

1 ECOLOGICAL INFORMATICS (ISSN: 1574-9541) 32: pp. 63-68. (2016)

2

3 **Once again on the components of pairwise beta diversity**

4

5 Running title: On components of pairwise beta diversity

6 Word count: ca. 5400

7

8 **Authors:**

9 **János Podani**

10 Department of Plant Systematics, Ecology and Theoretical Biology, Institute of

11 Biology, L. Eötvös University, Pázmány P. s. 1/c, H-1117, Budapest, Hungary

12 and

13 Ecology Research Group of the Hungarian Academy of Sciences, Budapest, Hungary

14 Corresponding author. Email: podani@ludens.elte.hu

15

16 and

17

18 **Dénes Schmera**

19 MTA Centre for Ecological Research, Tihany, Hungary

20 and

21 Section of Conservation Biology, Department of Environmental Sciences, University

22 of Basel, Basel, Switzerland

23

24

25 **Abstract**

26 Presence-absence based beta diversity defined for pairs of sites may be partitioned
27 into components following two different ways of thinking. Within the framework of
28 Baselga (abbreviated hereafter as BAS), nestedness is crucial and dissimilarity is
29 partitioned into replacement (turnover) and nestedness-resultant fractions. The
30 method proposed by Podani and Schmera (POD), however, places emphasis on the
31 mathematical additivity of components and divides dissimilarity into replacement and
32 richness difference components. A recent comparison by Baselga and Leprieur
33 (2015), on the example of the Jaccard family of indices, emphasizes the independence
34 of replacement component from absolute richness difference and concludes that the
35 replacement function of the BAS framework is the only true measure of species
36 replacement. As a response to this study, we show here that 1) the sacrifice one must
37 make for independence is that the components themselves are scaled differently and
38 are not always comparable ecologically, 2) absolute (raw) replacement and richness
39 difference are not independent, so that independence from the latter cannot be a
40 fundamental criterion that a replacement measure should satisfy, 3) relativization
41 applied in the POD framework is ecologically interpretable, leading to a meaningful
42 conceptualization of species replacement, 4) the BAS and POD methods are linked
43 through a generalized replacement function, 5) both the BAS and POD approaches
44 may produce high correlations with environmental variables, whereas 6) the POD
45 approach offers in many respects more illuminating demonstrations of the underlying
46 changes of pattern than the graphs of Baselga and Leprieur for both artificial and
47 actual fish distribution data.

48 **Keywords:** Data structure; Diversity partitioning; Nestedness; Richness difference;
49 Species replacement; Turnover

50 **1. Introduction**

51 Describing species distribution patterns and underlying mechanisms is at the heart of
52 current biodiversity research. Beta diversity, summarising a wide variety of
53 conceptual and methodological approaches, is a key term for assessing community
54 variation in space and time. Baselga (2010) suggested that beta diversity can be
55 partitioned into spatial turnover (also termed replacement) and nestedness resultant
56 dissimilarity components. This partitioning framework (hereafter abbreviated as BAS,
57 following Legendre 2014) has been considered as a major contribution to this field
58 because the components may reflect the existence of different underlying mechanisms
59 driving beta diversity and thus the methodological framework can efficiently be used
60 for understanding central geographic, ecological and conservation issues.

61 Thanks to its novelty and soundness, this scheme triggered critical comments
62 (Schmera & Podani 2011, Almeida-Neto et al. 2012), an intensive dispute (Baselga
63 2012, Carvalho et al. 2013, Chen 2015) and the development of competing
64 approaches (Podani & Schmera 2011, Carvalho et al. 2012, Cardoso et al. 2014) as
65 well. For instance, according to a recent proposal, beta diversity can be partitioned
66 into replacement and richness difference components (hereafter abbreviated as POD,
67 following Legendre 2014, see also Podani & Schmera 2011, Carvalho et al. 2012,
68 Carvalho et al. 2013). A review by Legendre (2014) compared these frameworks and
69 found that they represent alternative approaches with different conceptual and
70 mathematical backgrounds.

71 Most recently, Baselga & Leprieur (2015) revisited the issue by evaluating the
72 performance of the BAS and POD frameworks under different conditions. They
73 emphasized that (1) the replacement components of the two approaches formalize
74 different concepts, and (2) only the replacement component of the BAS framework is

75 independent of richness difference. Based on their simulations and analyses of actual
 76 data, BAS proved to be the winner, leading to the conclusion that (3) only its
 77 replacement component is able to express “true” replacement. We disagree with the
 78 final conclusion by questioning the basic assumption that the replacement component
 79 should be independent of richness difference and by arguing that the relevant
 80 component of the POD framework expresses replacement at least as faithfully as the
 81 BAS framework. Regarding independence, we refer to Chao et al. (2012) who
 82 clarified the multiple meaning of this term. Here, we mean that knowing one diversity
 83 component one has no information about the value of another, and vice versa. Our
 84 disagreement is also due to the fact that in many instances Baselga & Leprieur (2015)
 85 still misunderstood, confused or neglected some basic aspects of the POD approach.
 86 To avoid the undesirable situation that potential users of either method are misguided,
 87 this rebuttal attempts to clarify the most important points.

88 **2. The BAS and POD frameworks**

89 For the benefit of the reader, we give a brief summary of the two frameworks
 90 following the Legendre’s (2014) system of symbols. Decomposition of beta diversity
 91 for a pair of sites, expressed by the Sorensen dissimilarity index (D_S) was first
 92 suggested by Baselga (2010):

93

$$94 \quad D_S = \frac{b+c}{2a+b+c} = \text{Repl}_{BS} + \text{NeS}_{BS} = \frac{\min(b,c)}{a+\min(b,c)} + \frac{a}{a+\min(b,c)} * \frac{|b-c|}{2a+b+c} \quad (1)$$

95

96 in which a refers to the number of species present in both sites, while b and c
 97 correspond to the number of species present only in the first and the second,
 98 respectively, such that $n=a+b+c$. The above equation implies that Simpson
 99 dissimilarity (Repl_{BS}), expressing spatial turnover is subtracted from Sorensen

100 dissimilarity to obtain the so-called nestedness resultant dissimilarity (Nes_{BS}). It has
 101 been suggested that Simpson dissimilarity reflects the amount of species replacement,
 102 whereas the remainder is “due to nestedness”. Later, the Jaccard dissimilarity index
 103 was decomposed in an analogous way (Baselga 2012):

104

$$105 \quad D_J = \frac{b+c}{a+b+c} = Repl_{BJ} + Nes_{BJ} = \frac{2 \min(b,c)}{a+2 \min(b,c)} + \frac{a}{a+2 \min(b,c)} * \frac{|b-c|}{a+b+c} \quad (2)$$

106

107 The POD framework consists of an absolute (or raw) and a relativized decomposition
 108 of beta diversity (Podani and Schmera 2011). In the absolute decomposition, Weiher -
 109 Boylen beta diversity (β_{WB} , see also Koleff et al 2003) is divided into (raw)
 110 replacement and (raw) richness difference components:

111

$$112 \quad \beta_{WB} = b + c = \text{replacement} + \text{richness difference} = 2\min(b,c) + |b-c|$$

113

114 The basis of the relativized POD framework (Podani and Schmera 2011) is the
 115 decomposition of pairwise beta expressed by the Jaccard dissimilarity (D_J) into two
 116 additive fractions according to

117

$$118 \quad D_J = \frac{b+c}{a+b+c} = Repl_J + Rich_J = \frac{2 \min(b,c)}{a+b+c} + \frac{|b-c|}{a+b+c} \quad (3)$$

119

120 The first part was termed the relativized species replacement ($Repl_J$), while the second
 121 was relativized richness difference ($Rich_J$). *It is important to point out that we never*
 122 *considered these components as stand-alone dissimilarity functions, only fractions of*
 123 *dissimilarity!* According to our proposal, these two fractions, together with the

124 complement of D_J , i.e., Jaccard similarity (S_J) may be used as scores to specify the
125 position of the site pair in a 2D simplex diagram. If the same operation is done for all
126 pairs of sites in a region, then we have a tool for visualizing structure of the entire
127 presence/absence matrix (Podani and Schmera 2011). Sorensen index may also be
128 decomposed in similar manner (Carvalho et al. 2012, Legendre 2014), which is not
129 detailed here because 1) the only difference from Eq. 3 is that $2a$ replaces a in the
130 denominators and 2) this paper is concerned only with the Jaccard family of indices,
131 in response to Baselga & Leprieur (2015) whose study is restricted to this group.

132 *The relativized replacement component in POD*

133 Before we discuss the relative merits and disadvantages of the two frameworks, a
134 terminological clarification is in order. This is important because Baselga and
135 Leprieur erroneously equate our relativized species replacement function ($Repl_J$, in
136 Eq. 3) with Williams' replacement index (see $\beta_{.3}$ in their Table 1). The formula
137 proposed by Williams (1996) reads:

138

$$139 \quad \beta_{.3} = 1 - \frac{a + \min(b, c) + |b - c|}{a + b + c} = \frac{\min(b, c)}{a + b + c}, \quad (4)$$

140

141 that is, $2\beta_{.3} = Repl_J$, a fact that cannot be neglected when our objective is
142 decomposition into additive fractions. Multiplication by 2 was first suggested by
143 Cardoso et al. (2009) to ensure that the index changes between 0 and 1. Zero obtains
144 in the situation when there are no unique species for either site or both so that there is
145 no replacement. The maximum value results if the species set of one site is
146 completely replaced by the species of the other, which is possible only if $b=c>0$ and
147 $a=0$.

148 Cardoso et al. (2009) used the same abbreviation (β_3) for the multiplied index,
149 whereas Podani & Schmera (2011) suggested $\frac{2 \min(b,c)}{a+b+c}$ as a relativized replacement
150 index without referring to the suggestions of Cardoso et al. (2009). Thus, it is partly
151 our fault that the multiplied version and the original Williams index are still confused.
152 However, since multiplication modifies the range of the measure, β_3 and Repl_J should
153 be treated as different functions. This is in agreement with Koleff et al. (2003) who
154 list several examples of beta diversity measures which differ only in a multiplying
155 factor of 2.

156 **3. Commensurability**

157 Decomposition of absolute or raw beta diversity has been rarely used in practice,
158 because the results from different surveys are hardly commensurable. Relativization
159 (or standardization) of raw values is a general strategy, as with dissimilarity
160 coefficients whose values range between 0 and 1. When dissimilarities themselves are
161 partitioned, however, another aspect of commensurability enters the scene. If we
162 examine the suggested components of D_J in Eq. 2 we find that they disagree in the
163 manner they are standardized, as demonstrated by the following example with $a=3$,
164 $b=5$ and $c=1$. $D_J (=6/9)$ is the ratio of the number of unique species and the total
165 number of species in the two sites being compared, and $\text{Repl}_{BJ} (=2/5)$ may be
166 conceived as the ratio of the number of replaced species and the total number of
167 species if both sites *were* equally poor (i.e. with $b=c=1$). The third one, Nes_{BJ}
168 $(=12/45)$ lacks any ecologically meaningful interpretation in terms of species numbers
169 or raw data values. Apparently, due to different relativization and weighting
170 operations applied to the component terms, there is no “common currency” for the
171 three terms involved. This is not so with the POD framework: D_J and both of its
172 fractions in Eq. 3 are interpretable as ratios, since $\text{Repl}_J = 2/9$ and $\text{Rich}_J = 4/9$. These

173 are expressed on the same scale as D_J itself. We think this is important even though
174 the dissimilarities themselves are dimensionless quantities. A conclusion is that
175 whereas the BAS approach is inconsistent, the POD framework is consistent in the
176 way component terms are measured. This holds true for the partitioning of Sorensen
177 index as well (see Schmera & Podani 2011). A natural question arises: what are then
178 the advantages of using inconsistent relativization in the BAS decomposition? We
179 look for the answer in the next section.

180 **4. Replacement and its independence from richness difference**

181 Baselga and Leprieur (2015) emphasize the main result of their analyses as follows:
182 the replacement component in the BAS framework is *independent* of richness
183 difference, whereas the analogous component in the POD approach is not. To prove
184 this, they apply a huge arsenal of statistical tools (comparison of parallel components,
185 random parameter method, the evaluation of joint probabilities etc., see their Figure
186 2). The above conclusion is straightforward and no detailed tests are required,
187 however, if we consider how Repl_{BJ} (and Repl_{BS}) are formulated. The denominator of
188 Repl_{BJ} , $a+2\min(b,c)$, contains information on the number of species in the poorer site
189 whereas the number of species in the richer site ($a+\max(b,c)$) does not appear in the
190 formula at all. Therefore, the result is, by definition, unaffected by the richer site and
191 in turn richness difference either (see also Appendix 1). Obviously, Simpson
192 dissimilarity (Repl_{BS}) and Baselga's Repl_{BJ} were constructed *on purpose* to satisfy
193 this external requirement. Richness difference as such is not part of their
194 decomposition procedure.

195 Baselga and Leprieur (2015) argued that the use of the term replacement for the two
196 respective components originating from different frameworks might be misleading
197 and suggested that the term replacement should be reserved to indices that are

198 independent of richness difference. While we agree that terminological clarity is
199 essential, we disagree with the suggestion itself. Here we show that independence
200 from absolute richness difference cannot be a criterion of a replacement measure. In
201 fact, replacement "implies the replacement of some species by others" (p. 135,
202 Baselga 2010) and we emphasized several times (see e.g., p. 155 in Schmera and
203 Podani 2011) that species replacement originally refers to an *absolute* measure, i.e., to
204 the number of species that are replaced between two sites ($2\min(b,c)$). This is not
205 independent mathematically from raw richness difference ($|b-c|$) as proved in
206 Appendix 1 and graphically demonstrated also by two simulations in the
207 Supplementary Material using random parameter approaches similar to those in
208 Baselga and Leprieur's randomization experiments with relativized measures. For
209 actual ecological data, these two components may indeed be uncorrelated, yet large
210 richness differences rarely associate with large replacement due to the constraining
211 effect of the maximum number of species (see Supplementary Material). Therefore, if
212 the above suggestion were accepted, then even raw replacement could not be called
213 "replacement" – which is conflicting with any intuitive views on species turnover. A
214 potential source of this contradiction is that decomposition of absolute (raw) beta
215 diversity is not included originally in the BAS framework. In sum, *independence*
216 *cannot be a fundamental criterion that a replacement measure should satisfy and thus*
217 *the replacement component of the BAS framework cannot be regarded as the sole*
218 *candidate for the use of the term replacement.*

219 **5. How to relativize?**

220 Coefficients Repl_{BS} and Repl_{BJ} are asymmetric in handling b and c , as shown by the
221 following example with $a = 1$, $b = 49$ and $c = 1$. Now, $\text{Repl}_{\text{BJ}} = 2/3$ suggesting high
222 level of turnover, which is fine with respect to the poorer site. It has only two species,

223 and the loss of one species and the gain of another is indeed a significant change. As
224 Carvalho et al. (2013) pointed out, however, “we do not see any reason why only the
225 poorer site should establish the theoretical upper limit of species replacement.” Thus,
226 from the “viewpoint” of the richer site, the same change is practically nothing if we
227 consider that it has as much as 50 species in total. One alternative is therefore a new
228 replacement function in the Jaccard index-family, which considers the richer site for
229 relativization (Eq. A1, in Appendix 2). This function takes the value of 0.02 – which
230 appears a more reasonable score than $2/3$.

231 In these functions, one of the sites plays little or no role in determining the index
232 value – an undesirable property for a dissimilarity function. The relativized
233 replacement index $Repl_J$ does provide a balanced result with respect to the total
234 species richness of the two sites. For the small example above, the value of $2/51 \sim$
235 0.04 indicates fairly well that only a negligible minority of potential species
236 participated in the replacement process.

237 Relativization by $a+b+c$ has the obvious consequence that $Repl_J$ will not be
238 independent of richness difference (Appendix 1). It is implicit in the denominator
239 (i.e., $n = a + 2\min(b,c) + |b-c|$) therefore, if n increases because $|b-c|$ increases then
240 the value of the index will decrease. If $|b-c|$ increases such that n remains constant,
241 then $\min(b,c)$ must decrease after a reached zero (and may decrease even before), so
242 that the value of the index also decreases. In the POD approach this is fundamental: S_J
243 and the two components of D_J mutually constrain one another (their sum is always 1),
244 these three are expressed on the same scale to provide a meaningful 2D display in
245 form of a simplex plot, and the inclusion of all site pairs in the same simplex plot
246 would be impossible without relativization. Notably, such kind of relativization is
247 commonplace in numerical ecology – often leading to loss of independence. For

248 example, if raw data are standardized by species totals, the scores, which were
249 originally independent, will now be dependent: each new value will be affected by all
250 the other scores for the given species. Appendix 2 shows that the two asymmetric
251 replacement functions represent extreme cases of a general replacement formula in
252 which $Repl_j$ takes an intermediate position.

253 **6. Correlations with environmental variables**

254 Although we disagree with the view implicit in Baselga and Leprieur (2015)
255 suggesting that the higher the correlation between an index value and an
256 environmental variable, the better the index, we planned to repeat their study of actual
257 data (North American fish dataset) with more measures involved, including raw
258 replacement and richness difference. This would clarify the question whether it is
259 worth switching from raw measures to relativized ones. However, the supplementary
260 material did not include the geographical distances in terms of the number of drainage
261 basin divides between river basins Baselga and Leprieur (2015) were using.
262 Furthermore, glacial history was presented as an ordinal variable to which, according
263 to our view (Podani 2005), correlation measures cannot be applied (but see Ricotta
264 and Feoli 2013, for alternative views on the usefulness of ordinal data). Thus, we
265 computed correlations between the matrices of eight components in the BAS and
266 POD approaches and only four environmental variables, two of them being PCA
267 summaries of 17 climatic variables (Table 1).

268

269 **Table 1.** Pearson product moment matrix correlations between various coefficients
270 and components of beta diversity and four variables reflecting inter-site distances in
271 terms of climate (PCA1 and PCA2) plus altitude and area for the North American fish
272 distribution data.

Measure	PCA1	PCA2	Altitude	Area
$ b-c $	0.155	-0.039	-0.042	0.475
$2\min(b,c)$	-0.082	-0.041	0.038	0.138
D_J	0.372	0.258	0.191	0.018
$Rich_J$	0.246	0.020	-0.018	0.102
$Repl_J$	-0.061	0.119	0.123	-0.099
$Repl_{BJ}$	0.367	0.284	0.216	-0.126
Nes_{BJ}	-0.248	-0.213	-0.167	0.191
New function (A1)	0.344	0.235	0.172	0.293

273

274 It is seen from the table that the BAS framework does not have the privilege to
275 produce relatively high correlations. The largest value (0.475) was obtained between
276 absolute richness difference and area difference – which agrees well with basic
277 biogeographical “rules”. The effect of area was less conspicuous for other
278 components, with the new turnover function (Eq. A1) as an exception ($r=0.293$).
279 Regarding the BAS and POD frameworks, Baselga’s replacement function produced
280 higher correlations with climate than POD. For example, with the first climatic
281 component both the absolute and relativized turnover ($Repl_J$) were slightly negatively
282 correlated, while $Repl_{BJ}$ and function A1, as well as D_J itself had positive correlations
283 larger than 0.3. Note the extremely high negative correlations – not shown in the table
284 - between components within the BAS approach (it is -0.873 for $Repl_{BJ}$ and Nes_{BJ}),
285 and the POD framework as well (-0.864 between $Rich_J$ and $Repl_J$). These are
286 empirical indications of interdependence between these components. These results
287 represent only one case study with four variables, and more actual cases should be
288 examined before any far-reaching conclusions can be made about the relationship
289 between diversity components and the environment.

290 **7. Monotonicity**

291 Baselga and Leprieur (2015) examined four different scenarios in which an initial
 292 configuration was gradually and systematically modified to see if beta diversity
 293 components respond monotonically to these changes. The results were inconclusive:
 294 components of the BAS and POD differed in performance. We agree with Baselga
 295 and Leprieur (2015) that gradual changes of data pattern need not be reflected by
 296 monotonic responses under all circumstances. There is one comment only:
 297 monotonicity should be distinguished from strict monotonicity, so the POD
 298 components were monotonous, while not always strictly monotonic, over the spatial
 299 sequence in cases of the extinction and homogenization scenarios.

300 **8. Advantages of the POD approach**

301 Unfortunately, while focusing on independence and monotonicity issues Baselga and
 302 Leprieur (2015) are silent about additional and important features of the POD
 303 framework. In this, decomposition of Jaccard dissimilarity is only part of a more
 304 general analytical toolkit: S_J and the two components of D_J are not restricted to
 305 express only a single contrast, the one between similarity and beta diversity along a
 306 1D simplex. There are other two possibilities for combining S_J with $Repl_J$ or $Rich_J$.
 307 The sum of the first two terms gives the relative richness agreement function:

308

$$309 \quad Agr_J = S_J + Repl_J = \frac{a}{a+b+c} + \frac{2 \min(b,c)}{a+b+c} \quad (5)$$

310

311 which represents a contrast with $Rich_J$ along another 1D simplex. Perhaps more
 312 important is the other combination

313

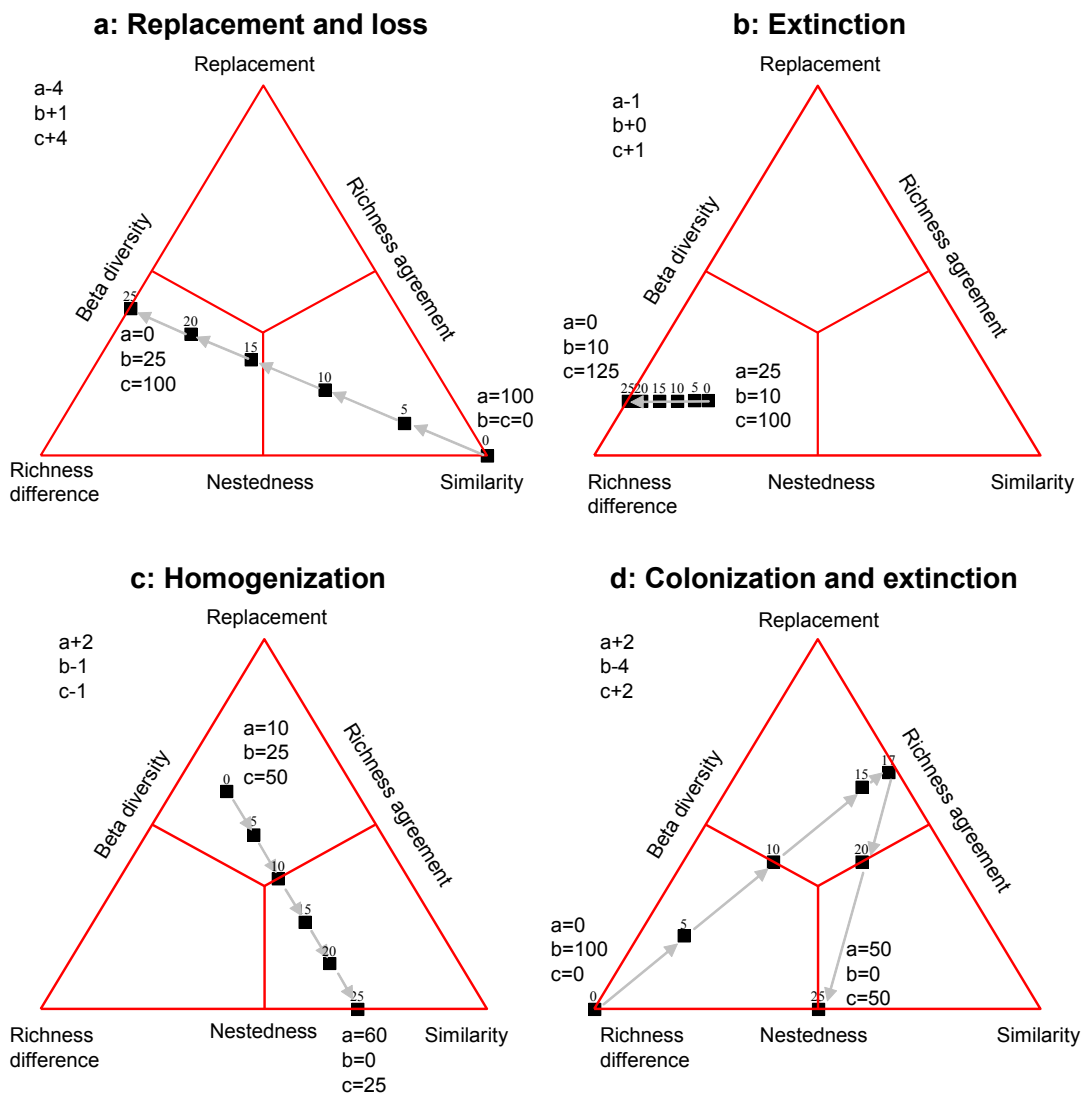
$$314 \quad Nes_J = S_J + Rich_J = \frac{a}{a+b+c} + \frac{|b-c|}{a+b+c} \text{ if } a > 0, \text{ otherwise } Nes_J = 0 \quad (6)$$

315

316 called the relativized nestedness index (Podani and Schmera 2011, 2012, Schmera and
317 Podani 2011). Baselga and Leprieur (2015, p. 1076 right column, top) refer to the left
318 part (i.e. without the $a > 0$ condition) as the “similarity index” counterpart of β_{-3} . The
319 problem is not merely that β_{-3} differs from the Repl_J function – as already said.
320 Equation 6 is *not similarity in the conventional sense*: it is a sum of similarity (S_J) and
321 a fraction of dissimilarity (Rich_J), both contributing with equal weight to pairwise
322 nestedness. Our nestedness concept is clear from this equation: it is determined by the
323 overlap between the species lists of the two sites and by how deeply the list of the
324 poorer site is “embedded” into the species list of the richer site, provided that they
325 have at least one species in common (Podani and Schmera 2012). In other words,
326 nestedness is neither similarity nor dissimilarity, and there is a condition (positive a)
327 involved in its calculation. The contrast between replacement and nestedness is thus
328 expressed on another 1D simplex, clearly showing that – in addition to the
329 disappearance of shared species – species replacement is the only pattern antagonistic
330 to nestedness.

331 Now we can examine the behaviour of the four artificial scenarios examined by
332 Baselga and Leprieur (2015) through our *entire* simplex approach. The ternary plots
333 of Figure 1 utilize the same information as Figure 4 of the above authors, yet these are
334 in many respects more illuminating demonstrations of the underlying changes of
335 pattern than the graphs of Baselga and Leprieur (2015). The “replacement and loss”
336 series starts with complete self similarity (right corner, Fig. 1a) and losses and
337 replacements direct the process into maximum beta diversity with richness difference
338 dominating over species replacement (i.e., the point representing the last step lies on
339 the beta diversity side, closer to the richness difference corner). The “extinction”

340 scenario involves a relatively short trajectory within the plot: the low initial similarity
341 further decreases down to zero while richness difference increases (Fig. 1b). This
342 sequence also ends up with maximum beta diversity, with richness difference even
343 more dominating than in the previous example. The trajectory is near and parallel to
344 the nestedness side, reflecting that species replacement is low and constant all the
345 way. The “homogenization” scenario starts from a position mostly determined by high
346 species replacement (Fig. 1c). The series involves a steady increase in the number of
347 shared species, and the continuous disappearance of unique species. At the end, the
348 species set of one site becomes completely nested within the other: the point
349 representing this site pair falls onto the nestedness side of the triangle. It is closer to
350 the similarity corner, because the number of shared species (i.e. overlap) is higher
351 than the number of species exclusively occurring in the richer site. The “colonization
352 and extinction” scenario begins with a species rich and an empty site, therefore the
353 first point falls right onto the richness difference corner (Fig. 1d). During the
354 colonization process species migrate to the empty site, half of them eventually
355 disappearing from the first one. This involves increasing similarity and species
356 replacement, so that the two sites converge in species richness. In step 17, they are
357 very close to the richness agreement side of the plot (for a total of 99, rather than 100
358 species full coincidence would have been possible). After this point, however, the
359 initially rich site suffers further losses by which it becomes increasingly nested within
360 the “daughter” site. The series finishes at the nestedness side of the plot, with
361 similarity and richness difference equally contributing to the perfect nestedness of the
362 last pair.

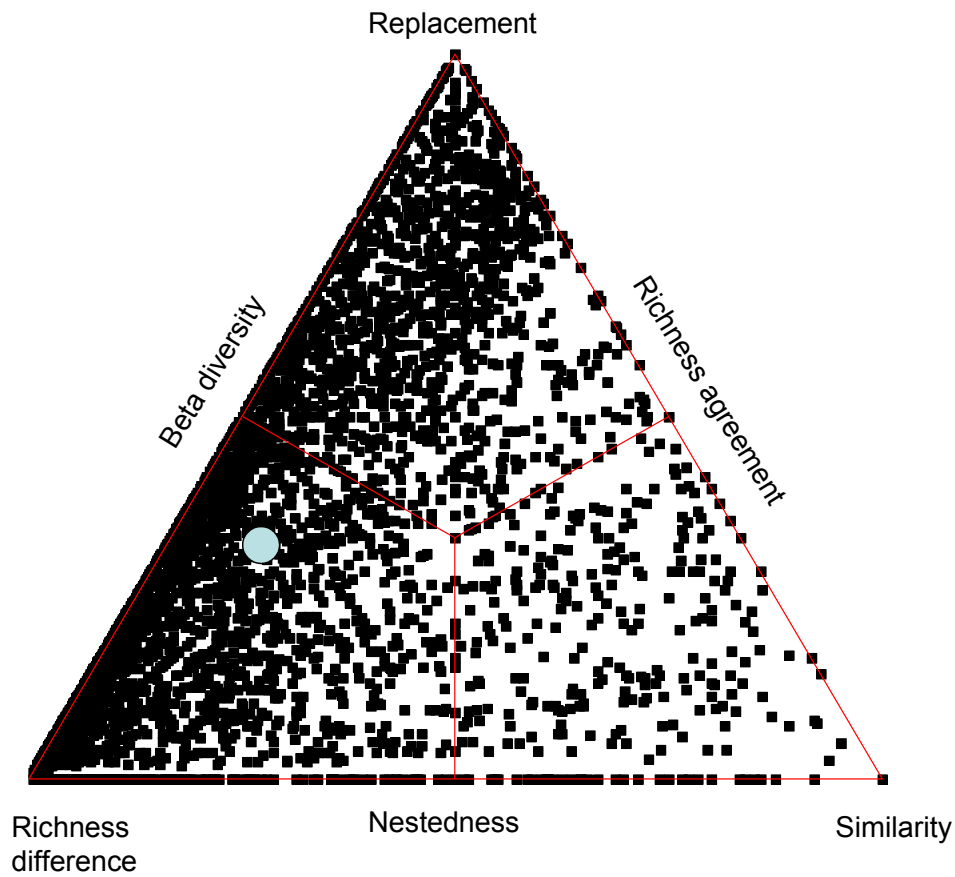


363

364

365 **Fig. 1.** Simplex diagrams illustrating the four 25-step artificial scenarios examined by
 366 Baselga and Leprieur (2015). Points labeled by 0 are the initial situations, and then
 367 every fifth stage is shown, except for **d** to which an extreme position is also added to
 368 show that the trajectory does not reach the richness agreement edge. Arrows illustrate
 369 direction of the processes. Changes involved in each step are shown in top left of each
 370 diagram, the values of *a*, *b* and *c* are shown for steps 0 and 25 only.

371 For comparison, we also calculated the 2D simplex for the North American fish data
372 as well. Matrix fill in the 171 sites by 549 species data matrix was very low (7.2%)
373 which is by itself an early indication of high beta diversity. The ternary plot (Fig. 2)
374 confirms this expectation: most of the site pairs are concentrated near the beta
375 diversity edge. Percentage beta diversity is as high as 92.8%, with species
376 replacement contributing with 32.1% and richness difference with 60.7%. Overall
377 similarity is only 7.2%. Therefore, the centroid of the point cloud falls close to the left
378 edge, nearer to the richness difference than the replacement corner. The anti-
379 nestedness fraction, i.e., the contribution of points which lie exactly on the beta
380 diversity (left) edge is 40.7% showing that a very large number of site pairs have no
381 species in common at all. We feel that these results are important for a deeper
382 understanding of pattern in the North American fish data: while the BAS approach
383 deliberately ignores richness differences, the POD framework clearly shows that
384 richness difference is the major factor governing beta diversity.



385

386 **Fig. 2.** Simplex diagram for the North American fish data. A circle indicates the
 387 centroid of the point cloud.

388

389 **Acknowledgements**

390 This research was supported by the Hungarian Scientific Research Fund (OTKA
 391 104279).

392

393

394

395 **References**

- 396 Almeida-Neto, M., Frensel, D.M.B. & Ulrich, W. (2012) Rethinking the relationship
397 between nestedness and beta diversity: a comment on Baselga (2010). *Global*
398 *Ecology and Biogeography*, 21, 727-777.
- 399 Baselga, A. (2010) Partitioning the turnover and nestedness components of beta
400 diversity. *Global Ecology and Biogeography*, 19, 134-143.
- 401 Baselga, A. (2012) The relationship between species replacement, dissimilarity
402 derived from nestedness, and nestedness. *Global Ecology and Biogeography*,
403 21, 1223-1232.
- 404 Baselga, A. & Leprieur, F. (2015) Comparing methods to separate components of beta
405 diversity. *Methods in Ecology and Evolution*, 6, 1069-1079.
- 406 Cardoso, P., Borges, P. A. V. and J. A. Veech. (2009). Testing the performance of
407 beta diversity measures based on incidence data: the robustness to
408 undersampling. *Diversity and Distributions*, 15, 1081-1090.
- 409 Cardoso, P., Rigal, F., Carvalho, J.C., Fortelius, M., Borges, P.A.V., Podani, J. and
410 Schmera, D. (2014) Partitioning taxon, phylogenetic and functional beta
411 diversity into replacement and richness difference components. *Journal of*
412 *Biogeography*, 41, 749-761.
- 413 Carvalho, J.P., Cardoso, P. & Gomes, P. (2012) Determining the relative roles of
414 species replacement and species richness differences in generating beta-
415 diversity patterns. *Global Ecology and Biogeography*, 21, 760-771.
- 416 Carvalho, J.C., Cardoso, P., Borges, P.A.V., Schmera, D. & Podani, J. (2013)
417 Measuring fractions of beta diversity and their relationships to nestedness: a
418 theoretical and empirical comparison of novel approaches. *Oikos*, 122, 825-834.

419 Chao, A., Chiu, C.H. and Hsieh, T.C. (2012) Proposing a resolution to debates on
420 diversity partitioning. *Ecology*, 93, 2037-51.

421 Chen, Y. 2015. Independence of total species richness and richness difference are not
422 the same thing. *Ecological Informatics*, 30, 119-123.

423 Gaston, K.J. & Blackburn, T.M. (2000) *Pattern and Process in Macroecology*.
424 Blackwell,, Oxford.

425 Koleff, P., Gaston, K.J. & Lennon, J.J. (2003) Measuring beta diversity for presence-
426 absence data. *Journal of Animal Ecology*, 72, 367–382.

427 Legendre, P. (2014) Interpreting the replacement and richness difference components
428 of beta diversity. *Global Ecology and Biogeography*, 23, 1324–1334.

429 Podani, J. (2000) *Introduction to the Exploration of Multivariate Biological Data*.
430 Backhuys, Leiden.

431 Podani, J. (2005) Multivariate exploratory analysis of ordinal data in ecology: pitfalls,
432 problems and solutions. *Journal of Vegetation Science*, 15, 497-510.

433 Podani, J. & Schmera, D. (2011) A new conceptual and methodological framework
434 for exploring and explaining pattern in presence-absence data. *Oikos*, 120,
435 1625-1638.

436 Podani, J. & Schmera, D. (2012) A comparative evaluation of pairwise nestedness
437 measures. *Ecography*, 35, 889-900.

438 Ricotta, C. and Feoli, E. (2013). Does ordinal cover estimation offer reliable quality
439 data structures in vegetation ecological studies? *Folia Geobotanica*, 48, 437-
440 447.

441 Schmera, D. & Podani, J. (2011) Comments on separating components of beta
442 diversity. *Community Ecology*, 12, 153-160.

443 Ulrich, W., Almeida-Neto, M. & Gotelli, N.J. (2008) A consumer's guide to
444 nestedness analysis. *Oikos*, 118, 3-17.

445 Williams, P.H. (1996) Mapping variations in the strength and breadth of
446 biogeographic transition zones using species turnover. *Proceedings of the Royal*
447 *Society London, B* 263, 579-588.

448

449

450 **Appendix 1 - Independence between components of beta diversity**

451 We shall set $n = a + b + c$ as in the main text, while – without loss of generality – we
452 assume here that $b \leq c$. Except the first case, all functions come from the Jaccard
453 family of indices; proofs for the Sorensen family follow a similar logic. Here we
454 adopt the *fixed parameter approach* (Chao et al. 2012, p. 2040) which checks whether
455 one has information about the value of one component provided that the other
456 component is known. Results of the random parameter approach (Chao et al. 2012, p.
457 2040) and illustrations based on actual data are given in the Supplementary Material.

458 *Raw richness difference and raw species replacement*

459 These two components of Weiher-Boyle beta diversity are not independent. Since c
460 is the sum of b and raw richness difference ($D_r = c - b$), we have

461
$$D_r = n - a - 2b,$$

462 so that for fixed values of n and a , richness difference and replacement mutually
463 constrain each other. To put it differently: raw richness difference can only increase
464 on account of raw species replacement.

465

466 *BAS replacement function and raw richness difference*

467 The function $\text{Repl}_{\text{BJ}} = b/(a + b)$ is obviously independent of raw richness difference.
468 For any value of a and b , the same result is obtained no matter what values n and c ,
469 and therefore D_r take, since these parameters do not appear in the formula.

470

471 *POD replacement functions and raw richness difference*

472 The function $\text{Repl}_{\text{J}} = 2b/n$ is not independent of raw richness difference. After
473 rewriting we have $\text{Repl}_{\text{J}} = 2b/(a + 2b + D_r)$ so that for a given value Repl_{J} and a , $2b$
474 and D_r mutually constrain each other. If replacement is expressed in relation to the

475 richer site, i.e., by $\text{Repl}_{PJ} = 2b/(a + c) = 2b/(a + b + D_r)$ (see formula A2 in
476 Appendix 2), it is easy to see by the same reasoning that dependence holds true. In
477 fact, for fixed n and a , the remaining b and c mutually constrain each other in any
478 expression incorporating these parameters.

479

480 **Appendix 2 – Generalized replacement functions**

481 For the Jaccard index family, the new formula for calculating pairwise turnover with
482 respect to the richer site is given by

483

$$484 \text{Repl}_{PJ} = \frac{2 \min(b, c)}{a + 2 \max(b, c)} \quad (\text{A1})$$

485

486 This is an extreme case (with $\alpha=0$) of the following general formula in which the
487 minimum and the maximum of b and c are weighted by a scale factor α with a range
488 of $[0, 1]$:

489

$$490 \text{Repl}_{\alpha J} = \frac{2 \min(b, c)}{a + 2[\alpha \min(b, c) + (1 - \alpha) \max(b, c)]} \quad (\text{A2})$$

491

492 For $\alpha=1$, we have Baselga's Repl_{BJ} coefficient. For $\alpha=0.5$, equation A2 reduces to the
493 relativized replacement index Repl_J . Thus, it may be conceived as a function
494 intermediate between two extremes which consider the number of unique species
495 asymmetrically. In other words, Repl_J measures turnover *with respect to the total*
496 *species richness of the two sites*. If $\alpha < 1$, then both b and c are considered, and
497 therefore the measure is not independent of raw richness difference (see Appendix 1
498 for proofs, and Supplementary Material for graphical demonstrations).

499 For the Sorensen family of indices we may also define a general formula with α
500 ranging from 0 to 1:

501

$$502 \text{Repl}_{\alpha S} = \frac{\min(b,c)}{a + \alpha \min(b,c) + (1 - \alpha) \max(b,c)} \quad (\text{A3})$$

503

504 If $\alpha=0$, we get the analogue of equation A1, known as the Savage index (its
505 complement is called the Braun-Blanquet similarity, see e.g. Háyek 1994, Tulloss
506 1997). For $\alpha=1$, it is easy to see that the general formula becomes identical to the
507 Simpson dissimilarity function, while $\alpha=0.5$ yields the replacement component of
508 Sorensen dissimilarity in the POD framework

509

$$510 \text{Repl}_S = \frac{\min(b,c)}{a + 0.5b + 0.5c} = \frac{2 \min(b,c)}{2a + b + c} \quad (\text{A4})$$

511

512 not discussed in this paper. This latter index is therefore a replacement measure with
513 respect to the total number of presences in the two sites.

514 Note that general formulae similar to A2 and A3 have been known from numerical
515 ecology. For example, we can mention the intermediate coefficient suggested by Faith
516 (1984) as generalized by Podani (2000, p. 108). This represents a continuum from
517 Kendall dissimilarity to the Manhattan metric.

518

519 **References to appendices**

520 Chao, A., Chiu, C.H. and Hsieh, T.C. (2012) Proposing a resolution to debates on
521 diversity partitioning. *Ecology*, 93, 2037-51.

522 Faith, D. P. (1984) Patterns of sensitivity of association measures in numerical
523 taxonomy. *Mathematical Biosciences*, 69, 199-207.

524 Hayek, L.-A. C. (1994) Analysis of amphibian biodiversity data. *Measuring and*
525 *Monitoring Biological Diversity. Standard Methods for Amphibians*. (eds. W. R.
526 Heyer et al.), pp. 207-269. Smithsonian Institution, Washington, D. C.

527 Podani, J. 2000. *Introduction to the Exploration of Multivariate Biological Data*.
528 Backhuys, Leiden.

529 Tulloss, R.E. (1997) Assessment of similarity indices for undesirable properties and a
530 new tripartite similarity index based on cost functions. *Mycology in Sustainable*
531 *Development: Expanding Concepts, Vanishing Borders*. (eds. M.E. Palm & I.H.
532 Chapela), pp. 122-143. Parkway Publ., Boone, NC.

533

534

535