



**HAL**  
open science

## Ecological versatility and the assembly of multiple competitors: cautionary notes for assembly inferences

Victor Saito, Fabien Laroche, Tadeu Siqueira, Sandrine Pavoine

### ► To cite this version:

Victor Saito, Fabien Laroche, Tadeu Siqueira, Sandrine Pavoine. Ecological versatility and the assembly of multiple competitors: cautionary notes for assembly inferences. *Ecology*, 2018, 99 (5), pp.1173-1183. 10.1002/ecy.2197. hal-02551812

**HAL Id: hal-02551812**

**<https://hal.science/hal-02551812>**

Submitted on 23 Apr 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 Running head: Assembly of multispecies communities

2 **Ecological versatility and the assembly of multiple competitors: cautionary notes for**  
3 **assembly inferences**

4 Victor S. Saito<sup>a,b</sup>, Fabien Laroche<sup>c</sup>, Tadeu Siqueira<sup>d</sup>, Sandrine Pavoine<sup>b</sup>

5 *<sup>a</sup>Departamento de Ciências Ambientais - DCAm and Programa de Pós-Graduação em Ecologia e Recursos,*  
6 *Universidade Federal de São Carlos, São Carlos SP, Brazil, VSS E-mail: victor.saito@gmail.br*

7 *<sup>b</sup>Centre d'Ecologie et des Sciences de la Conservation (CESCO UMR7204), Sorbonne Universités, MNHN,*  
8 *CNRS, UPMC, CP51, 43-61 rue Buffon, 75005, Paris, France, SP E-mail : sandrine.pavoine@mnhn.fr*

9 *<sup>c</sup>Irstea, UR EFNO, F-45290 Nogent-sur-Vernisson, France, FL E-mail: fabien.laroche@irstea.fr*

10 *<sup>d</sup>Universidade Estadual Paulista (UNESP), Instituto de Biociências, Rio Claro, Brazil, TS E-mail:*  
11 *tsiqueira@rc.unesp.br*

12 \*Corresponding author: [victor.saito@gmail.com](mailto:victor.saito@gmail.com)

13 Full address : Rodovia Washington Luis, km 235, SP-310, CEP 13565-905, São Carlos, São  
14 Paulo, Brazil.

15

16

17

18

19

20

21

22

23

24

25

26 **ABSTRACT**

27 The role of niche differences and competition is invoked when one finds coexisting species to  
28 be more dissimilar in trait composition than expected at random in community assembly  
29 studies. This approach has been questioned as competition has been hypothesized to either  
30 lead to communities assembled by similar or dissimilar species, depending on whether  
31 species similarity reflects fitness or niche differences, respectively. A current problem is that  
32 the arguments used to draw relationships between competition and species similarity are  
33 based on pairwise theoretical examples, while in nature competition can occur among a  
34 constellation of species with different levels of versatility in resources used. By versatility we  
35 mean the documented ability of some species to escape competition for commonly used  
36 resources by changing for marginal and unused resources. Thus, a versatile species will have  
37 the ability to decrease niche overlap with all other species when facing strong competitors.  
38 When these species are embedded in multiple interactions the role of pairwise niche and  
39 fitness differences could be reduced due to indirect effects and thus competition would not be  
40 detectable. Here we developed a coexistence theory where competition occurs simultaneously  
41 among multiple species with different levels of versatility and then used it in a simulation to  
42 unravel patterns of species similarity during community assembly. We found that simulated  
43 communities can be assembled by species with more, less or equal similarity compared to a  
44 null model when using a mean distance based metric (SES.MPD). However, contrasting these  
45 varied results, we consistently found species overdispersion using a nearest neighbor-based  
46 metric (SES.MNTD), even when species differences reflected more directly their competitive  
47 abilities than their niche differences. Strong tendency to overdispersion emerged when high  
48 ecological versatility promoted large niche differences and enabled coexistence. This is  
49 because versatility to use marginal resources compensates possible fitness differences among  
50 species. Our findings provide mounting evidence of the important role of minimum niche

51 differences and versatility in resource consumption for species embedded in multiple direct  
52 and indirect interactions.

53

54 *Keywords:* assembly rules, multispecies competition, functional diversity, limiting similarity,  
55 functional versatility, niche theory, clustering, overdispersion.

56

## 57 **INTRODUCTION**

58 It has become common practice in community ecology to assume that communities  
59 composed of more dissimilar species than expected at random (hereafter species  
60 overdispersion) are predominantly assembled by interspecific competition (Weiher and  
61 Keddy 1995; Webb 2000; HilleRisLambers et al. 2012; D'Andrea and Ostling 2016). This  
62 inference is grounded on MacArthur and Levins's (1967) limiting similarity hypothesis,  
63 which claims the existence of a limit to the resource use overlap between species for their  
64 stable coexistence. Accordingly, communities composed of similar species (hereafter species  
65 clustering) are seen as evidence for relaxed competition for resources and strong abiotic  
66 filtering selecting phenotypes well adapted to the environment (Weiher and Keddy 1995).

67 After years of intense application of this approach, Mayfield and Levine (2010)  
68 challenged it by explaining that species clustering can also be the result of interspecific  
69 competition if species similarities (e.g. calculated using morphologic, behavioral, metabolic  
70 traits) reflect their competitive hierarchy (see also HilleRisLambers et al. 2012; Adler et al.  
71 2013; D'Andrea and Ostling 2016; Cadotte and Tucker 2017). This happens because, in  
72 contrast to the traditional view, not all species differences should promote coexistence.  
73 According to Chesson (2000), species coexistence depends on two kinds of interspecific  
74 differences. First, stabilizing niche differences (hereafter niche differences) favor coexistence  
75 by reducing the overlap in resource use between species in time or space, making

76 intraspecific competition more intense than interspecific competition (Chesson 1991).  
77 Second, species-level average fitness differences (hereafter fitness differences) determine  
78 which species will outcompete others in absence of niche differences (Chesson 2000). A  
79 balance between niche and fitness differences determines stable coexistence and the pairwise  
80 coexistence condition is that niche differences must overcome fitness differences (Chesson  
81 2000). So, if species similarities relate to their fitness equalities, then one could expect to find  
82 species clustering around an optimum trait value when competition rules the assembly  
83 (Mayfield and Levine 2010).

84         Understanding how differences in species morphology, behaviors and strategies  
85 influence niche and fitness differences between them is a complex challenge (Godoy et al.  
86 2014; Herben and Goldberg 2014; Kraft et al. 2015; D’Andrea and Ostling 2016). Several  
87 traits can impact both niche and fitness differences, even if the net impact on one is higher  
88 than on the other (Kraft et al. 2015). For example, two animal species competing for different  
89 prey can have distinct preferences for prey size and taxonomic identity of prey. Differences  
90 in taxonomic preferences could favor coexistence through resource partitioning but size  
91 preferences could lead to competitive asymmetry given that feeding on larger animals may  
92 improve species fitness. Thus, differences in taxonomic and size preferences would  
93 contribute to both niche and fitness differences. Therefore, it is not clear whether analyzing  
94 together traits that drive stabilizing and equalizing mechanisms simultaneously should give  
95 clustered, segregated or random patterns.

96         Beyond the importance of niche and fitness differences for competitive outcomes, two  
97 characteristics of species competition are usually overlooked in community assembly studies.  
98 First, competition generally happens between several species in nature (Pianka 1974), not  
99 only in pairwise situations, as commonly considered in conceptual and experimental  
100 examples (Mayfield and Levine 2010; Chu and Adler 2015). The addition of a third (or more)

101 species to a community composed of two species modifies the competitive network. This  
102 happens because three species can experience indirect effects of each other (Levine 1976;  
103 Lawlor 1979). For example, let us consider species A, B and C: species A has a direct effect  
104 on species B by directly competing for resources, but also has an indirect effect by competing  
105 with the third species C, which is a common competitor of species A and B – thus the full  
106 impact of species A on B would be the sum of the direct and indirect effects. Given that, the  
107 pairwise coexistence condition – i.e. niche differences must exceed fitness differences  
108 (Chesson 2000) – is not the only determinant of species coexistence, as it is also influenced  
109 by the other species that compose such a community.

110         Second, one of the most used models of species competition considers that  
111 competition occurs over a linear array of resources and each species have a trait that makes  
112 them able to consume a range of this array of resources (the MacArthur’s consumer-resource  
113 model, 1970). In this case, trait differences determine the resource use overlap and thus the  
114 interaction strength among species (see Fig. 1). However, competition in nature can happens  
115 in a different way. In other documented cases, all species will be able to consume one  
116 common resource (in general an abundant resource), but they may vary in the ability to  
117 consume other resources, changing their dependency on this common resource when facing  
118 competition (Liem 1980; 1990; Bellwood et al. 2006). For example, Liem (1980) found that  
119 morphological differentiation in mouth apparatus in cichlids determine the ability of species  
120 to escape competition for common preys (by changing for a different resource) in the  
121 presence of strong competitors. In this case, differences in traits impacts their versatility and  
122 thus could determine niche differences with all other species simultaneously using the  
123 coexistence framework (see Fig. 1 for a schematic explanation of the differences between  
124 MacArthur’s model and our model considering species versatility).

125 So, based on the potential impact that multispecies competition and species versatility  
 126 may have on community assembly, we investigated the strength of commonly used metrics of  
 127 species dispersion (e.g. mean pairwise and nearest neighbor metrics, Weiher and Keddy  
 128 1995; Webb 2000) to detect signatures of competition under a mechanistic model of  
 129 competition. We first developed the underlying theory of competition considering species  
 130 versatility upon Chesson's (2000) coexistence theory and then used it within a simulation  
 131 based on sequential colonization from a species pool. We observed the emergence of species  
 132 clustering, overdispersion and random patterns in simulated communities fueling the ongoing  
 133 discussion on the use of species dispersion metrics to infer assembly processes.

134

## 135 **METHODS**

### 136 **Theory development**

137 We used a model to simulate community assembly based on modern coexistence theory and  
 138 under multispecies competition (see Appendix S2 in Supporting Information for a detailed  
 139 version of the model and proofs). We aimed a reference model that respects Chesson's (2000)  
 140 following to the long-term low-density growth rate of a species  $i$  in a multispecies system  
 141 (Eq. 4 in Chesson 2000):

$$142 \quad \bar{r}_i \approx b_i \left( k_i - \frac{\sum_{s \neq i} k_s}{n-1} \right) + b_i \frac{(1-\rho)}{n-1} D \quad (1)$$

143 where  $n$  is the number of species in the system, the  $k_s$  are measures of fitness of individual  
 144 species,  $(\sum_{s \neq i} k_s)/(n-1)$  is the average fitness of the competitors of  $i$ ,  $\rho$  is a measure of  
 145 niche overlap, and  $D$  is a positive constant. This equation is composed of an equalizing term

$$146 \quad b_i \left( k_i - \frac{\sum_{s \neq i} k_s}{n-1} \right) \quad (2)$$

147 and a stabilizing term

$$148 \quad b_i \frac{(1-\rho)}{n-1} D \quad (3)$$

149 In this kind of community model, a set of species is said to coexist if all individual species  
 150 can grow from low density in the community made of the  $n - 1$  other species at equilibrium  
 151 (i.e. community feasibility).

152 We started with MacArthur's consumer-resource model where  $n$  species compete for  $m$   
 153 resources (Eqs. 1 and 2 in Chesson 1990):

$$154 \quad \begin{cases} \frac{1}{X_i} \frac{dX_i}{dt} = b_i [\sum_{l=1}^m c_{il} w_l R_l - m_i] \\ \frac{1}{R_l} \frac{dR_l}{dt} = r_l \left( 1 - \frac{R_l}{K_l} \right) - \sum_{i=1}^n c_{il} X_i \end{cases} \quad (4)$$

155 where  $i$  is a consumer species;  $X_i$  is population density;  $R_l$  is the density of food resource  $l$ ;  $w_l$   
 156 is the value of one unit of resource  $l$  to the consumer;  $c_{il}$  is the rate at which consumer  
 157 species  $i$  captures resource  $l$  per unit abundance of resource  $l$ ;  $m_i$  is the total value of resource  
 158 that must be harvested per capita for the growth rate to be exactly 0;  $b_i$  is a factor converting  
 159 the resource excess into the per capita growth rate;  $r_l$  is the growth rate and  $K_l$  the carrying  
 160 capacity for resource  $l$ .

161 Assuming a separation of time scales between resource and consumer dynamics yields (see  
 162 Appendix S2):

$$163 \quad \frac{1}{\tilde{X}_i} \frac{d\tilde{X}_i}{dt} = \tilde{b}_i [\tilde{k}_i - \sum_{j=1}^n \tilde{\alpha}_{ij} \tilde{X}_j] \quad (5)$$

164 where  $\tilde{X}_i$  and  $\tilde{b}_i$  are rescaled quantities corresponding to  $X_i$  and  $b_i$  from equation (4),  $\tilde{k}_i$  is a  
 165 rescaled quantity corresponding to  $\sum_{l=1}^m c_{il} w_l R_l - m_i$  in equation (4) and  $\tilde{\alpha}_{ij}$  is a rescaled  
 166 interaction coefficient between species  $i$  and  $j$ , comprised between 0 and 1, which  
 167 corresponds to a resource overlap between the two species (in particular  $\tilde{\alpha}_{ii} = 1$ ).

168 Consider species  $i$  at low density, and assume that the  $n - 1$  other species can coexist. Then  
 169 all the other species  $j \neq i$  harboring their equilibrium density  $\tilde{X}_{ji}$  (second subscript “ $i$ ”  
 170 meaning “in the absence of species  $i$ ”). Then the long-term growth rate at low density of  
 171 species  $i$ ,  $\bar{r}_i$ , is obtained by neglecting  $\tilde{X}_i$  and setting the  $\tilde{X}_j$ s to  $\tilde{X}_{ji}$ s in (5):



$$172 \quad \bar{r}_i = \tilde{b}_i \left[ \tilde{k}_i - \sum_{\substack{j=1 \\ j \neq i}}^n \tilde{\alpha}_{ij} \tilde{X}_{ji} \right] \quad (6)$$

173 To derive an equation for the long-term low-density growth rate of a species in the form of  
 174 Chesson's approximation for multispecies system (Eq. 1), we considered a particular type of  
 175 interactions among species that reflect species versatility: for all pairs of species  $i \neq j$ , we  
 176 assumed that  $\tilde{\alpha}_{ij} = \theta_i \theta_j$ , where  $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_n)$  is a vector of positive real numbers  
 177 between 0 and 1. In particular,  $\tilde{\alpha}_{ii} = 1$ . Empirically,  $\boldsymbol{\theta}$  can be a measure of resource use  
 178 restriction or versatility (e.g. Bellwood et al. 2006 used the species distance to the centroid of  
 179 a multidimensional space of species distances as a measure of versatility). High values of  
 180  $\boldsymbol{\theta}$  means that a species will depend mostly on the common resource all species can consume,  
 181 while species with low  $\boldsymbol{\theta}$  will be those with the versatility to use marginal resources when  
 182 facing competition (see Fig. 1).

183 Importantly, we show in Appendix S2 that any set of interactions of this type can be obtained  
 184 by adjusting the basic parameters of the model (resource renewal dynamics and consumption  
 185 vectors), as long as the number of resource types is equal or greater than the number of  
 186 species. As similar to the MacArthur's model, resource types can be anything that can vary  
 187 and can be depleted for other species, i.e. seed size ranging from 1 to 100 cm can be  
 188 considered 100 resources for birds if each size encompasses a resource that provide niche  
 189 differences and can be limiting:

$$190 \quad \bar{r}_i = \tilde{b}_i \left( \tilde{k}_i - \frac{\theta_i}{\left(1 + \sum_{\substack{l=1 \\ l \neq i}}^n \frac{\theta_l^2}{1 - \theta_l^2}\right)} \sum_{\substack{j=1 \\ j \neq i}}^n \frac{\theta_j}{1 - \theta_j^2} \tilde{k}_j \right) \quad (7)$$

191 We now seek to decompose equation (7) into stabilizing and equalizing terms as in Chesson's  
 192 (2000) framework. To find the stabilizing term, we first looked at the long-term growth rate  
 193 at low density when all the species have the same fitness  $\tilde{k}$ . In this case, only the stabilizing  
 194 should act on coexistence. When all species have equal fitness, (7) becomes:

195 
$$\bar{r}_i = \tilde{b}_i \tilde{k} \left( 1 - \frac{\theta_i \sum_{j=1}^n \frac{\theta_j}{1-\theta_j^2}}{\sum_{l=1}^n \frac{\theta_l^2}{1-\theta_l^2}} \right) \quad (8)$$

196 We now search for the corresponding equalizing term in an additive decomposition as in  
 197 Chesson's (2000) framework (Eq. 1 above). Starting from (7), we obtained (see Appendix S2  
 198 for detailed proofs) the following final expression for the long-term growth rate of a species:

199 
$$\bar{r}_i = \tilde{b}_i \left( \tilde{k}_i - \frac{\sum_{j=1}^n \frac{\theta_j}{1-\theta_j^2} \tilde{k}_j}{\sum_{j=1}^n \frac{\theta_j}{1-\theta_j^2}} + \frac{\sum_{j=1}^n \frac{\theta_j}{1-\theta_j^2} \tilde{k}_j}{\sum_{j=1}^n \frac{\theta_j}{1-\theta_j^2}} \left( \mathbf{1} - \frac{\sum_{j=1}^n \frac{\theta_i \theta_j}{1-\theta_j^2}}{1 + \sum_{l=1}^n \frac{\theta_l^2}{1-\theta_l^2}} \right) \right) = \tilde{b}_i (\tilde{k}_i - \bar{k}_i + (1 - \rho_i) \bar{k}_i) \quad (9)$$

200 where

201 
$$\bar{k}_i = \frac{\sum_{j=1}^n \frac{\theta_j}{1-\theta_j^2} \tilde{k}_j}{\sum_{j=1}^n \frac{\theta_j}{1-\theta_j^2}}$$

202 
$$\rho_i = \frac{\theta_i \sum_{j=1}^n \frac{\theta_j}{1-\theta_j^2}}{1 + \sum_{l=1}^n \frac{\theta_l^2}{1-\theta_l^2}}$$

203

204 We thus retrieved in (9) a stabilizing term with a form consistent with Chesson's (2000)  
 205 equation (Eq. 1).

206 The main difference with Chesson's (2000) framework is that  $\bar{k}_i$  is not a simple arithmetic  
 207 mean of the fitness values but a weighted mean, where weights  $(\frac{\theta_j}{1-\theta_j^2})$  depend on some  
 208 interaction effects with other species. Changing interactions among species thus changes the  
 209  $\bar{k}_i$ . Our theory provides a multispecies extension of the pairwise Chesson's theory in the case  
 210 when interactions strength among species simultaneously affects in the same direction the  
 211 competition with all the other species. This is because all species can consume a common  
 212 resource, but vary in their versatility to consume other resources (see Fig. 1).

213

## 214 **Simulation description**

215 To simulate community assembly, we created a species pool with 100 species. For  
216 each species we settled values of  $k_i$  and  $\theta_i$  – the parameters of species in our coexistence  
217 model (Eq. 9). The values of these species parameters determine niche and fitness differences  
218 and thus both are necessary to determine coexistence in communities using equation (Eq. 9).  
219 Biologically,  $k_i$  would be related to one or several morphological, behavioral or other traits  
220 that influence species competitive ability. Theta, however, represents a synthetic measure of  
221 ecological versatility indicating the extent to which a species depends on a common resource  
222 that all other species can consume (see Fig. 1). We calculated fitness differences and niche  
223 differences resulting in two species distance matrix, one that represented species niche  
224 overlap ( $1-\theta_i\theta_j$ ) (from 0 to 1) and another that represented competitive differences ( $|\tilde{k}_i -$   
225  $\tilde{k}_j|$ ). Subsequent analyses of species similarity were based on these two species distance and  
226 not on  $\tilde{k}_i$ s and  $\theta_i$ s directly. So, species competitive ability determines competitive exclusion  
227 in the lack of niche differences, while species versatility determines the niche overlap among  
228 species.

229 We simulated values for parameters  $k_i$  and  $\theta_i$  using the beta distribution, which is a  
230 continuous probability distribution defined on  $[0, 1]$  with two positive parameters:  $\alpha$  and  $\beta$ .  
231 We used  $\alpha$  and  $\beta$  to control the means of  $k_i$  and  $\theta_i$  values and their variation. It is important to  
232 stress that low  $\alpha$  values for  $\theta_i$  means that, overall, species are highly versatile. Controlling the  
233  $\alpha$  and  $\beta$  parameters in the beta distribution allowed us to simulate various scenarios for the  
234 distribution of fitness and niche differences in the species pool (Fig. 2). Differences between  
235 the parameters of the beta distribution (Fig. 2) are due to differences in the way fitness and  
236 interaction strength interfere in the model, fitness in an additive manner and interaction  
237 strength in a multiplicative manner.

238           Once the species pool was defined, we built the model to simulate community  
239 assembly. The model requires a given *a priori* number of colonization events to occupy a  
240 community. The species pool is unchanged so that a single species can try to invade a  
241 community several times. The first species is randomly selected from the pool and occupies a  
242 community. Then, the model randomly selects a second species to invade the community and  
243 applies the coexistence condition (Eq. 9) to decide whether the first and the second species  
244 can coexist in that community. If these species can coexist, the next (third) invader would be  
245 tested using the multispecies coexistence condition (Eq. 9). Subsequent colonizing species  
246 are tested for the coexistence condition with all species that already coexist there. The  
247 process of assembly ends when all the colonization events defined *a priori* were done. We  
248 also simulated scenarios where species have different colonization probabilities depending on  
249 their  $k_i$  and  $\theta_i$ , one with a colonization-competition trade-off and another where species that  
250 rely mostly on a common resource have higher chance to colonize new patches. However,  
251 these simulations provided highly similar results to the random dispersal approach. Thus,  
252 these are only described in Appendix S3.

253           In this model, when the coexistence condition was not satisfied and the long-term  
254 growth rate of the invader was negative, then the invader was excluded. If the coexistence  
255 was not satisfied and the long-term growth rate of the invader was positive, then the species  
256 with the minimum long-term growth rate was excluded and the coexistence condition was  
257 tested in the new resulting community. This process was repeated until the coexistence  
258 condition was satisfied. We relied on previous work (Tregonning and Roberts 1979; Wilson  
259 and Lundberg 2004) to assume that the iterative removal of species with negative growth  
260 rates provides a reasonable approximation of the equilibrium community.

261           We defined a simulation as each round starting from simulating the species pool to  
262 the assembling of 100 communities. So, each simulation has one singular species pool to

263 assemble 100 communities. A simulation thus ends with a set of 100 communities assembled  
264 under the same conditions.

265

## 266 **Community metrics**

267 To analyze the communities, we calculated species richness, beta diversity (Jaccard  
268 dissimilarity) and two metrics of species dispersion. Among species dispersion metrics, we  
269 used the standardized effect size of mean pairwise distances (SES.MPD) and that of the mean  
270 distance to the most similar species within communities (SES.MNTD , Webb 2000). We used  
271 the ‘taxa shuffle’ null model that randomizes 1,000 times the rows and columns of the species  
272 distance matrix. An additional metric of functional dispersion was also calculated. However  
273 this metric provided very similar results from SES.MNTD and SES.MPD, so we included  
274 these results in Appendix S3.

275 As a first step, we ran 100 simulations, with 100 invasions in each community, for  
276 each combination of fitness and niche distributions (Fig. 2). In these simulations, we first  
277 applied the species dispersion metrics (SES.MNTD and SES.MPD) to each assembled  
278 community, but separately to fitness distances and niche distances. Next, we applied the  
279 dispersion metrics to a combined niche and fitness-based distance between species named  
280 'overall species distance' and defined as:  $\frac{1}{2\max(\text{Fitness distances})} + \frac{1}{2\max(\text{Niche distances})}$ , where  
281 maxima are over all species pairs in the species pool. Finally, to simulate studies gathering  
282 traits more related to niche or fitness differences among species (see Kraft et al. 2015), we  
283 modified the way we defined overall species distances as:

284  $\frac{a}{a+b\max(\text{Fitness distances})} + \frac{b}{a+b\max(\text{Niche distances})}$ , where maxima also are over all species pairs in  
285 the regional species pool and  $a$  and  $b$  are positive parameters. This simulation represents  
286 studies gathering different number of traits related to niche differences (larger values of  $b$ )  
287 and fitness differences (larger values of  $a$ ), or also simulates studies where one singular trait

288 relates to both niche and fitness differences but with different weights (e.g. Godoy et al.  
289 2014; Kraft et al. 2015). In a second step, we varied the number of colonization events per  
290 community considering 30, 50, 100, 300 and then 500 colonization events.

291

## 292 **RESULTS**

293         When communities were assembled by 100 colonization events, we found a positive  
294 trend in species richness by increasing the niche differences (and versatility) and the fitness  
295 values at the species pool scale (Fig. 3a). Beta diversity was high on average, independently  
296 of niche differences and species fitness at the species pool scale (Fig. 3b).

297         Considering the species dispersion metrics, when species distances were calculated  
298 using niche differences only, we found consistent overdispersion (Fig. 3c,d). With larger  
299 niche differences and higher fitness at the species pool scale, overdispersion tended to be  
300 more evident. When species distances were calculated using fitness differences only, we  
301 found clustering for most simulations (Fig. 3e,f). The exceptions were a tendency to  
302 randomness when mean species  $k$  were low at the species pool scale (Fig 3c first boxplots  
303 within each color) and when niche differences were high at the species pool scale for the  
304 SES.MNTD metric (Fig. 3e, yellow boxplots). When overall species distances were  
305 calculated, we found a strong tendency towards overdispersion using the SES.MNTD metric  
306 (Fig. 3g), especially when niche differences at the species pool scale were large (species  
307 where highly versatile) (Fig 3g yellow boxes). The only exception happened when the species  
308 pool had small  $ks$  values (Fig. 3g, first blue box).

309         When using the SES.MPD metric, results were more variable: 1) overdispersion when  
310 species had large niche differences at the species pool (Fig. 3h, yellow boxes) or similarly  
311 high  $ks$  at the species pool (Fig. 3h, third boxes within each color); 2) a tendency to clustering  
312 when species  $ks$  had a U-shaped distribution (Fig. 2h, forth blue and red boxes); 3) random

313 patterns when species had low or medium  $k$  and/or niche differences at the species pool (Fig.  
314 3h, first and second blue and red boxes).

315         When varying the weights of niche differences and fitness differences in species  
316 overall distance (i.e. varying the parameters  $a$  and  $b$  when calculating overall distances), we  
317 found that the weights strongly impacted dispersion metrics, independently of the species  
318 pool properties. For brevity, we only show the results when species have medium niche  
319 differences and medium  $k$ s values at the species pool scale. We found that SES.MNTD was  
320 overdispersed even when we gave more weight to fitness differences than niche differences  
321 (Fig. 4a). There was a tendency towards randomness and clustering when fitness differences  
322 were four to five times more important than niche differences (Fig. 4a). Using SES.MPD, we  
323 found random patterns when the weights of niche and fitness differences was the same, but  
324 clustering and overdispersion when there were more weights to fitness and niche differences,  
325 respectively (Fig. 4b).

326         Varying the number of sequential colonization events led to the same general  
327 conclusions (Fig .5). For brevity, we only show the results obtained with medium niche  
328 differences and fitness values at the species pool scale. Community richness increased with  
329 more colonization events, but stabilized after some point (Fig. 5a). Beta diversity decreased  
330 with increasing colonization reaching very low levels after 300 and 500 colonization events  
331 (Fig. 5b). We found consistent niche overdispersion and fitness clustering (Fig. 5c,d,e,f).  
332 Using overall species distance we found overdispersion using SES.MNTD with a minimum  
333 of 50 colonization events and only randomness using SES.MPD (Fig. 5g,h). Overall,  
334 dispersion patterns became more evident with increasing number of colonization events.

335

336 **DISCUSSION**

337 Community assembly studies have the ambitious aim of inferring processes by  
338 analyzing similarities in traits or in evolutionary history among species in a given community  
339 (Weiher and Keddy 1995; Webb 2000). We show, however, that such approaches can suffer  
340 from two types of Achilles' heels when linking patterns to processes. The first is when  
341 different processes result in the same pattern. We found that communities containing very  
342 similar species – a pattern used to infer environmental filtering (Weiher and Keddy 1995;  
343 Webb 2000) – can also be the result of competition alone (Mayfield and Levine 2010). The  
344 second is when one process leads to different patterns. We found that competition can result  
345 in communities containing very distinct species (i.e. an overdispersed pattern, as predicted by  
346 the traditional framework), but also in communities with clustered (as predicted by Mayfield  
347 and Levine 2010) and even random structure. These problems likely arise because we  
348 simulated communities with the inclusion of two scenarios rarely considered in previous  
349 studies: simultaneous competition among multiple competitors (e.g. Godoy et al. 2017) and  
350 species with different levels of versatility to exploit marginal resources. These results bring  
351 concerns to the use of the traditional community assembly approach, but enabled us to do a  
352 step forward in understanding recurrent patterns of communities in nature.

353 Here, we introduced an extension of the pairwise theory of coexistence using a  
354 variation of the MacArthur's consumer-resource model (Chesson 1990) that considers species  
355 versatility to determine the interaction strength among species. This new model differs in  
356 important points from the traditional view of competition (Fig. 1) but yield concordant  
357 patterns with previous inferences on the impacts of ecological versatility in community  
358 assembly. For example, species richness increased with species versatility in the community  
359 and in the species pool scale (Bellwood et al. 2006; Belmaker et al. 2012). Additionally, we  
360 found that versatility provides important niche differences and promotes coexistence when  
361 competitors from the pool are very strong (high fitness), as already suggested for empirical



362 communities (e.g. cichlids, Liem 1980). Overall, these concordances from theory and data  
363 emphasize the important role that versatility may play in the assembly of natural  
364 communities.

365         Our findings indicate that empirical studies able to separate niche from fitness traits  
366 (e.g. Kraft et al. 2015) could have found species overdispersion if species distances reflected  
367 niche differences, but clustering if it reflected fitness differences. However, in the case traits  
368 impact both fitness and niche differences (Adler et al. 2013; Kraft et al. 2015), our results  
369 suggest that MPD metrics can also be random. Previous empirical studies indeed found  
370 clustering and overdispersion using different traits with an overall species clustering, but the  
371 interpretation relied on the opposing effects of environmental filtering and limiting similarity  
372 (e.g. Swenson and Enquist 2009). Here, we demonstrated how these same patterns can arise  
373 due to competition only, when fitness have a U-shaped distribution (Fig. 3h). The U-shaped  
374 distribution is the only bimodal distribution we used and it demonstrated to promote different  
375 patterns compared to when fitness has unimodal distributions. Empirically, traits with  
376 bimodal distribution are recurrent in nature (e.g. marine zooplankton and stream biotas with  
377 bimodal size distribution [Warwick and Clarke 1984; Poff et al. 1993]). So, as these patterns  
378 can be found in nature, it is likely that clustering in some traits (related to competitive ability)  
379 and overdispersion in others (related to species versatility) are also possible when  
380 competition rules the assembling.

381         The distributions of niche and fitness differences in the species pool had pivotal  
382 effects in determining the level of species dispersion, especially in SES.MPD. With large  
383 niche differences in the species pool (i.e. species are highly versatile), overdispersion  
384 emerged consistently when using overall species distances. This is likely the result of a  
385 possible coexistence when most species have large niche differences, independently of their  
386 fitness differences (Chu and Adler 2015). A strong tendency towards overall overdispersion

387 also emerged with high fitness at the species pool scale. This contradicts the idea that  
388 communities dominated by strong competitors should have clustered patterns (Mayfield and  
389 Levine 2010), at least when species vary in their dependency for a common resource.  
390 Actually we found that, in these cases, species would tend to be high versatile to overcome  
391 possible large fitness disadvantages. In a more conservative case, when the species pool has  
392 medium fitness values and medium niche differences, overall distance was in most cases  
393 random (Figs. 2h and 4h). Given the commonness of a bell-shaped distribution of traits in  
394 nature (e.g. traits evolving under a Brownian motion commonly result in normal distributions  
395 around an ancestral trait value, Kraft et al. [2007]), it is not surprising that most studies to  
396 date have observed random patterns when analyzing mean distance-based metrics (82%  
397 according to a review by Götzenberger et al. [2012]). Thus, we support recent studies  
398 cautioning the inference of stochastic processes, e.g. random variation in birth and death  
399 rates, as the main processes governing communities with random dispersion patterns (Gallien  
400 2017).

401 Besides the cautionary notes, we found one consistent result across most model  
402 variations: consistent overdispersion when using the SES.MNTD metric. This finding  
403 supports the importance of minimum niche differences between species in stable multispecies  
404 communities (Godoy et al. 2017). This is likely because a high ability to escape competition  
405 for a common resource will help overshadow the competitive differences with all other  
406 species (e.g. Liem 1980). After years of criticism (Mayfield and Levine, 2010; Adler et al.  
407 2013; Gallien 2017), we add new evidence that the signature of pairwise niche differences in  
408 communities assembled by multispecies competition is the same overdispersion hypothesized  
409 by Weiher and Keddy (1995), even when indirect effects due to multiple interactions could  
410 act to weaken such pattern (Godoy et al. 2017). It is also important to note that Godoy et al  
411 (2017) used a different model of competition (an annual plant species model, Godoy et al.

412 2014), but also found the similar pattern of overdispersion. This highlights the general  
413 importance of minimum niche differences even when considering competition in different  
414 ways.

415 In this recent study, Godoy et al. (2017) also discussed the role of niche differences  
416 and intransitivity (rock-paper-scissor dynamic) in maintaining multispecies communities,  
417 explaining that both would give similar empirical signals of overdispersion. However, for  
418 intransitive competition to promote coexistence, species need to differ in their ability to  
419 compete for different resources, thus making possible that a superior competitor for resource  
420 A is worse when competing for resource B (Gallien et al. 2017). However, this scenario  
421 cannot arise in our model as competitive ability is a general characteristic, not related to each  
422 specific resource; i.e. we would never have a species with  $k=10$  being outcompeted by a  
423 species with  $k=5$ . In our model, the coexistence of species with larger differences than  
424 expected at random is not related to possible intransitivity but specifically to stabilizing  
425 versatility in resources used. So, by using a model that isolates the role of niche differences,  
426 we are able to support the idea that intransitivity emerging from multispecies competition  
427 should not be pivotal for species coexistence in the absence of sufficient pairwise niche  
428 differences (Godoy et al. 2017).

429 Finally, we must punctuate some limits of our approach. First, our model is an  
430 interaction chain, where a species A has an indirect effect on species C by also competing  
431 with the common competitor B. Another type of multispecies competition considers higher  
432 order interactions, where pairwise competition changes in the presence of a third competitor,  
433 so that the impact of A on B is different when in pairwise or in triplets. The distinction  
434 between interaction chains and higher-order interactions is determined by whether the  
435 indirect effect emerges from a change in competitor density (interaction chain) or a change in  
436 the per capita competitive effects (higher-order interaction). As we did not consider such

437 dynamics in our model, we cannot anticipate the impacts that these would have on the  
438 patterns we found (see Levine et al. 2017). Second, in our study, we did not explicitly include  
439 temporal or spatial variation mechanisms of coexistence. With more varying niche  
440 dimensions, it is possible that local differences in resource use became less important and  
441 thus local overdispersion may become weaker (Kraft et al. 2015). Third, our model is  
442 strongly deterministic, not including neutral population dynamics that could modify  
443 community patterns. Studies on these dynamics have shown that removing niche differences  
444 can have strong impacts on species coexistence. For example, D’Andrea and Ostling (2017)  
445 found that neutral cases can even harbor larger number of coexisting species than niche based  
446 scenarios, especially if the number of available niches is limited. So, a way forward would be  
447 to investigate how such dynamics would modify the patterns found here.

448         Concluding, here we introduced a new model of competition where species versatility  
449 promotes niche differences. This model proved to be an interesting variation of the  
450 MacArthur’s consumer-resource model when species vary in their dependency for a  
451 commonly exploited resource. Competition in nature is complex and should happen, not only  
452 in both ways, but also in many other ways. Besides this high complexity, we found that  
453 overdispersion using nearest neighbor metrics was a recurrent pattern and thus could  
454 signalize a master role that competition plays during community assembly. After some years  
455 of debate, our model and results highlight the role of species versatility and niche differences  
456 and bring new perspectives for those aiming to infer the role of competition in molding  
457 community patterns.

458

## 459 **ACKNOWLEDGEMENTS**

460 We are thankful to Raul Costa Pereira, Andros Gianuca, Franck Jabot, Ian Carroll, François  
461 Massol, Matthew Leibold and Peter Chesson for comments that sharpened our ideas. We also

462 thank Annette Ostling, Nathan Kraft and an anonymous referee for comments and  
463 discussions during the review process that greatly improved our manuscript. This study was  
464 funded by grants #2013/50424-1, #2013/20540-0, #2014/24532-4 São Paulo Research  
465 Foundation (FAPESP) and by grant 403723/2012-4, Conselho Nacional de Desenvolvimento  
466 Científico e Tecnológico (CNPq). VSS thanks the Centre d'Ecologie et de Sciences de la  
467 Conservation (CESCO) from the Muséum National d'Histoire Naturelle in Paris for kindly  
468 receiving him during part of the elaboration of this manuscript.

469

## 470 **REFERENCES**

- 471 Adler, P. B., A. Fajardo, A. R. Kleinhesselink and N. J. B. Kraft. 2013. Trait-based tests of  
472 coexistence mechanisms. *Ecology Letters* 16:1294–1306.
- 473 Bellwood, D. R., P.C. Wainwright, C.J. Fulton and A.S. Hoey. 2006. Functional versatility  
474 supports coral reef biodiversity. *Proceedings of the Royal Society of London B:  
475 Biological Sciences* 273: 101-107.
- 476 Belmaker, J., C.H. Sekercioglu and W. Jetz. 2012. Global patterns of specialization and  
477 coexistence in bird assemblages. *Journal of Biogeography* 39 , 193-203.
- 478 Cadotte, M. W. and C. M. Tucker 2017. Should Environmental Filtering be Abandoned?  
479 *Trends in Ecology and Evolution* 32: 429-437.
- 480 Chesson, P. 1991. A need for niches? *Trends in Ecology and Evolution* 6:26–8.
- 481 Chesson, P. 2000. Mechanisms of Maintenance of Species Diversity. *Annual Review of  
482 Ecology, Evolution and Systematics* 31:343–358.
- 483 Chesson, P. 1990. MacArthur's consumer-resource model. *Theoretical Population  
484 Biology* 37:26-38.
- 485 Chu, C. and P. B. Adler. 2015. Large niche differences emerge at the recruitment stage to  
486 stabilize grassland coexistence. *Ecological Monographs* 85:373–392.

487 Devictor, V., R. Julliard and F. Jiguet. 2008. Distribution of specialist and generalist species  
488 along spatial gradients of habitat disturbance and fragmentation. *Oikos* 117: 507-514.

489 D'Andrea, R., and A. Ostling. 2016. Challenges in linking trait patterns to niche  
490 differentiation. *Oikos* 125: 1369-1385.

491 D'Andrea, R., and A. Ostling. 2017. Biodiversity maintenance may be lower under partial  
492 niche differentiation than under neutrality. *Ecology*. Early view.

493 Gallien, L. 2017. Intransitive competition and its effects on community functional diversity.  
494 *Oikos* 126: 615-623.

495 Gallien, L., N. E. Zimmermann, J. M. Levine and, P. B. Adler. 2017. The effects of  
496 intransitive competition on coexistence. *Ecology Letters* 20: 791–800.

497 Godoy, O., N. J. B. Kraft and J. M. Levine. 2014. Phylogenetic relatedness and the  
498 determinants of competitive outcomes. *Ecology Letters* 17:836–844.

499 Godoy, O., D. B. Stouffer, N. J. Kraft and J. M. Levine. 2017. Intransitivity is infrequent and  
500 fails to promote annual plant coexistence without pairwise niche differences. *Ecology*  
501 98: 1193-1200.

502 Götzenberger, L., F. de Bello, K. A. Bråthen, J. Davison, A. Dubuis, A. Guisan et al. 2012.  
503 Ecological assembly rules in plant communities--approaches, patterns and prospects.  
504 *Biological Reviews* 87:111–27.

505 Herben, T. and D. E. Goldberg. 2014. Community assembly by limiting similarity vs.  
506 competitive hierarchies: Testing the consequences of dispersion of individual traits.  
507 *Journal of Ecology* 102:156–166.

508 HilleRisLambers, J., P. B. Adler, W. S. Harpole, J. M. Levine and M. M. Mayfield. 2012.  
509 Rethinking Community Assembly through the Lens of Coexistence Theory. *Annual*  
510 *Review of Ecology, Evolution and Systematics* 43:227–248.

511 Kraft, N. J. B., W. K. Cornwell, C.O. Webb and D. D. Ackerly. 2007. Trait evolution,

512 community assembly, and the phylogenetic structure of ecological communities.  
513 *American Naturalist* 170:271–83.

514 Kraft, N. J. B., O. Godoy, and J. M. Levine. 2015. Plant functional traits and the  
515 multidimensional nature of species coexistence. *Proceedings of the National Academy*  
516 *of Sciences* 112:797– 802.

517 Lawlor, L. R. 1979. Direct and indirect effects of n-species competition. *Oecologia* 43:355-  
518 364.

519 Levine, S. H. 1976. Competitive interactions in ecosystems. *American Naturalist* 110:903-  
520 910.

521 Levine, J. M., J. Bascompte, P.B. Adler and S. Allesina. 2017. Beyond pairwise mechanisms  
522 of species coexistence in complex communities. *Nature* 546: 56-64.

523 Liem, K. F. 1980. Adaptive significance of intra-and interspecific differences in the feeding  
524 repertoires of cichlid fishes. *American zoologist* 20: 295-314.

525 Liem, K. F. 1990. Aquatic versus terrestrial feeding modes: possible impacts on the trophic  
526 ecology of vertebrates. *American Zoologist* 30: 209-221.

527 MacArthur, R. 1970. Species packing and competitive equilibrium for many  
528 species. *Theoretical Population Biology* 1: 1-11.

529 MacArthur, R. H. and R. Levins. 1967. The Limiting Similarity, Convergence, and  
530 Divergence of Coexisting Species. *American Naturalist* 101:377–385.

531 Mayfield, M. M. and J. M. Levine. 2010. Opposing effects of competitive exclusion on the  
532 phylogenetic structure of communities. *Ecology Letters* 13:1085–93.

533 Pianka, E. R. 1974. Niche overlap and diffuse competition. *Proceedings of the National*  
534 *Academy of Sciences* 71:2141-2145.

535 Poff, N. L., M. A. Palmer, P. L. Angermeier, R. L. Vadas, C. C. Hakenkamp, A. Bely, P.  
536 Arensburger, and A. P. Martin. 1993. Size structure of the metazoan community in a

537 Piedmont stream. *Oecologia* 95:202-209.

538 Swenson, N. G. and B. J. Enquist. 2009. Opposing assembly mechanisms in a Neotropical  
539 dry forest: implications for phylogenetic and functional community ecology. *Ecology*  
540 90:2161–2170.

541 Tregonning, K. and A. Roberts. 1979. Complex systems which evolve towards homeostatis.  
542 *Nature* 281:563–564.

543 Warwick, R. M. and K. R. Clarke. 1984. Species size distributions in marine benthic  
544 communities. *Oecologia* 61:32-41.

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

547 Weiher, E. and P. A. Keddy. 1995. Assembly Rules, Null Models, and Trait Dispersion: New  
548 Questions from Old Patterns. *Oikos* 74:159–164.

549 Wilson, W. G. and P. Lundberg. 2004. Biodiversity and the Lotka–Volterra theory of species  
550 interactions: open systems and the distribution of logarithmic densities. *Proceedings*  
551 *of the Royal Society of London B: Biological Sciences* 271:1977-1984.

552

553

554

555

556

557

558

559

560

561



562 **Figures legends**

563 **Figure 1.** The differences between MacArthur's consumer-resource model and our new  
564 model with ecological versatility. (A) In MacArthur's model, species have trait differences  
565 that make their interactions less intense, given that competition happens for a linear array of  
566 resources and each species is able to consume part of this array. (B) In nature, this would be  
567 the case when, for example, birds have different beak sizes and each size enables the  
568 consumption of a range of seed sizes. So, differences in beak sizes always decrease the  
569 interaction strength. (C) In our model, species also have traits that enable them to consume  
570 part of the array of resources, but all species can consume one common resource. (D) This  
571 would be the case when all species of birds can eat a common seed or insect, but not all  
572 species are dependent on these items. Some species have a more distinct beak morphology  
573 that enables them to also eat other items. In this case, the interaction strength and thus niche  
574 overlap are determined by how much a species depends on this common resource.

575

576 **Figure 2.** Typical distributions used for  $k$  and  $\theta$  parameters using beta distributions. The  $k$ s  
577 relate to species fitness and  $\theta$ s relate to species interaction strength. Using the simulated  
578 species we calculated fitness-based distances between any two species  $i$  and  $j$  as  $|\tilde{k}_i - \tilde{k}_j|$  and  
579 niche-based distances as  $1 - \theta_i \theta_j$ . (a) A positive skewed distribution of  $k$ s with most species  
580 having low fitness. (b) A U-shaped distribution of  $k$ s with most species having low or high  
581 fitness and few species having medium fitness. (c) A for bell-shaped distribution of  $k$ s with  
582 most species having medium fitness. (d) A negative skewed distribution of  $k$ s with most  
583 species having high fitness. (e) A negative skewed distribution of  $\theta$ s resulting in most species  
584 having low niche differences between them. (f) A bell-shaped distributions of  $\theta$ s resulting in  
585 most species having medium niche differences between them. (g) A uniform distribution of  
586  $\theta$ s resulting in most of species having high niche differences between them.

587

588 **Figure 3.** Variation in species richness, Jaccard dissimilarity and species dispersion metrics  
589 when simulating community assembly using different species pools. Blue boxplots are results  
590 for species pools with average low niche differences, red boxplots for medium niche  
591 differences and yellow for high niche differences. Within each color, the first boxplot is when  
592 species have low fitness at the species pool scale, the second when species have medium  
593 fitness values, the third when species have high fitness values and the fourth when species  
594 have a U-shaped distribution. We first applied indices SES-MNTD and SES-MPD separately  
595 to niche distances and fitness distances between species and then to the overall functional  
596 distances which combines niche and fitness as defined in the main text.

597

598 **Figure 4.** Result of species dispersion metrics when species distance is calculated using  
599 different weights for niche and fitness differences. Community assembly was performed  
600 using a species pool with medium niche differences and medium fitness values at the species  
601 pool scale.

602

603 **Figure 5.** Results of richness, Jaccard dissimilarity and species dispersion metrics when we  
604 varied the number of sequential colonization events during community assembly. Community  
605 assembly was performed using a species pool with medium niche differences and fitness  
606 values at the species pool scale. We first applied indices SES-MNTD and SES-MPD  
607 separately to niche distances and fitness distances between species and then to the overall  
608 functional distances which combines niche and fitness as defined in the main text.

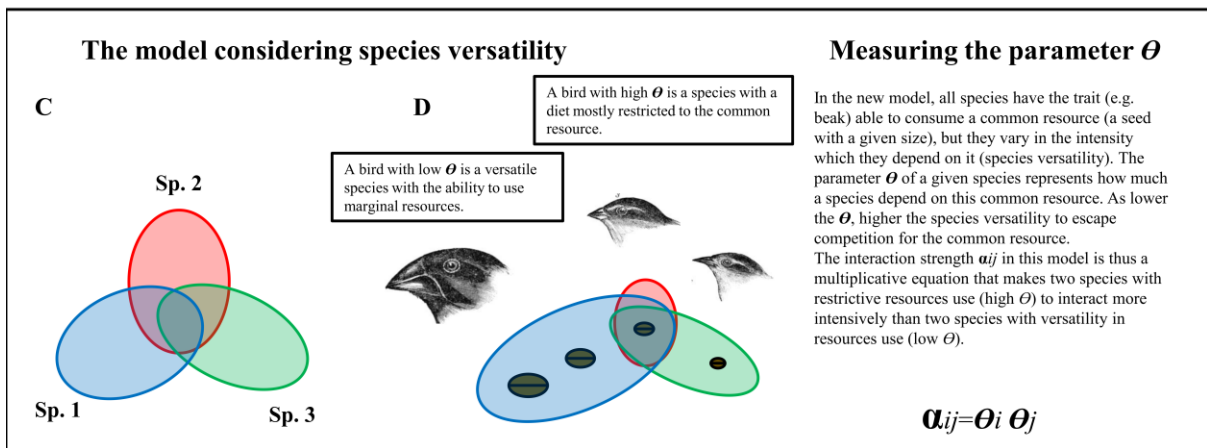
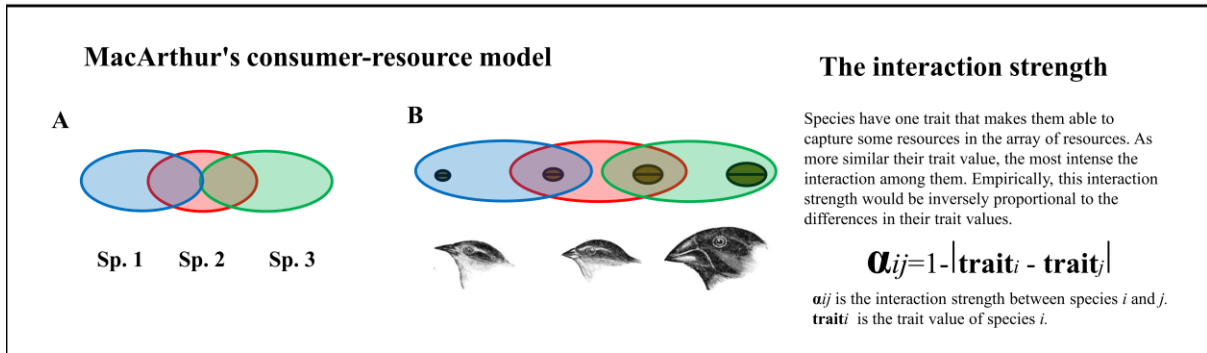
609

610

611

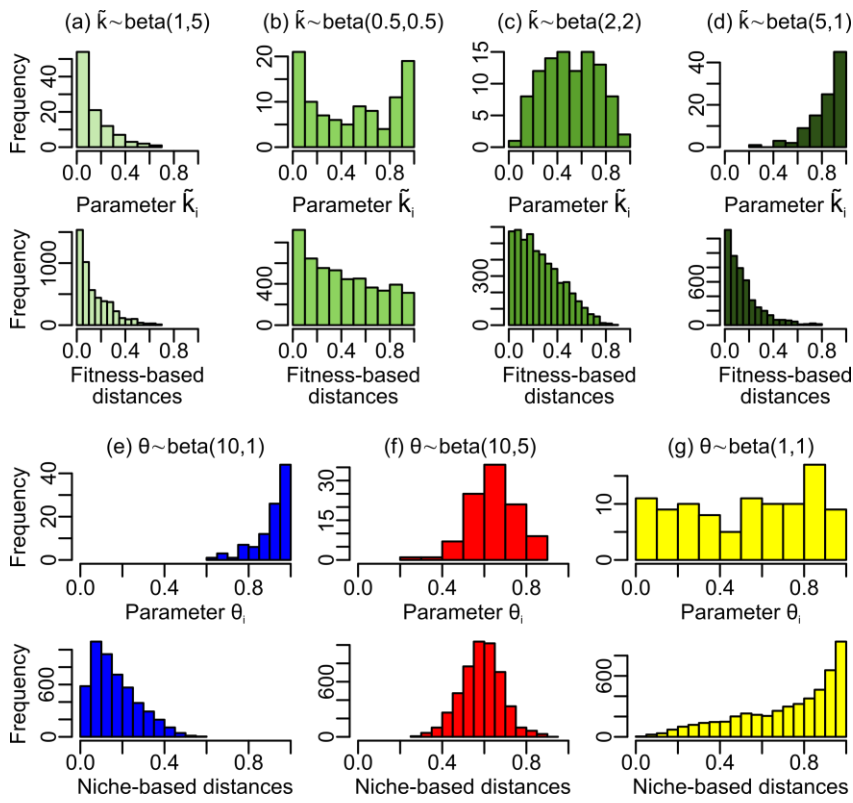
612 **Figures**

613 Figure 1.



614  
 615  
 616  
 617  
 618  
 619  
 620  
 621  
 622  
 623  
 624  
 625  
 626  
 627  
 628  
 629

630 Figure 2.



631

632

633

634

635

636

637

638

639

640

641

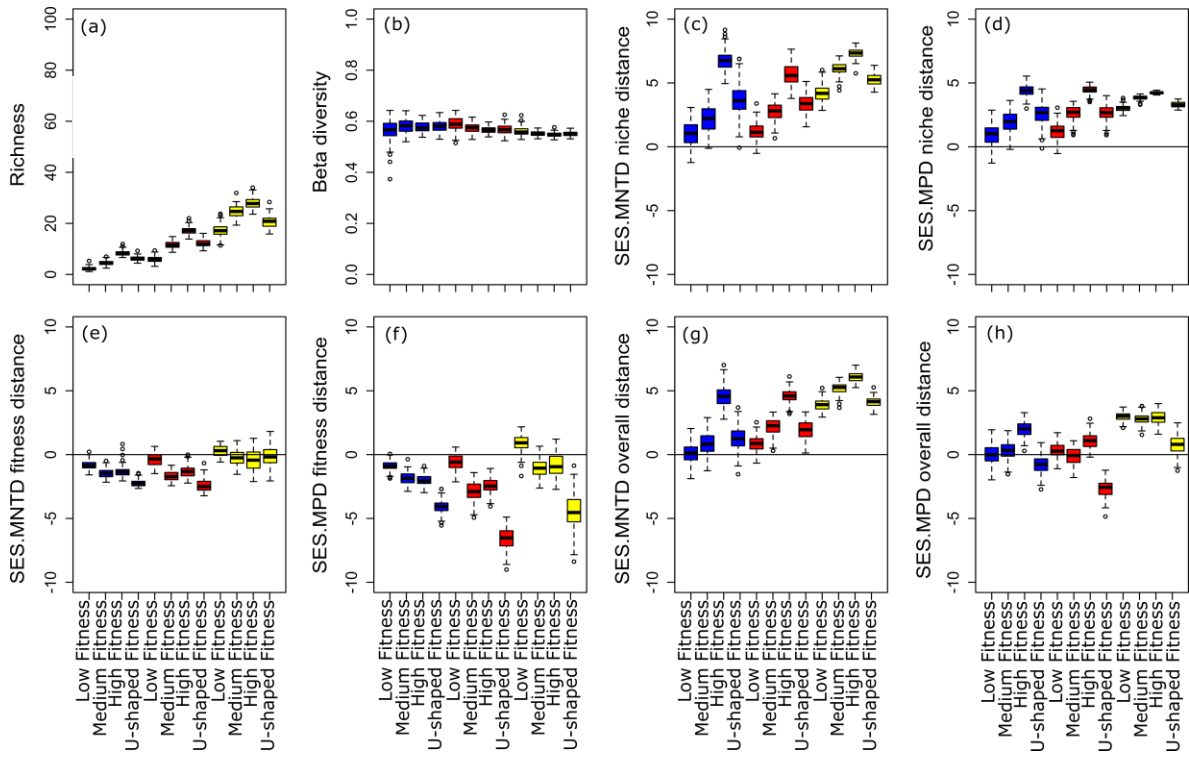
642

643

644

645

646 Figure 3.



647

648

649

650

651

652

653

654

655

656

657

658

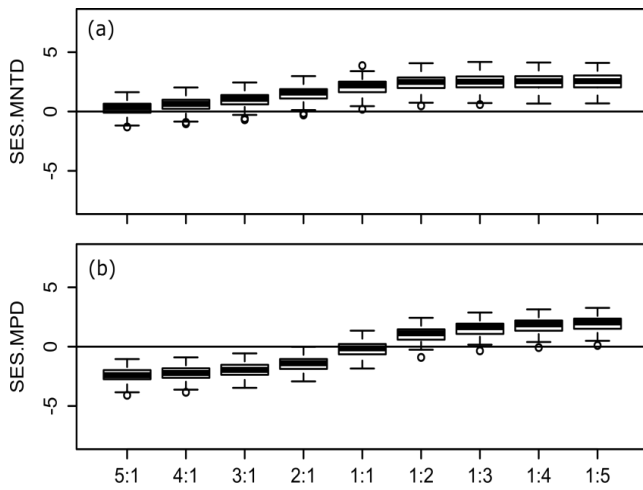
659

660

661

662

663 Figure 4.



Weights of  
fitness:niche  
distances in  
overall species  
distance  
calculation

664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678

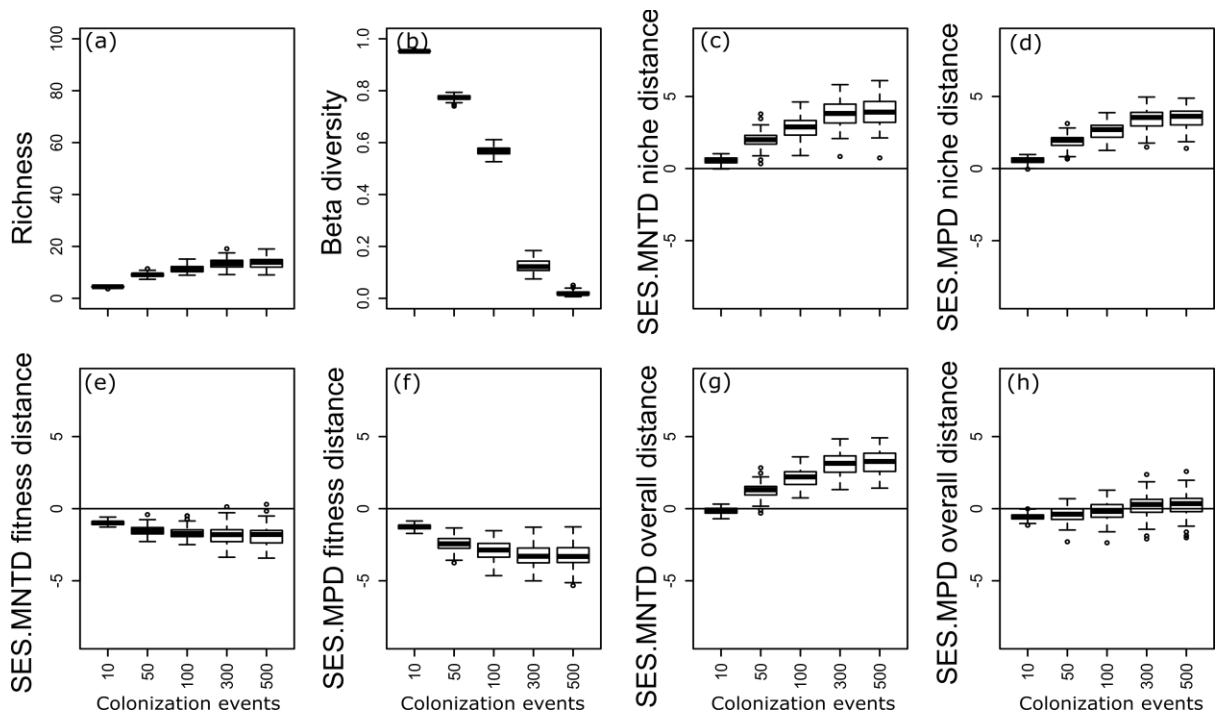
679

680

681

682

683 Figure 5.



684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

## 699 **Appendix S1 Manual to use functions of community assembly included in DataS1**

700

701

### 702 **## Usage**

703

704 `scenario_1_multispecies(a, t1, t2, f1, f2, nsim, tol=1e-8)`

705 `scenario_2_multispecies(a, t1, t2, f1, f2, nsim, tol=1e-8)`

706 `scenario_3_multispecies(nsim, t1, t2, f1, f2, tol=1e-8)`

707

708 • `a` is the number of invasions (a numeric, integer);

709 • `nsim` is the number of times the simulations are repeated: from the definition of the species pool to the calculation of the indices SES-MPD and SES-MNTD (a numeric, integer);

710 • `t1, t2, f1, f2`, are the parameters, bounded between 0 and 1, and used for beta distribution. `t1` and `t2` are respectively the  $\alpha$  and  $\beta$  parameters of the beta distribution used for  $\theta$ 's values. `f1` and `f2` are respectively the  $\alpha$  and  $\beta$  parameters of the beta distribution used for  $\tilde{k}_i$ 's values. We considered various distributions for  $\tilde{k}_i$ 's and  $\theta$ 's values as indicated in the main text..

711 • `tol` is a tolerance threshold for null values: a value lower than `tol` is considered as null (a numeric)

712

### 713 **Value**

714 In function `scenario_1_multispecies` we varied the relative weight given to niche versus fitness traits in overall functional distances when calculating SES.MNTD and SES.MNTD (Figure 3 of the main text).

715 In function `scenario_2_multispecies` we calculated all diversity indices for a fixed number of colonization events: richness; beta diversity; and SES.MNTD and SES.MNTD, applied to niche distances, fitness distances and overall functional distances with even weight given to niche versus fitness traits (Figure 2 of the main text). In function `scenario_3_multispecies` we also calculated all diversity indices but varying the number of colonization events (Figure 4 of the main text).

716 All functions end up with a list. Each value of the list contains the result of a simulation, where simulation means that a pool of species has been obtained, 100 communities have been obtained by colonization from the species pool as described in the main text and metrics have been calculated on each community (species richness, beta diversity, and SES MNTD, SES MPD using niche traits, fitness or both) and averaged over the 100 communities. A simulation thus corresponds to a particular species pool.

717 All functions calculate all metrics (species richness, beta diversity, and SES MNTD, SES MPD using niche traits, fitness or both) except `scenario_1_multispecies` where only SES MNTD, SES MPD on both niche and fitness are calculated.

718 When calculating averaged SES MNTD and averaged SES MPD over the 100 simulated communities, communities, if any, with a single species were discarded.

719

### 720 **Example of applications (to obtain similar figures as in the main text)**

721 Comments are in bold letters following #.

722

### 723 **## Scenario 1**

724



```

743 # Example of application with a distribution beta(10, 5) for  $\theta$ 's values and beta(2,2) for  $\tilde{k}_i$ 's.
744 This corresponds to bell-shaped distributions for interaction strength and for fitness.
745 scenario_1_res_10_5_2_2 <- scenario_1_multispecies(a=100, 10,5,2,2,
746 nsim=100)
747
748 # The result of the first simulation scheme (from the definition of a unique species pool to the
749 simulations of the 100 communities) can be obtained with the following instruction:
750 scenario_1_res_10_5_2_2[[1]]
751 $SES.MNTD
752 niche1.fitness5 niche1.fitness4 niche1.fitness3 niche1.fitness2
753     -1.0992815     -0.9177786     -0.6943723     -0.1880420
754 niche1.fitness1 niche2.fitness1 niche3.fitness1 niche4.fitness1
755     0.4595798     0.8757695     1.0033666     1.1138923
756 niche5.fitness1
757     1.2033374
758
759 $SES.MPD
760 niche1.fitness5 niche1.fitness4 niche1.fitness3 niche1.fitness2
761     -2.0664446     -1.9937556     -1.8735802     -1.4928701
762 niche1.fitness1 niche2.fitness1 niche3.fitness1 niche4.fitness1
763     -0.7268908     0.2662877     0.7062460     0.9866080
764 niche5.fitness1
765     1.1153460
766
767 # scenario_1_res_10_5_2_2[[1]] thus is a list of two vectors: the first one for metric MNTD, the
768 second one for metric MPD. In each of the vector, the first value is the average metric value
769 (either SES MNTD or SES MPD) over the 100 simulated communities when the fitness was
770 given 5 times more weight than niche when calculating functional dissimilarities among
771 species; the second value is the average metric value (either SES MNTD or SES MPD) over the
772 100 simulated communities when the fitness was given 4 times more weight than niche when
773 calculating functional dissimilarities among species, etc. "niche4.fitness1" thus for example
774 means 4 times more weight to niche than to fitness.
775
776 # scenario_1_res_10_5_2_2[[i]] and scenario_1_res_10_5_2_2[[j]] for any i different from j
777 correspond to distinct species pools.
778
779 # The following scripts led to Figure 4 of the main text
780
781 scenario_1_SES.MNTD_niche1.fitness5 <-
782 unlist(lapply(scenario_1_res_10_5_2_2, function(x)
783 x$SES.MNTD["niche1.fitness5"]))
784 scenario_1_SES.MNTD_niche1.fitness4 <-
785 unlist(lapply(scenario_1_res_10_5_2_2, function(x)
786 x$SES.MNTD["niche1.fitness4"]))
787 scenario_1_SES.MNTD_niche1.fitness3 <-
788 unlist(lapply(scenario_1_res_10_5_2_2, function(x)
789 x$SES.MNTD["niche1.fitness3"]))
790 scenario_1_SES.MNTD_niche1.fitness2 <-
791 unlist(lapply(scenario_1_res_10_5_2_2, function(x)
792 x$SES.MNTD["niche1.fitness2"]))
793 scenario_1_SES.MNTD_niche1.fitness1 <-
794 unlist(lapply(scenario_1_res_10_5_2_2, function(x)
795 x$SES.MNTD["niche1.fitness1"]))

```

```

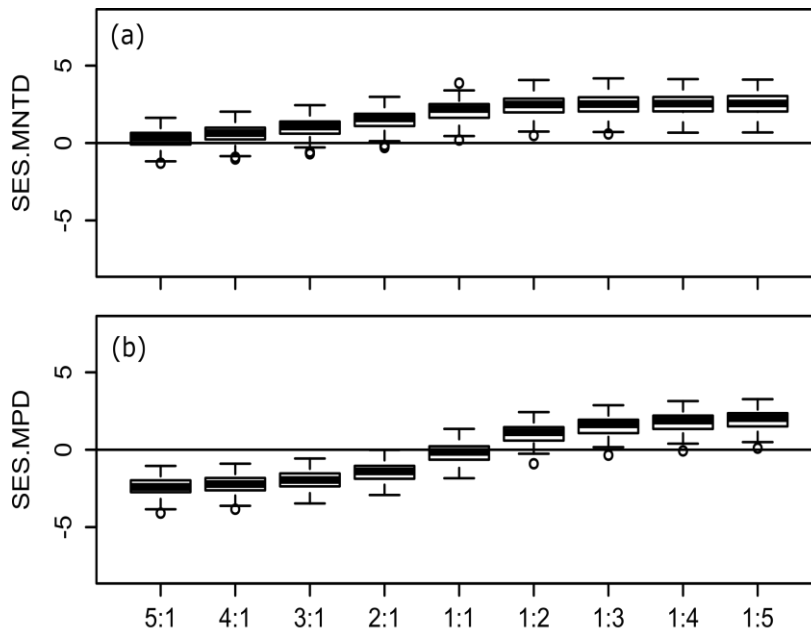
796 scenario_1_SES.MNTD_niche2.fitness1 <-
797 unlist(lapply(scenario_1_res_10_5_2_2, function(x)
798 x$SES.MNTD["niche2.fitness1"]))
799 scenario_1_SES.MNTD_niche3.fitness1 <-
800 unlist(lapply(scenario_1_res_10_5_2_2, function(x)
801 x$SES.MNTD["niche3.fitness1"]))
802 scenario_1_SES.MNTD_niche4.fitness1 <-
803 unlist(lapply(scenario_1_res_10_5_2_2, function(x)
804 x$SES.MNTD["niche4.fitness1"]))
805 scenario_1_SES.MNTD_niche5.fitness1 <-
806 unlist(lapply(scenario_1_res_10_5_2_2, function(x)
807 x$SES.MNTD["niche5.fitness1"]))
808
809 scenario_1_MNTD <- list(scenario_1_SES.MNTD_niche1.fitness5,
810 scenario_1_SES.MNTD_niche1.fitness4, scenario_1_SES.MNTD_niche1.fitness3,
811 scenario_1_SES.MNTD_niche1.fitness2, scenario_1_SES.MNTD_niche1.fitness1,
812 scenario_1_SES.MNTD_niche2.fitness1, scenario_1_SES.MNTD_niche3.fitness1,
813 scenario_1_SES.MNTD_niche4.fitness1, scenario_1_SES.MNTD_niche5.fitness1)
814
815
816 scenario_1_SES.MPD_niche1.fitness5 <-
817 unlist(lapply(scenario_1_res_10_5_2_2, function(x)
818 x$SES.MPD["niche1.fitness5"]))
819 scenario_1_SES.MPD_niche1.fitness4 <-
820 unlist(lapply(scenario_1_res_10_5_2_2, function(x)
821 x$SES.MPD["niche1.fitness4"]))
822 scenario_1_SES.MPD_niche1.fitness3 <-
823 unlist(lapply(scenario_1_res_10_5_2_2, function(x)
824 x$SES.MPD["niche1.fitness3"]))
825 scenario_1_SES.MPD_niche1.fitness2 <-
826 unlist(lapply(scenario_1_res_10_5_2_2, function(x)
827 x$SES.MPD["niche1.fitness2"]))
828 scenario_1_SES.MPD_niche1.fitness1 <-
829 unlist(lapply(scenario_1_res_10_5_2_2, function(x)
830 x$SES.MPD["niche1.fitness1"]))
831 scenario_1_SES.MPD_niche2.fitness1 <-
832 unlist(lapply(scenario_1_res_10_5_2_2, function(x)
833 x$SES.MPD["niche2.fitness1"]))
834 scenario_1_SES.MPD_niche3.fitness1 <-
835 unlist(lapply(scenario_1_res_10_5_2_2, function(x)
836 x$SES.MPD["niche3.fitness1"]))
837 scenario_1_SES.MPD_niche4.fitness1 <-
838 unlist(lapply(scenario_1_res_10_5_2_2, function(x)
839 x$SES.MPD["niche4.fitness1"]))
840 scenario_1_SES.MPD_niche5.fitness1 <-
841 unlist(lapply(scenario_1_res_10_5_2_2, function(x)
842 x$SES.MPD["niche5.fitness1"]))
843
844 scenario_1_MPD <- list(scenario_1_SES.MPD_niche1.fitness5,
845 scenario_1_SES.MPD_niche1.fitness4, scenario_1_SES.MPD_niche1.fitness3,
846 scenario_1_SES.MPD_niche1.fitness2, scenario_1_SES.MPD_niche1.fitness1,
847 scenario_1_SES.MPD_niche2.fitness1, scenario_1_SES.MPD_niche3.fitness1,
848 scenario_1_SES.MPD_niche4.fitness1, scenario_1_SES.MPD_niche5.fitness1)

```

```

849
850 par(mfrow=c(2,1))
851 boxplot(scenario_1_MNTD, ylim=c(-8,8), ylab="SES.MNTD", xlab="Nb of
852 fitness:niche traits",
853 names=c("5:1","4:1","3:1","2:1","1:1","1:2","1:3","1:4","1:5"))
854 abline(h=0)
855
856 boxplot(scenario_1_MPD, ylim=c(-8,8), ylab="SES.MPD", xlab="Nb of
857 fitness:niche traits",
858 names=c("5:1","4:1","3:1","2:1","1:1","1:2","1:3","1:4","1:5"))
859 abline(h=0)
860

```



Weights of fitness:niche  
distances in overall  
species distance  
calculation

861  
862 **Duplication of Figure 4 of the main text.** Result of species dispersion metrics when species  
863 distance is calculated using different weights for niche and fitness differences. Community  
864 assembly was performed using a species pool with medium niche differences and medium  
865 fitness values at the species pool scale.

866  
867  
868 **## Scenario 2**

869  
870 **# We considered 3 sets of parameters for the beta distributions of the  $\theta_i$ 's values and 4 for  
871 the beta distributions of the  $\tilde{k}_i$ 's as specified in the main text.**

```

872
873 scenario_2_res_10_1_1_5 <- scenario_2_multispecies(a=100, 10,1,1,5,
874 nsim=100)

```

```

875 scenario_2_res_10_1_05_05 <- scenario_2_multispecies(a=100, 10,1,0.5,0.5,
876 nsim=100)
877 scenario_2_res_10_1_2_2 <- scenario_2_multispecies(a=100, 10,1,2,2,
878 nsim=100)
879 scenario_2_res_10_1_5_1 <- scenario_2_multispecies(a=100, 10,1,5,1,
880 nsim=100)
881
882
883 scenario_2_res_10_5_1_5 <- scenario_2_multispecies(a=100, 10,5,1,5,
884 nsim=100)
885 scenario_2_res_10_5_05_05 <- scenario_2_multispecies(a=100, 10,5,0.5,0.5,
886 nsim=100)
887 scenario_2_res_10_5_2_2 <- scenario_2_multispecies(a=100, 10,5,2,2,
888 nsim=100)
889 scenario_2_res_10_5_5_1 <- scenario_2_multispecies(a=100, 10,5,5,1,
890 nsim=100)
891
892 scenario_2_res_1_1_1_5 <- scenario_2_multispecies(a=100, 1,1,1,5,
893 nsim=100)
894 scenario_2_res_1_1_05_05 <- scenario_2_multispecies(a=100, 1,1,0.5,0.5,
895 nsim=100)
896 scenario_2_res_1_1_2_2 <- scenario_2_multispecies(a=100, 1,1,2,2,
897 nsim=100)
898 scenario_2_res_1_1_5_1 <- scenario_2_multispecies(a=100, 1,1,5,1,
899 nsim=100)
900
901 # For example, with a distribution beta(10, 5) for  $\theta$ 's values and beta(2,2) for  $\tilde{k}_i$ 's, the result of
902 the first simulation schemes (from the definition of a unique species pool to the simulations of
903 the 100 communities, with 100% of niche range) can be obtained with the following
904 instruction:
905
906 scenario_2_res_10_5_2_2[[1]]
907      richness      beta  SES.MNTDniche  SES.MPDniche  SES.MNTDfitness
908 13.2400000  0.5992819  2.2651896  2.7483701  -1.3339585
909  SES.MPDfitness  SES.MNTDoverall  SES.MPDoverall
910  -1.9873402  2.4475813  0.5512119
911
912 # scenario_2_res_10_5_2_2[[1]] is thus a vector of eight values: the first gives the average
913 richness in the 100 simulated communities, the second one the average beta diversity among
914 the 100 communities (pairwise Jaccard dissimilarity), then the other values provide indices
915 SES MNTD and SES MPD as specified calculated with the niche trait, the fitness or both. For
916 example, "SES.MPDoverall" gives the SES MPD value applied to both niche and fitness and
917 averaged over the 100 simulated communities, while "SES.MNTDniche" gives the SES MNTD
918 value applied to niche only and averaged over the 100 simulated communities.
919
920 # scenario_2_res_10_5_2_2 [[i]] and scenario_2_res_10_5_2_2 [[j]] for any i different from j
921 correspond to distinct species pools.
922
923 # The following scripts led to Figure 3 of the main text
924
925 vecD =c("10_1_1_5", "10_1_05_05", "10_1_2_2", "10_1_5_1",
926 "10_5_1_5", "10_5_05_05", "10_5_2_2", "10_5_5_1",
927 "1_1_1_5", "1_1_05_05", "1_1_2_2", "1_1_5_1")

```

```

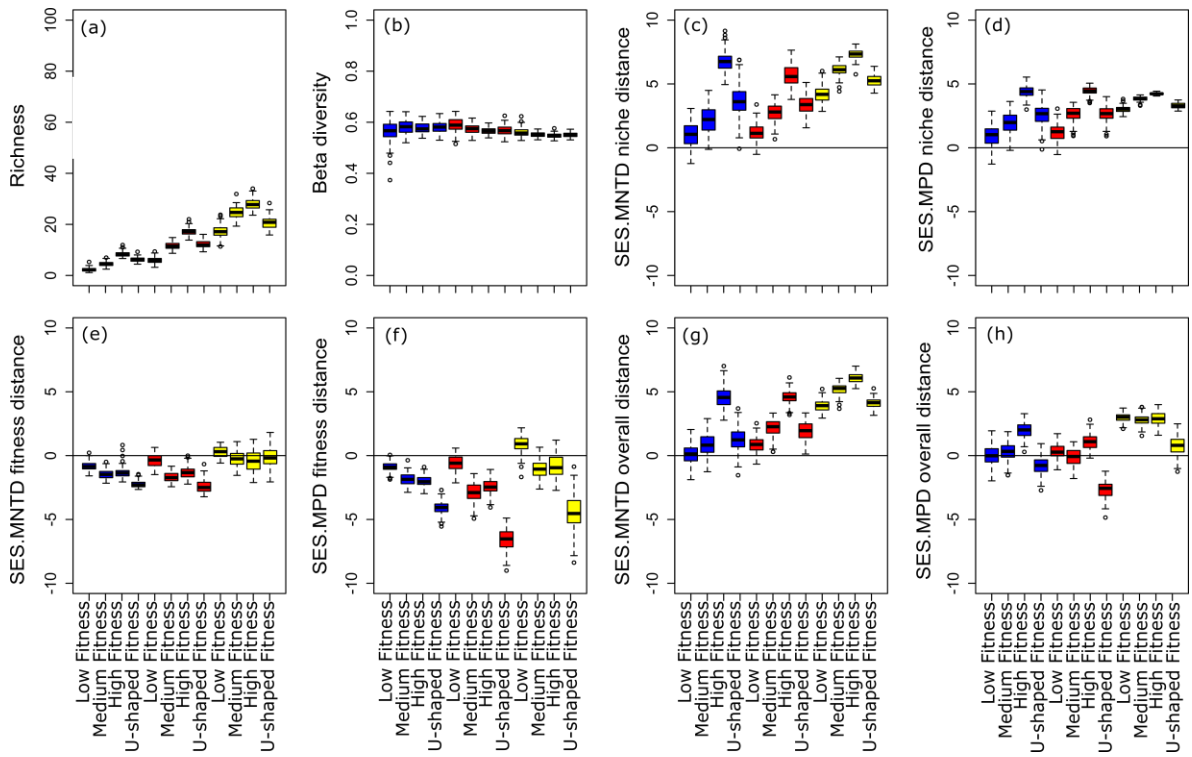
928
929 scenario_2_res <- list(
930 scenario_2_res_10_1_1_5, scenario_2_res_10_1_05_05,
931 scenario_2_res_10_1_2_2, scenario_2_res_10_1_5_1,
932 scenario_2_res_10_5_1_5, scenario_2_res_10_5_05_05,
933 scenario_2_res_10_5_2_2, scenario_2_res_10_5_5_1,
934 scenario_2_res_1_1_1_5, scenario_2_res_1_1_05_05, scenario_2_res_1_1_2_2,
935 scenario_2_res_1_1_5_1)
936
937 scenario_2_MNTDniche <- lapply(scenario_2_res, function(v)
938 as.vector(unlist(lapply(v, function(x) x["SES.MNTDniche"]))))
939
940 scenario_2_MNTDfitness <- lapply(scenario_2_res, function(v)
941 as.vector(unlist(lapply(v, function(x) x["SES.MNTDfitness"]))))
942
943 scenario_2_MNTDoverall <- lapply(scenario_2_res, function(v)
944 as.vector(unlist(lapply(v, function(x) x["SES.MNTDoverall"]))))
945
946 scenario_2_MPDniche <- lapply(scenario_2_res, function(v)
947 as.vector(unlist(lapply(v, function(x) x["SES.MPDniche"]))))
948
949 scenario_2_MPDfitness <- lapply(scenario_2_res, function(v)
950 as.vector(unlist(lapply(v, function(x) x["SES.MPDfitness"]))))
951
952 scenario_2_MPDOverall <- lapply(scenario_2_res, function(v)
953 as.vector(unlist(lapply(v, function(x) x["SES.MPDOverall"]))))
954
955 scenario_2_richness <- lapply(scenario_2_res, function(v)
956 as.vector(unlist(lapply(v, function(x) x["richness"]))))
957
958 scenario_2_beta <- lapply(scenario_2_res, function(v)
959 as.vector(unlist(lapply(v, function(x) x["beta"]))))
960
961 par(mfrow=c(2,4))
962 boxplot(scenario_2_richness, ylim=c(0,100), names=vecD, las=3,
963 ylab="Richness", col=c(rep("blue", 4), rep("red", 4), rep("yellow", 4)))
964 boxplot(scenario_2_beta, ylim=c(0,1), names=vecD, las=3, ylab="Beta
965 diversity", col=c(rep("blue", 4), rep("red", 4), rep("yellow", 4)))
966 boxplot(scenario_2_MNTDniche, ylim=c(-10,10), names=vecD, las=3,
967 ylab="SES.MNTD niche distance", col=c(rep("blue", 4), rep("red", 4),
968 rep("yellow", 4)))
969 abline(h=0)
970 boxplot(scenario_2_MPDniche, ylim=c(-10,10), names=vecD, las=3,
971 ylab="SES.MPD niche distance", col=c(rep("blue", 4), rep("red", 4),
972 rep("yellow", 4)))
973 abline(h=0)
974 boxplot(scenario_2_MNTDfitness, ylim=c(-10,10), names=vecD, las=3,
975 ylab="SES.MNTD fitness distance", col=c(rep("blue", 4), rep("red", 4),
976 rep("yellow", 4)))
977 abline(h=0)
978 boxplot(scenario_2_MPDfitness, ylim=c(-10,10), names=vecD, las=3,
979 ylab="SES.MPD fitness distance", col=c(rep("blue", 4), rep("red", 4),
980 rep("yellow", 4)))
981 abline(h=0)

```

```

982 boxplot(scenario_2_MNTDoverall, ylim=c(-10,10), names=vecD, las=3,
983 ylab="SES.MNTD overall distance", col=c(rep("blue", 4), rep("red", 4),
984 rep("yellow", 4)))
985 abline(h=0)
986 boxplot(scenario_2_MPDoverall, ylim=c(-10,10), names=vecD, las=3,
987 ylab="SES.MPD overall distance", col=c(rep("blue", 4), rep("red", 4),
988 rep("yellow", 4)))
989 abline(h=0)
990

```



991

992 **Duplication of Figure 3 of the main text.** Variation in species richness, Jaccard dissimilarity  
993 and species dispersion metrics when simulating community assembly using different species  
994 pools. Blue boxplots are results for species pools with average low niche differences, red  
995 boxplots for medium niche differences and yellow for high niche differences. Within each  
996 color, the first boxplot is when species have low fitness at the species pool scale, the second  
997 when species have medium fitness values, the third when species have high fitness values and  
998 the fourth when species have a U-shaped distribution. We first applied indices SES-MNTD  
999 and SES-MPD separately to niche distances and fitness distances between species and then to  
1000 the overall functional distances which combines niche and fitness as defined in the main text.

1001 **Legend:**

- 1002 Blue= Niche differences are high
- 1003 Red=Niche differences are medium
- 1004 Yellow=Niche differences are low

1005 For a given color, left=positive skewed fitness values (most species have low fitness, a few species  
 1006 have high fitness); middle left=U shaped distributed fitness values (many low and high fitness, less  
 1007 medium fitness values); middle left=Bell-shaped distributed fitness values (many medium fitness  
 1008 values); right=negative skewed fitness values (most species have high fitness, a few species have low  
 1009 fitness);

1010 **## Scenario 3**

1011  
 1012 **# Example of application with a distribution beta(10, 5) for  $\theta_i$ 's values and beta(2,2) for  $\tilde{k}_i$ 's.**  
 1013 **This corresponds to bell-shaped distributions for interaction strength and for fitness.**

```
1014  

  1015 scenario_3_res_10_5_2_2 <- scenario_3_multispecies(nsim=100, 10,5,2,2)
```

1016  
 1017 **# The result of the first simulation can be obtained with the following instruction:**

```
1018  

  1019 scenario_3_res_10_5_2_2[[1]]  

  1020 $richness  

  1021     10     50    100    300    500  

  1022  4.46 10.43 12.87 14.74 14.91  

  1023  

  1024 $beta  

  1025         10         50         100         300         500  

  1026 0.95226597 0.77407196 0.57021583 0.11711865 0.02511381  

  1027  

  1028 $SES.MNTDniche  

  1029         10         50         100         300         500  

  1030 0.3396579 1.4632582 2.4961132 3.8562099 3.6711381  

  1031  

  1032 $SES.MPDniche  

  1033         10         50         100         300         500  

  1034 0.4070142 1.6855537 2.5470740 3.0385411 3.1551790  

  1035  

  1036 $SES.MNTDfitness  

  1037         10         50         100         300         500  

  1038 -1.007009 -1.894202 -2.162191 -2.399947 -2.419053  

  1039  

  1040 $SES.MPDfitness  

  1041         10         50         100         300         500  

  1042 -1.289755 -2.456777 -2.815547 -3.365079 -3.337911  

  1043  

  1044 $SES.MNTDoverall  

  1045         10         50         100         300         500  

  1046 -0.4096832 0.9330254 1.9980024 2.7822315 2.6177442  

  1047  

  1048 $SES.MPDoverall  

  1049         10         50         100         300         500  

  1050 -0.7978737 -0.8030256 -0.4172586 -0.4110750 -0.5757312
```

1051  
 1052 **# scenario\_3\_res\_10\_5\_2\_2[[1]] is thus a list of eight vectors: the first gives the average**  
 1053 **richness in the 100 simulated communities, the second one the average beta diversity among**  
 1054 **the 100 communities (pairwise Jaccard dissimilarity), then the other vectors provide indices**  
 1055 **SES MNTD and SES MPD as specified calculated with the niche trait, the fitness or both. For**

1056 **example, \$SES.MPDoverall gives the SES MPD value applied to both niche and fitness and**  
1057 **averaged over the 100 simulated communities, while \$SES.MNTDniche gives the SES MNTD**  
1058 **value applied to niche only and averaged over the 100 simulated communities. The names**  
1059 **attributed above each value of a vector indicate the numbers of colonization events used in**  
1060 **the model.**

1061  
1062 **# scenario\_3\_res\_10\_5\_2\_2 [[i]] and scenario\_3\_res\_10\_5\_2\_2 [[j]] for any i different from j**  
1063 **correspond to distinct species pools.**

1064  
1065 **# The following scripts led to Figure 5 of the main text**

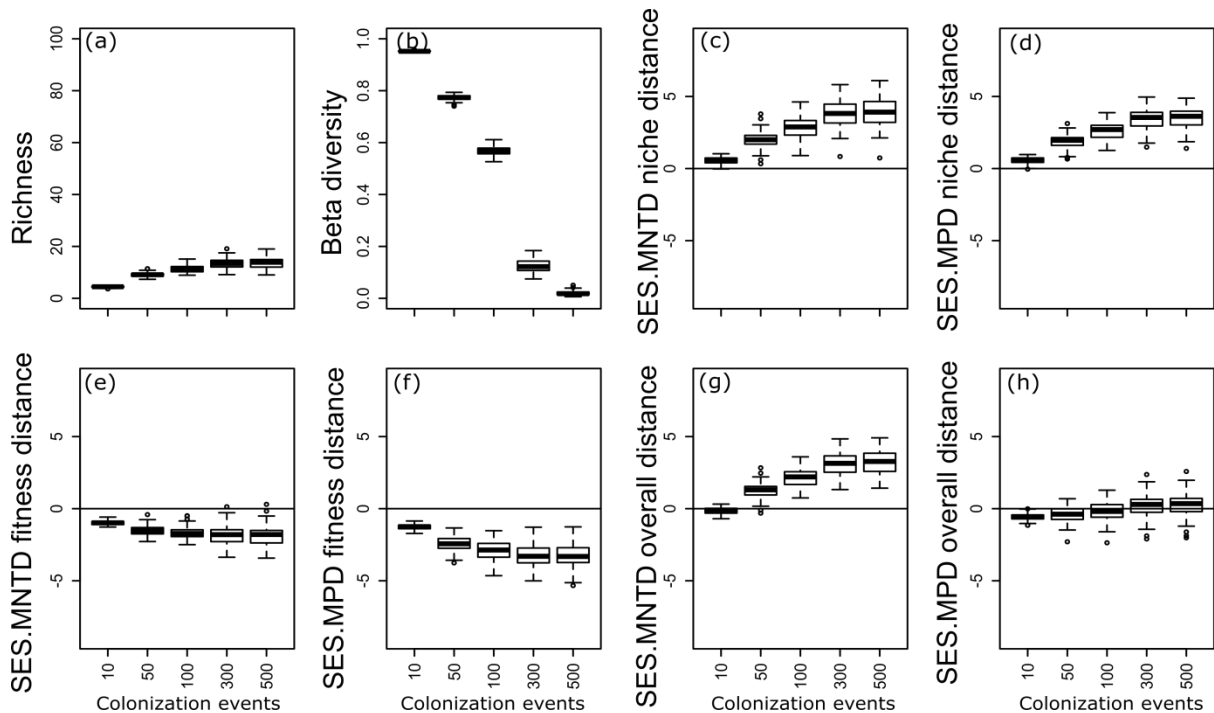
```
1066  
1067 vecC = as.character(c(10, 50, 100, 300, 500))  
1068  
1069 scenario_3_MNTDniche_10_5_2_2 <- lapply(vecC, function(v)  
1070 unlist(lapply(scenario_3_res_10_5_2_2, function(x)  
1071 x$SES.MNTDniche[v])))  
1072  
1073 scenario_3_MNTDfitness_10_5_2_2 <- lapply(vecC, function(v)  
1074 unlist(lapply(scenario_3_res_10_5_2_2, function(x)  
1075 x$SES.MNTDfitness[v])))  
1076  
1077 scenario_3_MNTDoverall_10_5_2_2 <- lapply(vecC, function(v)  
1078 unlist(lapply(scenario_3_res_10_5_2_2, function(x)  
1079 x$SES.MNTDoverall[v])))  
1080  
1081 scenario_3_MPDniche_10_5_2_2 <- lapply(vecC, function(v)  
1082 unlist(lapply(scenario_3_res_10_5_2_2, function(x)  
1083 x$SES.MPDniche[v])))  
1084  
1085 scenario_3_MPDfitness_10_5_2_2 <- lapply(vecC, function(v)  
1086 unlist(lapply(scenario_3_res_10_5_2_2, function(x)  
1087 x$SES.MPDfitness[v])))  
1088  
1089 scenario_3_MPDoverall_10_5_2_2 <- lapply(vecC, function(v)  
1090 unlist(lapply(scenario_3_res_10_5_2_2, function(x)  
1091 x$SES.MPDoverall[v])))  
1092  
1093 scenario_3_richness_10_5_2_2 <- lapply(vecC, function(v)  
1094 unlist(lapply(scenario_3_res_10_5_2_2, function(x) x$richness[v])))  
1095  
1096 scenario_3_beta_10_5_2_2 <- lapply(vecC, function(v)  
1097 unlist(lapply(scenario_3_res_10_5_2_2, function(x) x$beta[v])))  
1098  
1099  
1100  
1101 par(mfrow=c(2,4))  
1102 boxplot(scenario_3_richness_10_5_2_2, ylim=c(0,100), names=vecC,  
1103 las=3, xlab="Colonization intensity", ylab="Richness")  
1104 boxplot(scenario_3_beta_10_5_2_2, ylim=c(0,1), names=vecC, las=3,  
1105 xlab="Colonization intensity", ylab="Beta diversity")  
1106 boxplot(scenario_3_MNTDniche_10_5_2_2, ylim=c(-9,9), names=vecC,  
1107 las=3, xlab="Colonization intensity", ylab="SES.MNTD niche distance")  
1108 abline(h=0)  
1109 boxplot(scenario_3_MPDniche_10_5_2_2, ylim=c(-9,9), names=vecC, las=3,  
1110 xlab="Colonization intensity", ylab="SES.MPD niche distance")
```



```

1111 abline(h=0)
1112 boxplot(scenario_3_MNTDfitness_10_5_2_2, ylim=c(-9,9), names=vecC,
1113 las=3, xlab="Colonization intensity", ylab="SES.MNTD fitness distance")
1114 abline(h=0)
1115 boxplot(scenario_3_MPDfitness_10_5_2_2, ylim=c(-9,9), names=vecC,
1116 las=3, xlab="Colonization intensity", ylab="SES.MPD fitness distance")
1117 abline(h=0)
1118 boxplot(scenario_3_MNTDoverall_10_5_2_2, ylim=c(-9,9), names=vecC,
1119 las=3, xlab="Colonization intensity", ylab="SES.MNTD overall distance")
1120 abline(h=0)
1121 boxplot(scenario_3_MPDoverall_10_5_2_2, ylim=c(-9,9), names=vecC,
1122 las=3, xlab="Colonization intensity", ylab="SES.MPD overall distance")
1123 abline(h=0)
1124

```



1125  
1126 **Duplication of figure 5 of the main text.** Results of richness, Jaccard dissimilarity and  
1127 species dispersion metrics when we varied the number of sequential colonization events  
1128 during community assembly. Community assembly was performed using a species pool with  
1129 medium niche differences and fitness values at the species pool scale. We first applied  
1130 indices SES-MNTD and SES-MPD separately to niche distances and fitness distances  
1131 between species and then to the overall functional distances which combines niche and  
1132 fitness as defined in the main text.

1133  
1134

1135

1136

1137  
1138  
1139  
1140  
1141

1142 **Appendix S2 Deriving an equalizing/stabilizing additive decomposition in a multi-**  
1143 **species competition model starting with consumer-resource model (MacArthur 1970)**

1144 **In this appendix, we first provide the detailed demonstration of the method used in the**  
1145 **main text. Second, we justify the form of the interaction coefficients which is considered**  
1146 **in the model by relating it with a general resource competition model.**

1147 **Deriving the low-density long term growth rate and its stabilizing and equalizing**  
1148 **components in a multispecies competition model with interaction coefficients such as**  
1149  **$\alpha_{ij} = \theta_i \theta_j$**

1150 For our study, we searched a reference model that respects Chesson's (2000) following  
1151 approximation to the long-term low-density growth rate (equation 4 in Chesson 2000):

1152 
$$\bar{r}_i \approx b_i \left( k_i - \frac{\sum_{s \neq i} k_s}{n-1} \right) + b_i \frac{(1-\rho)}{n-1} D$$

1153 where  $n$  is the number of species in the system, the  $k_s$  are measures of fitness of individual  
1154 species,  $(\sum_{s \neq i} k_s)/(n-1)$  is the average fitness of the competitors of  $i$ ,  $\rho$  is a measure of  
1155 niche overlap, and  $D$  is a positive constant. This equation is composed of an equalizing term

$$b_i \left( k_i - \frac{\sum_{s \neq i} k_s}{n-1} \right)$$

1156 and a stabilizing term

1157 
$$b_i \frac{(1-\rho)}{n-1} D.$$

1158

1159 We started from the resource consumption model (equations 1 and 2 in Chesson 1990):

1160 
$$\begin{cases} \frac{1}{X_i} \frac{dX_i}{dt} = b_i [\sum_{l=1}^m c_{il} w_l R_l - m_i] \\ \frac{1}{R_l} \frac{dR_l}{dt} = r_l \left( 1 - \frac{R_l}{K_l} \right) - \sum_{i=1}^n c_{il} X_i \end{cases} \quad (\text{Eq. S1})$$
  
1161

1162 where  $i$  is a consumer species;  $X_i$  is population density;  $R_l$  is the density of food resource  $l$ ;  $w_l$   
 1163 is the value of one unit of resource  $l$  to the consumer;  $c_{il}$  is the rate at which consumer  
 1164 species  $i$  captures resource  $l$  per unit abundance of resource  $l$ ;  $m_i$  is the total value of  
 1165 resource that must be harvested per capita for the growth rate to be exactly 0;  $b_i$  is a factor  
 1166 converting the resource excess into the per capita growth rate;  $r_l$  is the growth rate and  $K_l$  the  
 1167 carrying capacity for resource  $l$ .

1168 Assuming a separation of time scales between resource and consumer dynamics, one obtains  
 1169 the quasi-static equilibrium of resources:

$$1170 \quad R_l(t) = \left[ 1 - \frac{1}{r_l} \sum_{i=1}^n c_{il} X_i(t) \right] K_l \quad (\text{Eq. S2})$$

1172 Introducing (Eq. S2) in the consumer dynamics presented in (Eq. S1) yields:

$$1173 \quad \frac{1}{X_i} \frac{dX_i}{dt} = b_i \left[ \sum_{l=1}^m c_{il} w_l K_l - m_i - \sum_{l=1}^m \sum_{j=1}^n c_{jl} c_{il} \frac{w_l K_l}{r_l} X_j \right] \quad (\text{Eq. S3})$$

1174 (Eq. S3) can be written using vectors (in bold letters) and matrices (in bold capital letters) as  
 1175 follows:

$$1176 \quad \frac{1}{X_i} \frac{dX_i}{dt} = b_i \left[ {}^t \mathbf{r} \mathbf{M} \mathbf{c}_i - m_i - \|\mathbf{c}_i\|_M \sum_{j=1}^n \frac{{}^t \mathbf{c}_j \mathbf{M} \mathbf{c}_i}{\|\mathbf{c}_i\|_M \|\mathbf{c}_j\|_M} \|\mathbf{c}_j\|_M X_j \right] \quad (\text{Eq. S4})$$

1177 where  $\mathbf{M}$  is a diagonal matrix such as  $M_{ll} = \frac{w_l K_l}{r_l}$  and  $\|\cdot\|_M$  is the corresponding Euclidean  
 1178 norm. Then, we can define the “standardized density”  $\tilde{X}_i = \|\mathbf{c}_i\|_M X_i$  so that (Eq. S4) yields:

$$1179 \quad \frac{1}{\tilde{X}_i} \frac{d\tilde{X}_i}{dt} = \tilde{b}_i \left[ \tilde{k}_i - \sum_{j=1}^n \tilde{\alpha}_{ij} \tilde{X}_j \right] \quad (\text{Eq. S5})$$

1181 where we have defined the standardized parameters:

$$1182 \quad \tilde{b}_i = \|\mathbf{c}_i\|_M b_i$$

$$1183 \quad \tilde{k}_i = \frac{{}^t \mathbf{r} \mathbf{M} \mathbf{c}_i - m_i}{\|\mathbf{c}_i\|_M} = \left[ \frac{{}^t \mathbf{r} \mathbf{M} \mathbf{c}_i}{\|\mathbf{c}_i\|_M \|\mathbf{r}\|_M} \right] \|\mathbf{r}\|_M - \frac{m_i}{\|\mathbf{c}_i\|_M}$$

$$1184 \quad \tilde{\alpha}_{ij} = \frac{{}^t \mathbf{c}_j \mathbf{M} \mathbf{c}_i}{\|\mathbf{c}_i\|_M \|\mathbf{c}_j\|_M}$$

1185 Importantly,  $\tilde{\alpha}_{ii} = 1$  and the Cauchy-Schwartz inequality ensures that  $\tilde{\alpha}_{ij} < 1$  when  $i \neq j$ .  
 1186 Consider species  $i$  at low density, and assume that the  $n - 1$  other species can coexist. Then  
 1187 all the other species  $j \neq i$  harboring their equilibrium density  $\tilde{X}_{ji}$  (second subscript “ $i$ ”  
 1188 meaning “in the absence of species  $i$ ”). Then the long-term growth rate at low density of  
 1189 species  $i$ ,  $\bar{r}_i$ , is obtained by neglecting  $\tilde{X}_i$  and setting the  $\tilde{X}_j$ s to  $\tilde{X}_{ji}$ s in (Eq. S5):

$$1190 \quad \bar{r}_i = \tilde{b}_i \left[ \tilde{k}_i - \sum_{\substack{j=1 \\ j \neq i}}^n \tilde{\alpha}_{ij} \tilde{X}_{ji} \right] \quad (\text{Eq. S6})$$

1192 Note that this is reminiscent of equation (58) of Chesson (1994), but in a standardized  
 1193 formulation of the system.

1194 We now consider  $n > 2$  species in competition. We define  $\tilde{\mathbf{A}}$  the matrix of standardized  
 1195 interaction coefficients among species ( $\tilde{A}_{ij} = \tilde{\alpha}_{ij}$ ). Then for any pair of species  $i \neq j$ , taking  
 1196 the equilibrium of  $j$  in the  $n - 1$  species system where  $i$  is absent in equation (Eq. S5) yields:

$$1197 \quad \tilde{\mathbf{A}}\tilde{\mathbf{X}} = \tilde{\mathbf{K}} + \Delta \quad (\text{Eq. S7})$$

1198 where we use the convention that  $\tilde{X}_{ii} = 0$ , we define  $\tilde{K}_{ij} = \tilde{k}_i$  and  $\Delta$  a diagonal matrix with  
 1199 unknown coefficients. (Eq. S7) implies that:

$$1200 \quad \tilde{\mathbf{X}} = \tilde{\mathbf{A}}^{-1}(\tilde{\mathbf{K}} + \Delta) \quad (\text{Eq. S8})$$

1202 Without making any assumptions on the coefficients of  $\tilde{\mathbf{A}}$ , combining (Eq. S6) and (Eq. S7)  
 1203 yields:

$$1204 \quad \tilde{r}_i = -\tilde{b}_i \Delta_{ii} \quad (\text{Eq. S9})$$

1206  $\Delta_{ii}$  can be computed by considering the  $(i, i)$  coefficient in (Eq. S8) and using that  $\tilde{X}_{ii} = 0$ :

$$1207 \quad [\Delta]_{ii} = -\frac{[\tilde{\mathbf{A}}^{-1}\tilde{\mathbf{K}}]_{ii}}{[\tilde{\mathbf{A}}^{-1}]_{ii}} \quad (\text{Eq. S10})$$

1209 Using (Eq. S10) in (Eq. S9) yields:

$$1210 \quad \tilde{r}_i = \tilde{b}_i \frac{[\tilde{\mathbf{A}}^{-1}\tilde{\mathbf{K}}]_{ii}}{[\tilde{\mathbf{A}}^{-1}]_{ii}} = \tilde{b}_i \left( \tilde{k}_i + \sum_{j=1, j \neq i}^n \frac{[\tilde{\mathbf{A}}^{-1}]_{ij}}{[\tilde{\mathbf{A}}^{-1}]_{ii}} \tilde{k}_j \right) \quad (\text{Eq. S11})$$

1211 We now consider a broader class of interaction matrices with the following form:

$$1212 \quad \tilde{\mathbf{A}} = \boldsymbol{\theta} \mathbf{t} \boldsymbol{\theta} + \mathbf{D} \quad (\text{Eq. S12})$$

1214 where:

- 1215 -  $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_n)$  is a vector of positive real numbers such as for all  $i$ ,  $\theta_i < 1$ ;
- 1216 -  $\mathbf{D}$  is a diagonal matrix such as  $D_{ii} = 1 - \theta_i^2$  for all  $i$ .

1217 (Eq. S12) implies that  $\alpha_{ij} = \theta_i \theta_j < 1$  for all  $i \neq j$ . We develop in the next section how such  
 1218 interaction matrices naturally emerge from competition for resources among species.

1219 If  $\tilde{\mathbf{A}}$  verifies (Eq. S12) then:

$$1220 \quad \tilde{\mathbf{A}}^{-1} = \left[ \mathbf{I} - \frac{\mathbf{D}^{-1} \boldsymbol{\theta} \mathbf{t} \boldsymbol{\theta}}{1 + \|\boldsymbol{\theta}\|_{\mathbf{D}^{-1}}^2} \right] \mathbf{D}^{-1} \quad (\text{Eq. S13})$$

1222 Using (Eq. S13) in (Eq. S11) yields:

$$\begin{aligned}
1223 \quad \bar{r}_i &= \tilde{b}_i \left( \tilde{k}_i - \sum_{j \neq i}^n \frac{\frac{\theta_i \theta_j [D^{-1}]_{jj}}{1 + \|\theta\|_{D^{-1}}^2} \tilde{k}_j}{\left(1 - \frac{[D^{-1}]_{ii} \theta_i^2}{1 + \|\theta\|_{D^{-1}}^2}\right)} \right) \\
1224 \quad &= \tilde{b}_i \left( \tilde{k}_i - \frac{\theta_i}{\left(1 + \|\theta\|_{D^{-1}}^2 - \frac{\theta_i^2}{1 - \theta_i^2}\right)} \sum_{j \neq i}^n \frac{\theta_j}{1 - \theta_j^2} \tilde{k}_j \right) \\
1225 \quad &= \tilde{b}_i \left( \tilde{k}_i - \frac{\theta_i}{\left(1 + \sum_{l \neq i}^n \frac{\theta_l^2}{1 - \theta_l^2}\right)} \sum_{j \neq i}^n \frac{\theta_j}{1 - \theta_j^2} \tilde{k}_j \right) \tag{Eq. S14}
\end{aligned}$$

1226 We now seek to decompose equation (Eq. S14) in stabilizing and equalizing terms as in  
1227 Chesson's (2000) framework. To find the stabilizing term, we first looked at the long-term  
1228 growth rate at low density when all the species have the same fitness  $\tilde{k}$ . In this case, only the  
1229 stabilizing should act on coexistence. When all the fitnesses are equal, (Eq. S14) becomes:

$$1230 \quad \bar{r}_i = \tilde{b}_i \tilde{k} \left( 1 - \frac{\theta_i \sum_{j \neq i}^n \frac{\theta_j}{1 - \theta_j^2}}{1 + \sum_{l \neq i}^n \frac{\theta_l^2}{1 - \theta_l^2}} \right) \tag{Eq. S15}$$

1231 We now search for the corresponding equalizing term in an additive decomposition as in  
1232 Chesson's (2000) framework (Eq. 1 above). Starting from (Eq. S14), we obtained the  
1233 following final expression for the long-term growth rate of a species:

$$1234 \quad \bar{r}_i = \tilde{b}_i \left( \tilde{k}_i - \frac{\sum_{j \neq i}^n \frac{\theta_j}{1 - \theta_j^2} \tilde{k}_j}{\sum_{j \neq i}^n \frac{\theta_j}{1 - \theta_j^2}} + \frac{\sum_{j \neq i}^n \frac{\theta_j}{1 - \theta_j^2} \tilde{k}_j}{\sum_{j \neq i}^n \frac{\theta_j}{1 - \theta_j^2}} \left( 1 - \frac{\sum_{j \neq i}^n \frac{\theta_i \theta_j}{1 - \theta_j^2}}{1 + \sum_{l \neq i}^n \frac{\theta_l^2}{1 - \theta_l^2}} \right) \right) = \tilde{b}_i (\tilde{k}_i - \bar{k}_i + (1 - \rho_i) \bar{k}_i) \tag{Eq. S16}$$

1235 where

$$\begin{aligned}
1236 \quad \bar{k}_i &= \frac{\sum_{j \neq i}^n \frac{\theta_j}{1 - \theta_j^2} \tilde{k}_j}{\sum_{j \neq i}^n \frac{\theta_j}{1 - \theta_j^2}} \\
1237 \quad \rho_i &= \frac{\theta_i \sum_{j \neq i}^n \frac{\theta_j}{1 - \theta_j^2}}{1 + \sum_{l \neq i}^n \frac{\theta_l^2}{1 - \theta_l^2}}
\end{aligned}$$

1238 We thus retrieve in (Eq. S16) a stabilizing term with a form consistent with Chesson (2000;  
1239 equation (4)). Note that changing  $\theta_i$  affects  $\rho_i$  but not  $\bar{k}_i$ , which is a desirable property of the  
1240 equalizing/stabilizing decomposition. The main difference with Chesson's (2000) framework  
1241 is that the  $\bar{k}_i$  is not a simple arithmetic mean of the fitnesses but also has to include some  
1242 interaction effects of other species. Changing interactions among species thus changes the  $\bar{k}_i$ s  
1243 even of fitnesses are maintained constant.

1244 [Making explicit the resource competition model underpinning the choice  \$\alpha\_{ij} = \theta\_i \theta\_j\$](#)

1245 We note  $\mathbf{c}_i = (c_{i1}, \dots, c_{im})$

1246 In any resource competition multispecies model,  $\mathbf{c}_i$ , the consumption vectors of species  $i$  can  
1247 be written as follows:

1248 
$$\mathbf{c}_i = \mathbf{c} + \mathbf{u}_i \tag{Eq. S17}$$

1249 Where  $\mathbf{c}$  is a component of the consumption vector which is shared among all the species and  
1250  $\mathbf{u}_i$  is a component of the consumption vector which is specific to species  $i$ . (Eq. S17) implies  
1251 that the interaction coefficients verify:

1252 
$$\tilde{\alpha}_{ij} = \frac{{}^t \mathbf{c}_j M \mathbf{c}_i}{\|\mathbf{c}_i\|_M \|\mathbf{c}_j\|_M} = \frac{{}^t \mathbf{c} M \mathbf{c} + {}^t \mathbf{u}_i M \mathbf{c} + {}^t \mathbf{u}_j M \mathbf{c} + {}^t \mathbf{u}_i M \mathbf{u}_j}{\|\mathbf{c}_i\|_M \|\mathbf{c}_j\|_M} \tag{Eq. S18}$$
  
1253 S18)

1254 (Eq. S18) implies that choosing the  $\mathbf{u}_i$ s such as:

1255 for all  $i$ ,  ${}^t \mathbf{u}_i M \mathbf{c} = 0$  (Eq. S19a)  
1256 S19a)

1257 for all  $i \neq j$ ,  ${}^t \mathbf{u}_i M \mathbf{u}_j = 0$  (Eq. S19b)  
1258 S19b)

1259 yields:

1260 
$$\tilde{\alpha}_{ij} = \frac{\|\mathbf{c}\|_M^2}{\|\mathbf{c}_i\|_M \|\mathbf{c}_j\|_M} = \theta_i \theta_j \tag{Eq. S20}$$

1261 where  $\theta_i = \frac{\|\mathbf{c}\|_M}{\|\mathbf{c}_i\|_M} = \frac{\|\mathbf{c}\|_M}{\sqrt{\|\mathbf{c}\|_M^2 + \|\mathbf{u}_i\|_M^2}} < 1$ . (Eq. S20) is the form of interaction coefficients we  
1262 chose in our model.

1263 A family of vectors  $\mathbf{u}_i$  verifying (Eq. S19a,b) always exists as long as the dimension of the  
1264 resource space verifies  $m \geq n$ . In words, there must be a number of resource types at equal or  
1265 superior to the number of competing species.

1266 Biologically speaking, (Eq. S18) and (Eq. S19a,b) makes competition among species happen  
1267 as follows: all the species share the same basic needs captured by  $\mathbf{c}$ . In addition to these basic  
1268 needs, each species  $i$  has some specific additional needs captured by  $\mathbf{u}_i$ , which differentiate  
1269 its ecological behavior (and its niche) from the others'.

1270 What makes our model not absolutely general is imposing (Eq. S19b), which states that  
1271 differences among species in their resource consumption must be orthogonal according to  $M$ .  
1272 We believe that this particular case is a fruitful and simple starting point for the study  
1273 multispecies competition models.

1274 **References**

1275 Chesson, P. (1990). MacArthur's consumer-resource model. *Theor. Popul. Biol.* 37, 26-38.

1276 Chesson, P. (1994). Multispecies competition in variable environments. *Theor. Pop. Biol.*, 45,  
1277 227-276.

1278 Chesson, P. (2000). Mechanisms of Maintenance of Species Diversity. *Annu. Rev. Ecol. Evol.*  
1279 *Syst.*, 31, 343–358.

1280 MacArthur, R. (1970). Species packing and competitive equilibrium for many species. *Theor.*  
1281 *Pop. Biol.*, 1, 1-11.

1282

1283

1284

1285

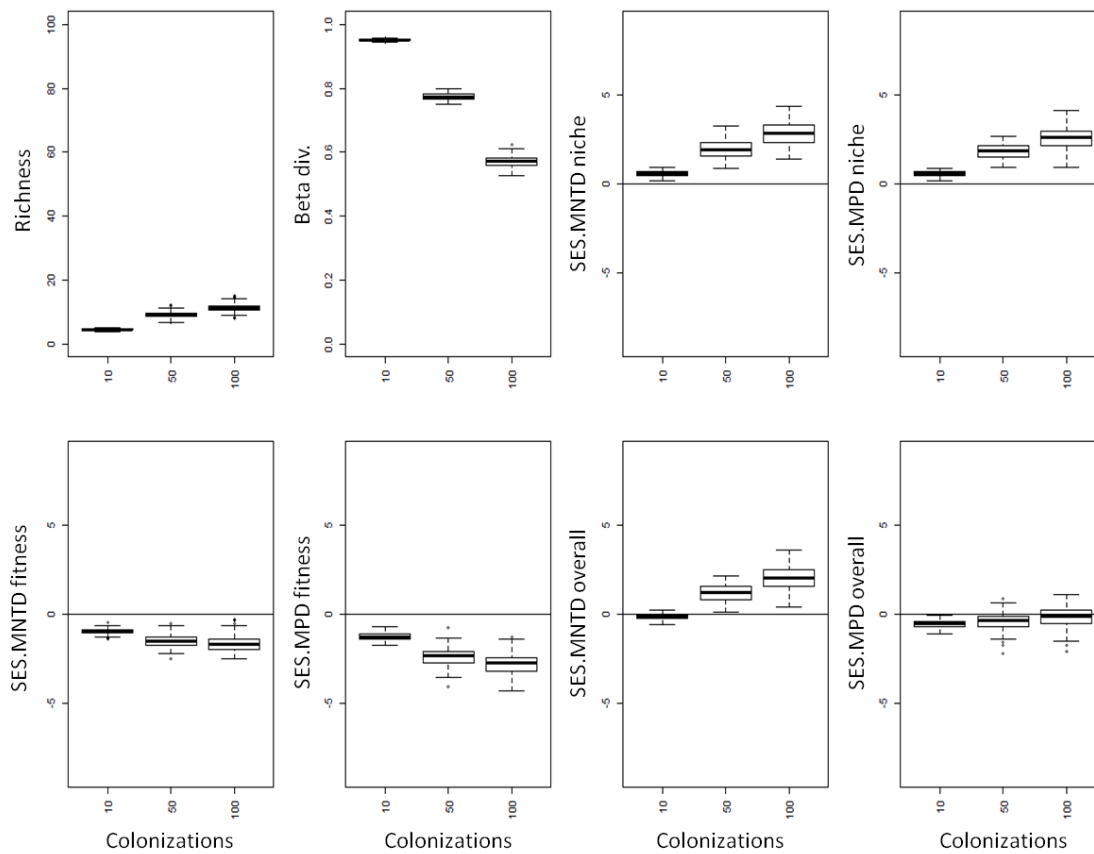
1286

### 1287 **Appendix S3 Supplementary results**

1288

#### 1289 **Additional scenarios considering different dispersal strategies**

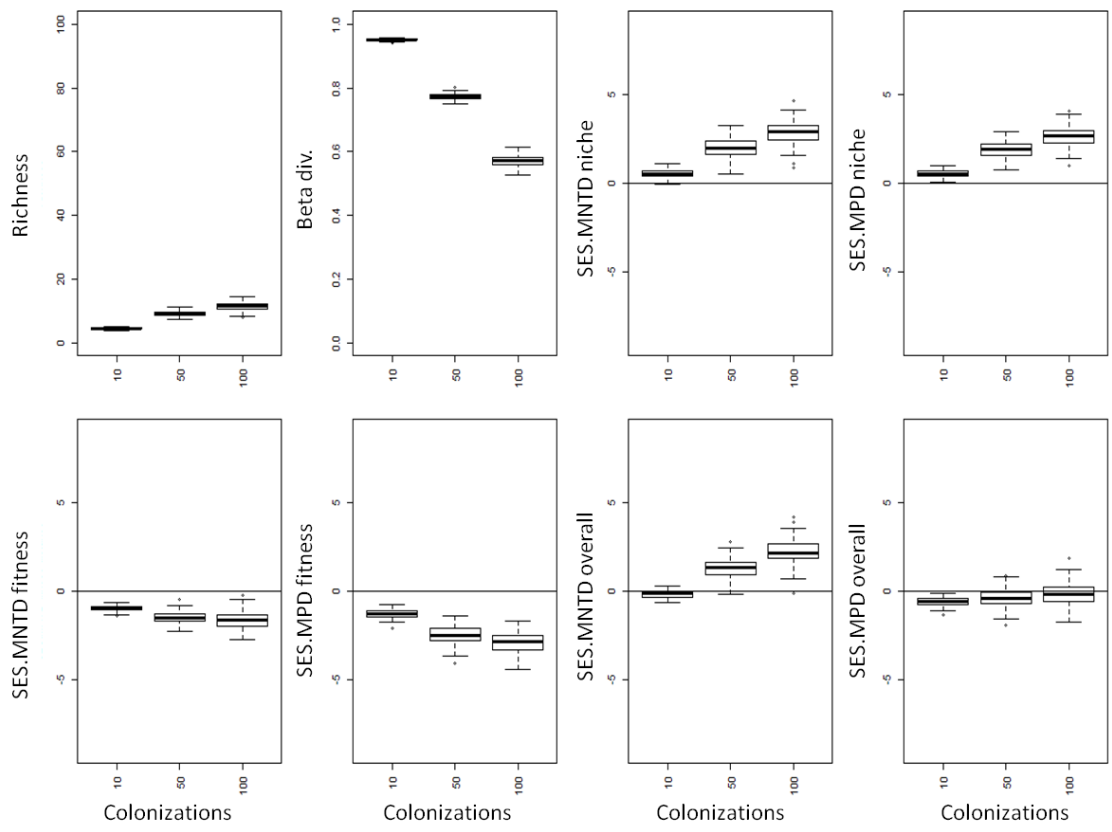
1290 First, we included a colonization-competition trade-off, where species chance to disperse was  
1291 inversely proportional to their  $k_s$ . So, species that are strong competitors are poor dispersers  
1292 (as already shown by empirical studies, as Cadotte 2007). Second, we simulated a  
1293 restricted/versatile scenario, where species chance to disperse is proportional to their  $\theta$ s.  
1294 In this case, species with higher dependence to common resources (generalists with high  
1295  $\theta$ s) would come first and late colonizers would be those with specialized resource  
1296 requirements (Bonte et al. 2003). This simulates the case where the resources for generalist  
1297 species are available more quickly (as would be expected if a similar species already occurs  
1298 in that community). As high  $\theta$  means that a species competes for the common resource, it  
1299 makes sense that this species would reach a patch more quickly, because its resources would  
1300 always be available (Reinhardt et al. 2005). So, in this case, marginal and specialized species  
1301 would come late in the colonization sequence.



1302

1303 **Figure S1.** Results of richness, Jaccard dissimilarity and species dispersion metrics when we  
 1304 varied the number of sequential colonization events during community assembly in the  
 1305 colonization-competition trade-off scenario. Community assembly was performed using a  
 1306 species pool with medium niche differences and fitness values at the species pool scale. We  
 1307 first applied indices SES-MNTD and SES-MPD separately to niche distances and fitness  
 1308 distances between species and then to the overall functional distances which combines niche  
 1309 and fitness as defined in the main text.





1310

1311 **Figure S2.** Results of richness, Jaccard dissimilarity and species dispersion metrics when we  
 1312 varied the number of sequential colonization events during community assembly in the  
 1313 restricted/versatile scenario. Community assembly was performed using a species pool with  
 1314 medium niche differences and fitness values at the species pool scale. We first applied  
 1315 indices SES-MNTD and SES-MPD separately to niche distances and fitness distances  
 1316 between species and then to the overall functional distances which combines niche and  
 1317 fitness as defined in the main text.

1318

1319

1320

1321

1322

1323

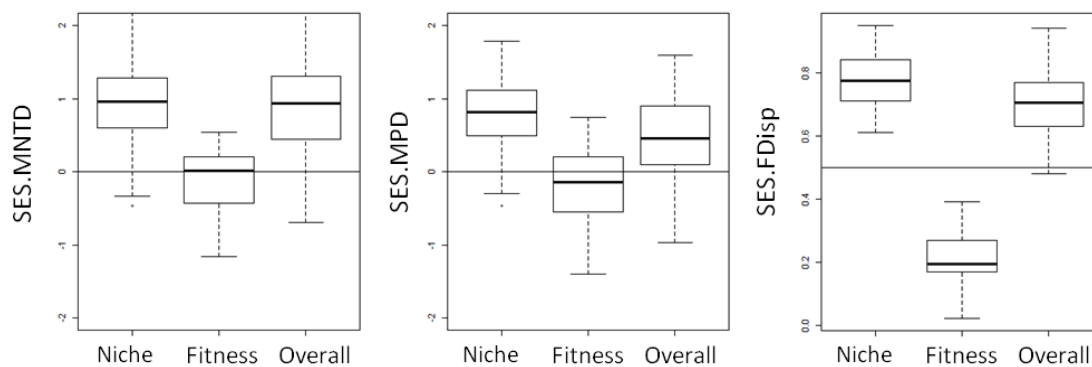
1324

1325

1326

1327 **Additional results using different functional diversity metrics**

1328 We calculated the Functional Dispersion (distance to centroid). We did that calculating FDisp  
1329 (Laliberté and Legendre 2010) and then using a randomization approach (Independent swap  
1330 by Gotelli and Entsminger 2003) in the distance matrix to provide a null model comparison.  
1331 The values in the figure below represent the number of times the real result was larger than  
1332 the null values, so values close to one means strong overdispersion (the real value was larger  
1333 than most of randomization). The results were similar to those of SES.MNTD and SES.MPD.



1334 **Figure S3.** Results for SES.MNTD, SES.MPD and SES.FDisp. The first box shows the  
1335 results of SES.MNTD, the second of SES.MPD and the third of the new SES.FDisp. The  
1336 horizontal line represents the null expectation. Values higher than the null expectation  
1337 represent more differences than expected by random and lower values the opposite. The  
1338 general pattern found in our study and used to support most of our discussion holds.  
1339 Sampling niche traits will always lead to trait overdispersion, fitness traits will be prone to  
1340 show clustering and random pattern and overall species distance, calculated using both niche  
1341 related and fitness related traits will tend strongly to overdispersion.

1343

1344

1345

1346

1347

1348 **References**

1349 Bonte, D., Vandenbroecke, N., Lens, L., and J. P. Maelfait. 2003. Low propensity for aerial

1350 dispersal in specialist spiders from fragmented landscapes. *Proceedings of the Royal*  
1351 *Society of London B: Biological Sciences* 270: 1601-1607.

1352 Cadotte, M. W. 2007. Concurrent niche and neutral processes in the competition–  
1353 colonization model of species coexistence. *Proceedings of the Royal Society of*  
1354 *London B: Biological Sciences* 274: 2739-2744.

1355 Reinhardt, K., Köhler, G., Maas, S., and P. Detzel. 2005. Low dispersal ability and habitat  
1356 specificity promote extinctions in rare but not in widespread species: the Orthoptera  
1357 of Germany. *Ecography* 28: 593-602.

1358 Laliberté, E., and P. Legendre. 2010. A distance based framework for measuring functional  
1359 diversity from multiple traits. *Ecology* 91: 299-305.

1360 Gotelli, N. J., and G. L. Entsminger. 2003. Swap algorithms in null model  
1361 analysis. *Ecology* 84: 532-535.

1362

1363