 (DOI: 10.5281/zenodo.2595322).

196 For each of the 254 datasets, we calculated additive  $\beta$ -diversity ( $\beta_{\text{Total}}$ ). Thus, the incidence  
197 matrices were not used; we simply used mean  $\alpha$  (the average richness of the set of islands in a  
198 dataset) and  $\gamma$  (the overall species richness of the set of islands in a dataset). We also  
199 calculated normalised  $\beta$ -diversity ( $\beta_{\text{Stan}}$ ) using the approach in Chao et al. (2012):

200 
$$\beta_{Stan} = \frac{\beta_{Total}}{(1-\frac{1}{N})^Y},$$

201 where N is the number of islands. We then calculated the  $\beta_{Nestdedness}$  and  $\beta_{Replacement}$  (Table 1)  
 202 partitions of  $\beta_{Total}$  (i.e. richness-only  $\beta$ -diversity partitioning) using the approach of Chiarucci  
 203 et al. (2010). This approach differs slightly from that employed by Crist & Veech (2006) in  
 204 that it uses the maximum observed richness of an island in the dataset rather than maximum  
 205 richness predicted by the power law SAR model. However, the two approaches produce  
 206 highly correlated values (Spearman's rho = 0.88,  $P < 0.001$ , for the correlation between the  
 207  $\beta_{Nestdedness}$  of Chiarucci et al. (2010) and the  $\beta_{Area}$  of Crist and Veech (2006) based on the 254  
 208 empirical matrices). Following Chiarucci et al. (2010)  $\beta_{Nestdedness}$  was calculated using the  
 209 equation:

210 
$$\beta_{Nestdedness} = \frac{1}{N} \sum_{i=1}^N (S_{max} - S_i), \quad (1)$$

211 where  $S_i$  is the number of species on the i-th island, and  $S_{max}$  is the number of species on the  
 212 most species rich island.  $\beta_{Replacement}$  can then be calculated using the equation:

213 
$$\beta_{Replacement} = \beta_{Total} - \beta_{Nestdedness}. \quad (2)$$

214 We note that eq. 1 can be reformulated to clarify the meaning of  $\beta_{Nestdedness}$ :

215 
$$\beta_{Nestdedness} = \frac{1}{N} \sum_{i=1}^N (S_{max} - S_i) = S_{max} - \frac{1}{N} \sum_{i=1}^N S_i = S_{max} - \alpha. \quad (3)$$

216 This is in accordance with the initial definition of  $\beta_{Area}$  by Crist & Veech (2006). As the  
 217 average local diversity,  $\alpha$ , is independent of the number of sites considered  $\beta_{Nestdedness}$  can be  
 218 derived from the knowledge of two sites having  $S_{max}$  and  $S_r$  species,

219 
$$\beta_{Nestdedness} = S_{max} - \frac{S_{max} + S_r}{2} = \frac{S_{max} - S_r}{2}, \quad (4)$$

220 Therefore,  $\beta_{\text{Nestedness}}$  reduces to a difference in species richness between two sites, irrespective  
 221 of the occurrence of joint species among these sites. This contradicts the basic definition of  
 222 nestedness, which asserts the existence of an ordered set of subsamples (Patterson & Atmar,  
 223 1986; Almeida-Neto, Guimarães, Guimarães, Loyola, & Ulrich, 2008). Taking the extreme  
 224 case of there being no shared species, and thus with the maximum possible species turnover,  
 225  $\beta_{\text{Nestedness}}$  can take any of the full range of possible values between zero and  $S_{\text{max}} - \alpha$ .

226 We also measured the compositional difference between islands in a dataset using the  
 227 Sørensen dissimilarity index computed on the full incidence matrix, using the ‘betapart’ R  
 228 package (version 1.4-1, Baselga, Orme, Villeger, De Bortoli, & Leprieur, 2017). Overall  
 229 compositional difference was calculated using Sørensen multi-site dissimilarity ( $\beta_{\text{Sor}}$ ; see  
 230 Table 1),

$$231 \quad \beta_{\text{Sor}} = \frac{\left[ \sum_{i < j} \min(b_{ij}, b_{ji}) \right] + \left[ \sum_{i < j} \max(b_{ij}, b_{ji}) \right]}{2 \left[ \sum_i S_i - S_T \right] + \left[ \sum_{i < j} \min(b_{ij}, b_{ji}) \right] + \left[ \sum_{i < j} \max(b_{ij}, b_{ji}) \right]} \quad (5)$$

232 where  $S_i$  is the total number of species in site  $i$ ,  $S_T$  is the total number of species in all sites  
 233 considered together and  $b_{ij}$ ,  $b_{ji}$  are the number of species exclusive to sites  $i$  and  $j$ ,  
 234 respectively. Sørensen multi-site dissimilarity was partitioned (i.e. incidence  $\beta$ -diversity  
 235 partitioning) into the turnover component (Simpson multi-site dissimilarity;  $\beta_{\text{Sim}}$ ; see Table  
 236 1),

$$237 \quad \beta_{\text{Sim}} = \frac{\left[ \sum_{i < j} \min(b_{ij}, b_{ji}) \right]}{\left[ \sum_i S_i - S_T \right] + \left[ \sum_{i < j} \min(b_{ij}, b_{ji}) \right]} \quad (6)$$

238 and the nestedness component (Nestedness-resultant multi-site dissimilarity;  $\beta_{\text{Sne}}$ , Baselga,  
 239 2010b, 2012),

$$240 \quad \beta_{\text{Sne}} = \frac{\left[ \sum_{i < j} \max(b_{ij}, b_{ji}) \right] - \left[ \sum_{i < j} \min(b_{ij}, b_{ji}) \right]}{2 \left[ \sum_i S_i - S_T \right] + \left[ \sum_{i < j} \min(b_{ij}, b_{ji}) \right] + \left[ \sum_{i < j} \max(b_{ij}, b_{ji}) \right]} * \frac{\sum_i S_i - S_T}{\left[ \sum_i S_i - S_T \right] + \left[ \sum_{i < j} \min(b_{ij}, b_{ji}) \right]} \quad (7)$$

241 As  $\beta_{\text{Sne}}$  is conceptually distinct from ‘true’ nestedness (Almeida-Neto, Frensel, & Ulrich,  
 242 2012; Baselga 2012), we also measured the nestedness of each dataset with the NODF  
 243 (‘nestedness metric based on overlap and decreasing fill’) metric (Almeida-Neto et al., 2008).  
 244 According to NODF, an incidence matrix sorted in decreasing order of marginal totals is  
 245 maximally nested when there is complete overlap of presence values (1s) from the right to the  
 246 left column and from the bottom to the top row, and no ties in both the row and column  
 247 marginal totals. We chose NODF as it is widely considered to be the most appropriate  
 248 measure of nestedness (Ulrich et al., 2009). NODF was computed on the full incidence  
 249 matrix, using the vegan R package (version 2.4-5, Oksanen et al., 2017). We used a  
 250 combination of graphical plots and Kendall rank correlation tests to assess the correlation  
 251 between the richness-only  $\beta$ -diversity partitioning metrics,  $\beta_{\text{Nestedness}}$  and  $\beta_{\text{Replacement}}$ , and the  
 252 incidence  $\beta$ -diversity partitioning metrics,  $\beta_{\text{Sim}}$  and  $\beta_{\text{Sne}}$ , and also NODF. We used the Kendall  
 253 rank correlation test rather than Spearman rank correlations as the former is argued to provide  
 254 better results with small sample sizes. To illustrate that the issues we highlight lie with  
 255 richness-only  $\beta$ -diversity partitioning and not with additive  $\beta$ -diversity ( $\beta_{\text{Total}}$ ) itself, we also  
 256 assessed the correlation between  $\beta_{\text{Stan}}$  and  $\beta_{\text{Sor}}$  using a Kendall rank correlation test.  
 257 To determine whether our results were consistent across different types of island datasets, we  
 258 divided our main dataset a number of different ways. First, we split the datasets into true  
 259 islands (oceanic and continental-shelf islands, and islands within natural lakes) and habitat

260 islands (all other datasets). Second, we calculated the quantiles, using all datasets, of each of:  
261 the number of islands,  $\alpha$ , and  $\gamma$ . For each factor, we then took the datasets representing the  
262 top and bottom quantiles (lowest and highest 25%) to be individual subsets. We then repeated  
263 the above analyses using each individual subset. Finally, we repeated the main correlation  
264 tests using the multisite version of Jaccard dissimilarity calculated using both the Baselga  
265 (2012) and the Carvalho et al. (2012) approaches. The above analyses were undertaken using  
266 the R programming language (Version 3.5.2, R Core Team, 2017), and the R code used to run  
267 the analyses is provided in a GitHub repository (txm676/ Partitioning-additive-beta).

## 268 **RESULTS**

269 Considering the full set of 254 empirical island incidence matrices, we found low (often very  
270 low) correlation between the relevant metrics. The lowest correlation was between  $\beta_{\text{Nestedness}}$   
271 and  $\beta_{\text{Sne}}$  (Kendall's tau = -0.04;  $P$ -value = 0.36), followed by  $\beta_{\text{Nestedness}}$  and NODF (Kendall's  
272 tau = -0.11;  $P$ -value = 0.01),  $\beta_{\text{Replacement}}$  and  $\beta_{\text{Sor}}$  (Kendall's tau = 0.23;  $P$ -value < 0.001) and  
273  $\beta_{\text{Replacement}}$  and  $\beta_{\text{Sim}}$  (Kendall's tau = 0.43;  $P$ -value < 0.001). As expected, there was a higher  
274 correlation between  $\beta_{\text{Stan}}$  and  $\beta_{\text{Sor}}$  (Kendall's tau = 0.56;  $P$ -value < 0.001). The plots of these  
275 relationships are provided in Figure 2; certain variables were logged prior to plotting (but not  
276 analysis) to ease the visual interpretation (see the legend of Figure 2). A complete set of  
277 pairwise scatter plots (along with their correlations) for all variables is provided as Figure S1  
278 in Appendix S1.

279 Re-running the analyses using the different dataset subsets generated broadly similar results  
280 (Table S1 in Appendix S1). The main difference was for the correlation between  $\beta_{\text{Nestedness}}$  and  
281  $\beta_{\text{Sne}}$ , whereby for three subsets ( $\alpha > 25$ ,  $\gamma < 20$  and  $\gamma > 77$ ) the correlation was positive and  
282 significant, although the correlation coefficient was less than 0.50 in all three cases (Table  
283 S1). The results of the analyses using the multisite version of Jaccard dissimilarity calculated

284 using both the Baselga (2012) and the Carvalho et al. (2012) approaches were similar to the  
285 main results and are not discussed further (Appendix S1).

286

## 287 **DISCUSSION**

288 We have explored to what extent nestedness and turnover calculated through richness-only  $\beta$ -  
289 diversity partitioning (i.e. partitioning  $\beta_{\text{Total}}$ ) are congruent with nestedness and compositional  
290 difference metrics calculated using incidence  $\beta$ -diversity partitioning. We find that neither  
291 partition of  $\beta_{\text{Total}}$  provides good measures of what is commonly regarded as nestedness or  
292 turnover. Our results lead us to caution the interpretation of the additive sub-components of  
293  $\beta_{\text{Total}}$  as metrics of nestedness and replacement/turnover. The problem is most acute when  
294 considering nestedness:  $\beta_{\text{Nestedness}}$  and NODF calculated using the empirical matrices were in  
295 fact weakly negatively correlated ( $\tau = -0.11$ ). A thought experiment that further illustrates  
296 the issue with  $\beta_{\text{Nestedness}}$  is provided in Appendix S2. Thus,  $\beta_{\text{Nestedness}}$  should simply be  
297 interpreted as representing area effects, as originally proposed by Crist & Veech (2006).  
298 Based on analyses of the empirical matrices, the correlation between  $\beta_{\text{Replacement}}$  and  $\beta_{\text{Sim}}$  was  
299 also low ( $\tau = 0.43$ ) and we do not recommend using  $\beta_{\text{Replacement}}$  in future studies to measure  
300 turnover. This is likely due to the fact that  $\beta_{\text{Replacement}}$  is not normalised (by either  $N$  or  $\gamma$ ).  
301 Rather, dissimilarity measures should be preferred when the full incidence matrix is available  
302 (Roden et al., 2018), and  $\beta_{\text{Stan}}$  or an equivalent metric (see Chao et al., 2012; Chao & Chiu,  
303 2016) should be used when it is not available.

304 Partitioning methods based on only  $\gamma$  and  $\alpha$ -diversity (i.e. richness-only  $\beta$ -diversity  
305 partitioning) intrinsically disregard the species composition of each site which, depending on  
306 the structure of the underlying incidence matrix, may have a disproportionate effect on the  
307 accuracy of partitioned measures. This is rather intuitive: for a given combination of  $\gamma$  and  $\alpha$ -

308 diversity values, one can generate a very large number of different matrices, due to the fact  
309 that neither  $\gamma$  nor  $\alpha$  include information on the number of sites in the system. However, even  
310 if we fix the number of sites to a given value  $N$  (in reality the number of islands in an  
311 archipelago should be known), the number of matrices with  $\gamma$  species,  $N$  sites and average  
312 species richness per site equal to  $\alpha$  might still be very large. In turn, this means that a given  
313 partitioned measure of  $\beta$ -diversity or nestedness may potentially correspond to a broad array  
314 of different values of their matrix-wide counterparts. To illustrate this concept, we chose a  
315 random block of 20 matrices from the set of 254 incidence matrices used in the main  
316 analysis, and we used a simple procedure to explore how much, for each matrix, we could  
317 modify the matrix structure towards either higher or lower  $\beta$ -diversity and nestedness relative  
318 to the observed values, without altering  $\gamma$ ,  $\alpha$  or  $N$  (see Appendix S3 for details). For clarity,  
319 results for a subset of five matrices are reported in Figure 3, whilst we provide separate plots  
320 for each of the 20 matrices in Figure S2 in Appendix S3. As expected, for a given matrix, the  
321 same  $\gamma$ ,  $\alpha$  and  $N$  can result in a wide range of internal matrix structure, especially in terms of  
322 nestedness. In turn, this makes it very difficult to draw parallels between the matrix-wide and  
323 the partitioned concepts of turnover and nestedness, as the latter might span a very large  
324 spectrum of cases that can only be finely discriminated using the former.

325 The results of our main analyses were largely consistent for the different subsets of datasets,  
326 although there were a few subsets ( $\alpha > 25$ ,  $\gamma < 20$  and  $\gamma > 77$ ) for which the correlation  
327 between  $\beta_{\text{Nestedness}}$  and  $\beta_{\text{Sne}}$  was positive and significant. In datasets with low  $\gamma$ , the richness  
328 differences between islands are likely constrained and thus  $\beta_{\text{Nestedness}}$  and  $\beta_{\text{Sne}}$  are both  
329 restricted to low values, which could explain the positive correlation between  $\beta_{\text{Nestedness}}$  and  
330  $\beta_{\text{Sne}}$  for the low  $\gamma$  subset. The reason for the positive correlations observed in the high  $\gamma$  and  
331 high  $\alpha$  subsets is unclear, but may point towards a joint dependency between the metrics and  
332  $\gamma$ . That being said, it should be noted that, whilst the correlations were significant, the

333 coefficients were relatively low (i.e. 0.23, 0.24 and 0.44) and thus our ability to make  
334 conclusions based on these results is limited.

335 The results of the present study clearly illustrate the issues with using richness-only  $\beta$ -  
336 diversity partitioning to measure species turnover and nestedness, and we recommend that, if  
337 using this approach, more accurate definitions are adopted for these terms in future studies.  
338 Readers are directed to Ulrich et al. (2009), Baselga (2012) and Chao & Chiu (2016) for  
339 discussion of other nestedness and  $\beta$ -diversity metrics. Perhaps more generally, the results of  
340 this study also highlight the benefits of the deposition of datasets from published studies in  
341 data archives.

342

#### 343 **REFERENCES**

344 Almeida-Neto, M., Guimarães, P., Guimarães, P. R., Loyola, R. D., & Ulrich, W. (2008). A  
345 consistent metric for nestedness analysis in ecological systems: reconciling concept and  
346 measurement. *Oikos*, *117*, 1227-1239.

347 Almeida-Neto, M., Frensel, D. M. B., & Ulrich, W. (2012). Rethinking the relationship  
348 between nestedness and beta diversity: a comment on Baselga (2010). *Global Ecology  
349 and Biogeography*, *21*, 772-777.

350 Atmar, W., & Patterson, B. D. (1995). The nestedness temperature calculator: a Visual Basic  
351 program. Chicago: AICS Research. Retrieved from [http://aics-  
352 research.com/nestedness/tempcalc.html](http://aics-research.com/nestedness/tempcalc.html)

353 Baselga, A. (2010a). Multiplicative partition of true diversity yields independent alpha and  
354 beta components; additive partition does not. *Ecology*, *91*, 1974-1981.

355 Baselga, A. (2010b). Partitioning the turnover and nestedness components of beta diversity.  
356 *Global Ecology and Biogeography*, *19*, 134-143.

357 Baselga, A. (2012). The relationship between species replacement, dissimilarity derived from  
358 nestedness, and nestedness. *Global Ecology and Biogeography*, *21*, 1223-1232.



- 359 Baselga, A., Orme, D., Villegger, S., De Bortoli, J., & Leprieur, F. (2017). betapart:  
360 partitioning beta diversity into turnover and nestedness components (1): R Package  
361 version 1.4-1.
- 362 Cabral, J. S., Weigelt, P., Kissling, W. D., & Kreft, H. (2014). Biogeographic, climatic and  
363 spatial drivers differentially affect  $\alpha$ -,  $\beta$ - and  $\gamma$ -diversities on oceanic archipelagos.  
364 *Proceedings of the Royal Society B: Biological Sciences*, 281, 20133246.
- 365 Carvalho, J.C., Cardoso, P. & Gomes, P. (2012). Determining the relative roles of species  
366 replacement and species richness differences in generating beta-diversity patterns.  
367 *Global Ecology and Biogeography*, 21, 760-771.
- 368 Chao, A., & Chiu, C.-H. (2016). Bridging the variance and diversity decomposition  
369 approaches to beta diversity via similarity and differentiation measures. *Methods in*  
370 *Ecology and Evolution*, 7, 919-928.
- 371 Chao, A., Chiu, C.-H., & Hsieh, T. C. (2012). Proposing a resolution to debates on diversity  
372 partitioning. *Ecology*, 93, 2037-2051.
- 373 Chiarucci, A., Bacaro, G., Ramón Arévalo, J., Domingo Delgado, J., & María Fernández-  
374 Palacios, J. (2010). Additive partitioning as a tool for investigating the flora diversity in  
375 oceanic archipelagos. *Perspectives in Plant Ecology, Evolution and Systematics*, 12, 83-  
376 91.
- 377 Cowie, R. H. (1995). Variation in species diversity and shell shape in Hawaiian Land Snails:  
378 in situ speciation and ecological relationships. *Evolution*, 49, 1191-1202.
- 379 Crist, T. O., & Veech, J. A. (2006). Additive partitioning of rarefaction curves and species-  
380 area relationships: unifying  $\alpha$ -,  $\beta$ - and  $\gamma$ -diversity with sample size and habitat area.  
381 *Ecology Letters*, 9, 923-932.
- 382 Crist, T. O., Veech, J. A., Gering, J. C., & Summerville, K. S. (2003). Partitioning species  
383 diversity across landscapes and regions: a hierarchical analysis of  $\alpha$ ,  $\beta$ , and  $\gamma$  diversity.  
384 *The American Naturalist*, 162, 734-743.
- 385 Jost, L. (2007). Partitioning diversity into independent alpha and beta components. *Ecology*,  
386 88, 2427-2439.

- 387 Koleff, P., Gaston, K. J., & Lennon, J. J. (2003). Measuring beta diversity for presence–  
388 absence data. *Journal of Animal Ecology*, *72*, 367-382.
- 389 Lande, R. (1996). Statistics and partitioning of species diversity, and similarity among  
390 multiple communities. *Oikos*, *76*, 5-13.
- 391 Legendre, P., & De Cáceres, M. (2013). Beta diversity as the variance of community data:  
392 dissimilarity coefficients and partitioning. *Ecology Letters*, *16*, 951-963.
- 393 Legendre, P., & Legendre, L. (1983). *Numerical ecology*. Amsterdam: Elsevier.
- 394 MacArthur, R., Recher, H., & Cody, M. (1966). On the relation between habitat selection and  
395 species diversity. *The American Naturalist*, *100*, 319-332.
- 396 Matthews, T. J. (2015). Analysing and modelling the impact of habitat fragmentation on  
397 species diversity: a macroecological perspective. *Frontiers of Biogeography*, *7*, 60-68.
- 398 Matthews, T. J., Cottee-Jones, H. E. W., & Whittaker, R. J. (2015). Quantifying and  
399 interpreting nestedness in habitat islands: a synthetic analysis of multiple datasets.  
400 *Diversity and Distributions*, *21*, 392-404.
- 401 Matthews, T. J., Guilhaumon, F., Triantis, K. A., Borregaard, M. K., & Whittaker, R. J.  
402 (2016). On the form of species–area relationships in habitat islands and true islands.  
403 *Global Ecology and Biogeography*, *25*, 847–858
- 404 Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlenn, D., . . . Wagner,  
405 H. (2017). vegan: community ecology package. R package version 2.4-5.
- 406 Patterson, B. D., & Atmar, W. (1986). Nested subsets and the structure of insular mammalian  
407 faunas and archipelagos. *Biological Journal of the Linnean Society*, *28*, 65-82.
- 408 R Core Team. (2017). R: a language and environment for statistical computing (Version  
409 3.5.2). Vienna, Austria: R foundation for statistical computing. Retrieved from  
410 <https://www.R-project.org/>
- 411 Roden, V. J., Kocsis, Á. T., Zuschin, M., & Kiessling, W. (2018). Reliable estimates of beta  
412 diversity with incomplete sampling. *Ecology*, *99*, 1051-1062.
- 413 Rosenzweig, M. L. (1995). *Species diversity in space and time*. Cambridge: Cambridge  
414 University Press.

- 415 Sfenthourakis, S., & Panitsa, M. (2012). From plots to islands: species diversity at different  
416 scales. *Journal of Biogeography*, *39*, 750-759.
- 417 Strona, G., Ulrich, W., & Gotelli, N. J. (2017). Bi-dimensional null model analysis of  
418 presence-absence binary matrices. *Ecology*, *99*, 103-115.
- 419 Triantis, K. A., Guilhaumon, F., & Whittaker, R. J. (2012). The island species–area  
420 relationship: biology and statistics. *Journal of Biogeography*, *39*, 215-231.
- 421 Tuomisto, H. (2010). A diversity of beta diversities: straightening up a concept gone awry.  
422 Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography*,  
423 *33*, 2-22.
- 424 Ulrich, W., Almeida-Neto, M., & Gotelli, N. J. (2009). A consumer's guide to nestedness  
425 analysis. *Oikos*, *118*, 3-17.
- 426 Veech, J. A., & Crist, T. O. (2010). Toward a unified view of diversity partitioning. *Ecology*,  
427 *91*, 1988-1992.
- 428 Veech, J. A., Summerville, K. S., Crist, T. O., & Gering, J. C. (2002). The additive  
429 partitioning of species diversity: recent revival of an old idea. *Oikos*, *99*, 3-9.
- 430 Whittaker, R. H. (1960). Vegetation of the Siskiyou Mountains, Oregon and California.  
431 *Ecological Monographs*, *30*, 279-338.
- 432 Whittaker, R. H. (1965). Dominance and diversity in land plant communities. *Science*, *147*,  
433 250–260.
- 434 Whittaker, R. J., & Fernández-Palacios, J. M. (2007). *Island biogeography: ecology,*  
435 *evolution, and conservation* (2nd ed.). Oxford: Oxford University Press.
- 436 Zhang, Q., Hou, X., Li, F. Y., Niu, J., Zhou, Y., Ding, Y., . . . Kang, S. (2014). Alpha, beta  
437 and gamma diversity differ in response to precipitation in the inner Mongolia grassland.  
438 *PloS ONE*, *9*, e93518.

439

#### 440 **SUPPORTING INFORMATION**

441 Additional supporting information may be found in the online version of this article at ...

442

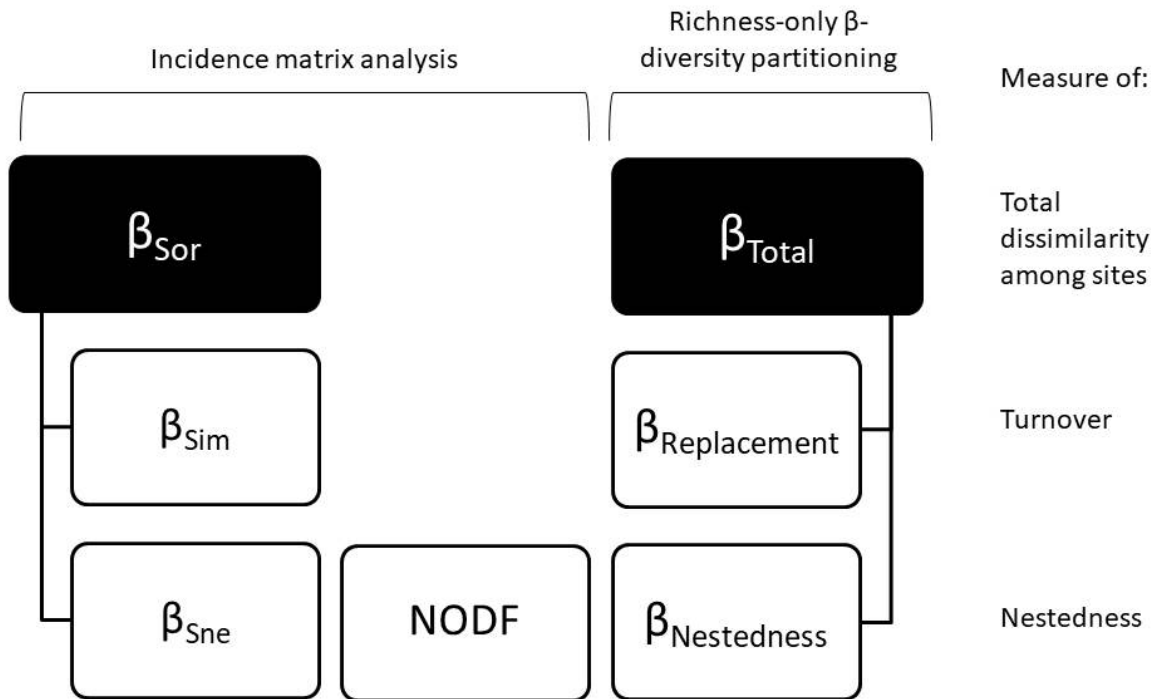
443 **TABLES**

444 **Table 1.** A glossary of the different  $\beta$ -diversity and nestedness metrics used in this study. The  
 445 metrics are split into those that are calculated using the incidence matrix and those that are  
 446 calculated using simple richness data (i.e.  $\gamma$  and  $\alpha$ -diversity).

Metric	Description	References
<b>Incidence Matrix Metrics</b>		
$\beta_{\text{Sor}}$	Sørensen dissimilarity; a $\beta$ -diversity metric that measures compositional differences between sites.	Baselga (2010, 2012)
$\beta_{\text{Sim}}$	Simpson dissimilarity; the turnover component of Sørensen dissimilarity.	Baselga (2010, 2012)
$\beta_{\text{Sne}}$	Nestedness-resultant fraction of Sørensen dissimilarity.	Baselga (2010, 2012)
NODF	A nestedness index based on the twin properties of standardized differences in matrix row and column fills and paired overlap.	Almeida-Neto et al. (2008)
<b>Richness-Only Metrics</b>		
$\beta_{\text{Total}}$	$\beta$ -diversity calculated using additive diversity partitioning; the amount that regional diversity exceeds the mean diversity of a set of sites.	MacArthur et al. (1966), Lande (1996), Veech et al. (2002)
$\beta_{\text{Stan}}$ $\beta_{\text{Nestedness}}$	$\beta_{\text{Total}}$ normalised by $\gamma$ and the number of sites. Hitherto interpreted as a measure of nestedness. Its calculation is almost identical to the $\beta_{\text{Area}}$ metric of Crist and Veech (2006) but uses the maximum observed richness of an island in the dataset rather than maximum richness predicted by the power law SAR model.	Chao et al. (2012) Chiarucci et al. (2010), Cabral et al. (2014), Zhang et al. (2014)
$\beta_{\text{Replacement}}$	The turnover component of $\beta_{\text{Total}}$ . Hitherto interpreted as a measure of the compositional differences across a set of sites. Similar in calculation to the $\beta_{\text{Replace}}$ of Crist and Veech (2006).	Chiarucci et al. (2010), Cabral et al. (2014), Zhang et al. (2014)
$\beta_{\text{Area}}$	Measures the portion of $\beta_{\text{Total}}$ that is due to area effects.	Crist & Veech (2006)
$\beta_{\text{Replace}}$	The portion of $\beta_{\text{Total}}$ that is due to factors other than area.	Crist & Veech (2006)
$\beta_{\text{Mult}}$	$\beta$ -diversity calculated using multiplicative diversity partitioning; the regional-to-local diversity ratio (true $\beta$ -diversity).	Whittaker (1960, 1965), Jost (2007), Tuomisto (2010)

447

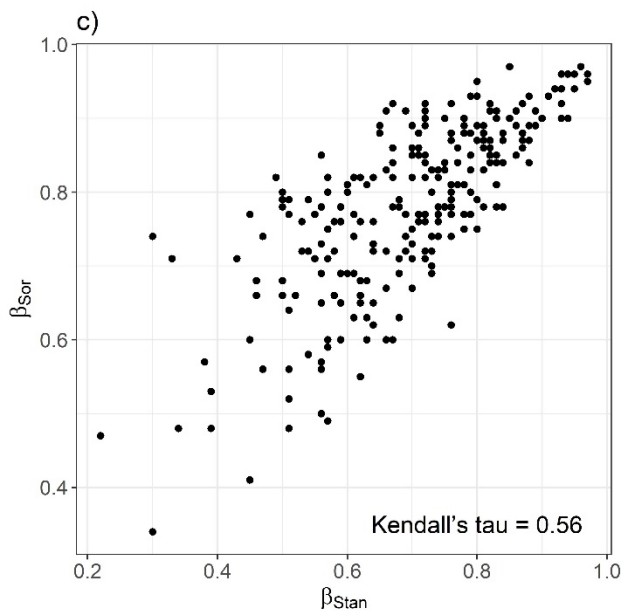
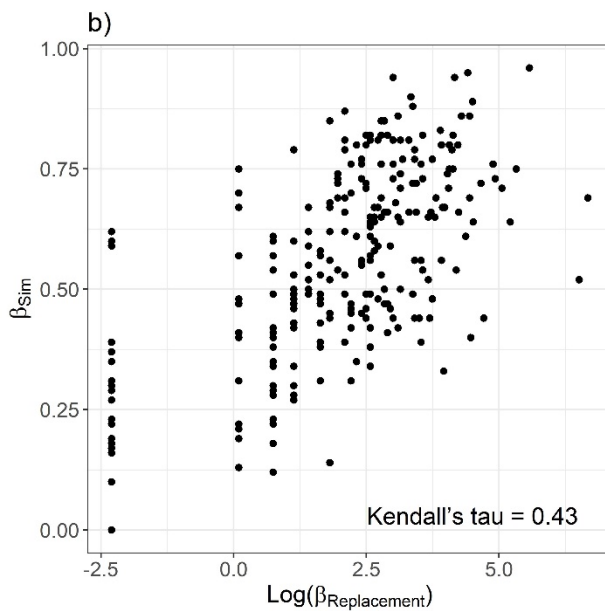
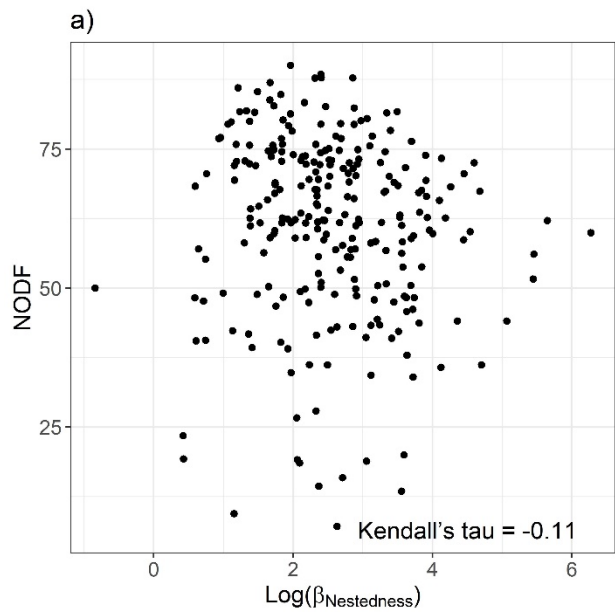
448



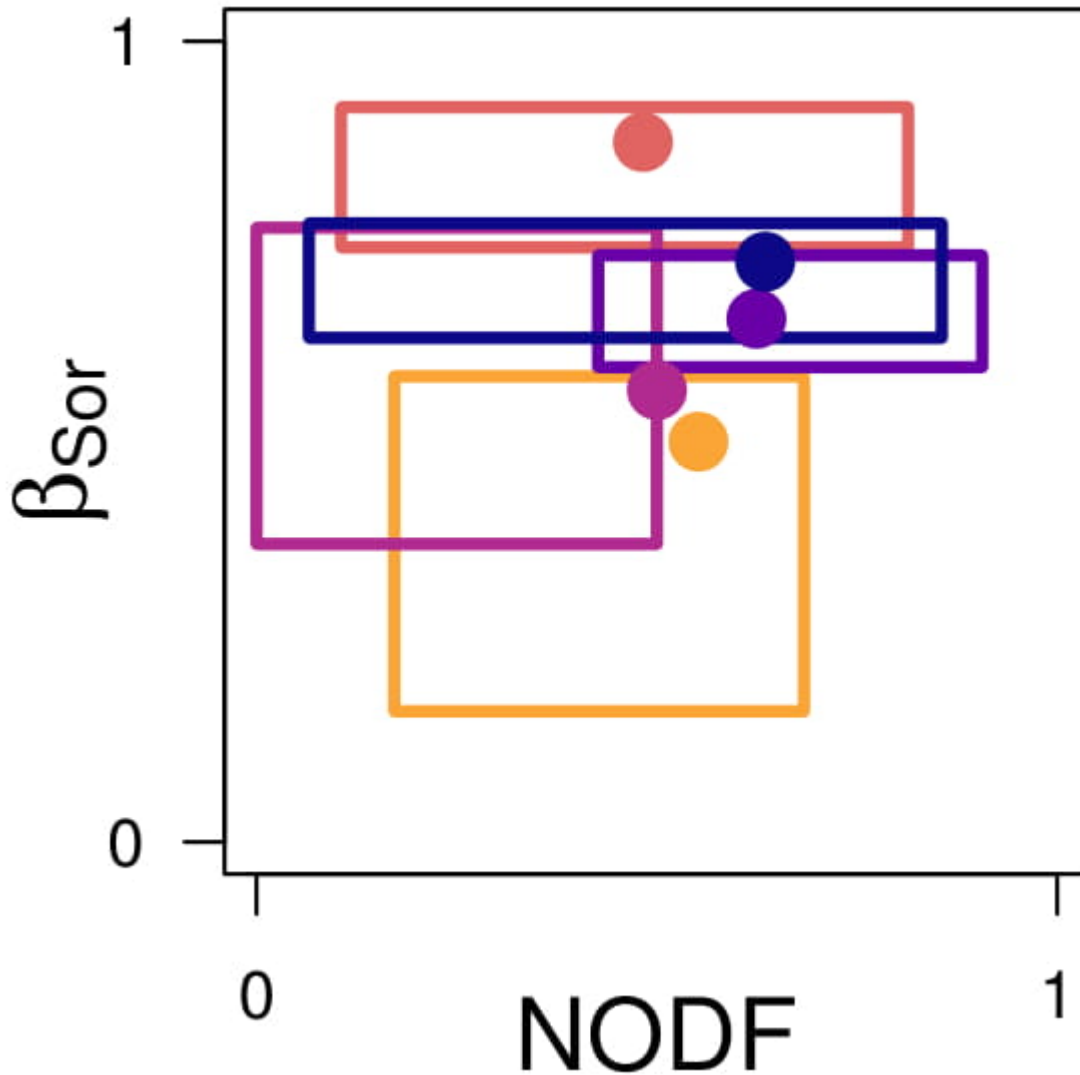
451

452 **Figure 1.** Overview of the various total  $\beta$ -diversity, turnover and nestedness metrics  
 453 discussed in the main text. The metrics are organised within the figure according to how they  
 454 are calculated (richness-only  $\beta$ -diversity partitioning and incidence matrix analysis), and to  
 455 what it is they are purported to be measuring (total  $\beta$ -diversity, turnover or nestedness). The  
 456 two solid black boxes indicate metrics that have been partitioned into two components  
 457 (connected by the black lines). Here, richness-only  $\beta$ -diversity partitioning relates to the  
 458 partitioning of additive  $\beta$ -diversity. For definitions of terms, see Table 1.

459



**Figure 2.** Scatter plots showing the relationship between different  $\beta$ -diversity and nestedness metrics, calculated using richness-only  $\beta$ -diversity partitioning and through analysing the full incidence matrix. The data are 254 empirical incidence matrices from different island systems (e.g. oceanic islands, habitat islands). The metrics on the x-axis in all three plots are those calculated using richness-only  $\beta$ -diversity partitioning, whereas those on the y-axis were calculated using the full incidence matrix. For the definitions of the variables, see Table 1.  $\beta_{\text{Nestedness}}$  and  $\beta_{\text{Replacement}}$  have been logged (base-e; a constant of 0.1 was added to all values to avoid zero values) for presentation purposes.



479

480 **Figure 3.** The range of possible NODF and  $\beta_{Sor}$  values for five incidence matrices, given the  
 481 constraints of a dataset's  $\gamma$ -diversity,  $\alpha$ -diversity and number of islands. For a given matrix  
 482 the coloured rectangle represents the range of possible values, and the same coloured dot  
 483 represents the observed values for that matrix. In each case, the procedure consisted of  
 484 attempting a large number of random swaps (10, 000) between species presences / absences  
 485 within sites, retaining only those swaps that moved the matrix structure in the desired  
 486 direction (i.e. towards the extremes of possible NODF and  $\beta_{Sor}$  values). Five incidence  
 487 matrices were randomly selected from the full set of 254.

488