Unravelling the small-island effect through phylogenetic community ecology

Matthews, Tom; Rigal, François; Kougioumoutzis, Kostas; Trigas, Panayiotis; Triantis, Kostas A.

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

Unravelling the small-island effect through phylogenetic community ecology

Thomas J. Matthews* [1, 2+], François Rigal [3,2+], Kostas Kougioumoutzis[4], Panayiotis Trigas[5] & Kostas A. Triantis*[4]

[1]GEES (School of Geography, Earth and Environmental Sciences) and Birmingham Institute of Forest Research, University of Birmingham, Birmingham, B15 2TT

[2]CE3C – Centre for Ecology, Evolution and Environmental Changes/Azorean Biodiversity Group and Universidade. dos Açores – Depto de Ciências Agráriase Engenharia do Ambiente, PT-9700-042, Angra do Heroísmo, Açores, Portugal.

[3]Environment and Microbiology Team, Université de Pau et des Pays de l'Adour, IPREM UMR CNRS 5254, BP 1155, 64013 Pau Cedex, France.

[4]Department of Ecology and Taxonomy, Faculty of Biology, National and Kapodistrian University of Athens, Athens GR-15784, Greece.

[5]Faculty of Crop Science, Agricultural University of Athens, Iera Odos, GR-118 55, Athens, Greece.

[+]These authors contributed equally to this work.

*Correspondence authors. Emails: TJM – txm676@gmail.com; KAT - ktriantis@biol.uoa.gr

Running title: Phylogenetic diversity and the small-island effect

22 **Abstract**

23 **Aim:** The small-island effect (SIE) describes a different relationship between island area and

24 species richness on smaller compared to larger islands. The pattern has recently gained

25 widespread support. However, few studies have attempted to identify the actual mechanisms

26 driving the SIE. Here, we use a phylogenetic community framework to study the SIE, based

27 on the assumption that if the dominant assembly processes differ between small and large

28 islands, patterns of phylogenetic community structure should shift across the area and habitat

29 diversity gradient.

30 **Location:** The Aegean Archipelago, Greece.

31 **Methods:** We used a large dataset of 3262 vascular plant species distributed across 173

32 islands, in combination with a species-level phylogeny. The phylogenetic community

33 structure of each island was calculated using a null modelling framework and was quantified

34 using effect sizes (ES); negative values indicating phylogenetic clustering and positive values

35 overdispersion. Habitat diversity, species richness, phylogenetic diversity (PD) and ES values

36 were regressed against $\log_{10}$ -transformed area and we tested for a SIE using piecewise

37 regression models. We also assessed differences in taxonomic and phylogenetic composition

38 between small and large islands using a beta-diversity framework.

39 **Results:** We found evidence of a SIE using species richness, PD, and phylogenetic

40 community structure (ES values). Small islands displayed low variation in habitat diversity

41 and tended to be more phylogenetically clustered, while large islands shifted from

42 phylogenetic clustering towards phylogenetic overdispersion with increasing area and habitat

43 diversity. In addition, we showed that phylogenetic composition tended to be more similar

44 between small islands than expected.

45 **Main conclusion:** Overall, our results provide an example of a SIE in the analysis of island

46 phylogenetic community structure, and point to a role of habitat diversity in driving the SIE

47 more generally.

48

49 **Keywords:** Breakpoints, island biogeography, phylogenetic diversity, small-island effect,

50 species–area relationship, beta-diversity

51

1.   INTRODUCTION

The small-island effect can be broadly described as a different relationship between island area and species richness (the island species–area relationship; ISAR) on smaller compared to larger islands (Triantis & Sfenthourakis, 2012). That is, there is a threshold point (or at least a transitional phase) in island area where the slope of the ISAR changes; either species richness on small islands varies independently of area or increases / decreases at a different rate than on larger islands. The first graphical representation of the pattern was provided by Niering in 1956 (see his Figure 32), and not by the same author in 1963 as is commonly reported. Discussion of the SIE in the context of the equilibrium theory of island biogeography was then later provided by MacArthur & Wilson (1967) and Whitehead & Jones (1969). Although there were SIE studies in the intervening years, the pattern has undergone a resurgence in interest since the start of the millennium (Lomolino & Weiser, 2001; Triantis et al., 2006; Burns, McHardy, & Pledger, 2009; Dengler, 2010; Triantis & Sfenthourakis, 2012) and has now been shown to be present in a variety of systems, including true islands and habitat islands (Matthews, Steinbauer, Tzirkalli, Triantis, & Whittaker, 2014; Morrison, 2014; Wang, Millien, & Ding, 2016), although a number of exceptions exist (Matthews et al., 2014; Wang et al., 2016).

Although the SIE has become an important part of the theoretical framework of island biogeography, there is still considerable debate about the mechanisms underpinning the SIE. A number of explanations have been proposed and these can be broadly split into two main hypotheses (Sfenthourakis & Triantis, 2009). The first hypothesis postulates that on small islands extinction rates vary independently of island area due, for example, to stochastic events and / or to episodic catastrophic disturbances such as storms that may result in the (near) complete extirpation of species on small islands (e.g. MacArthur & Wilson, 1967). One possible pattern emerging from such area-independent extinction rates on small islands is higher temporal species turnover on smaller compared to larger islands (MacArthur & Wilson, 1967). The second hypothesis refers to the primary role of habitat diversity and niche-based factors in driving the SIE (Triantis et al., 2006; Sfenthourakis & Triantis, 2009; Chisholm et al., 2016; Chen, Yang, Tan, & Wang, 2020). Here, the idea is that on small islands island area and environmental heterogeneity become decoupled (Triantis, Mylonas, Weiser, Lika, & Vardinoyannis, 2005; Triantis et al., 2006), and that below a threshold of island area the absence of major habitat types is the main constraint on richness

85   (Sfenthourakis & Triantis, 2009; Chen et al., 2020). For example, Niering (1963) postulated

86   that the lack of mature soils on islands smaller than 3.5 acres explained the SIE pattern he

87   observed; it should be noted that Niering did not explicitly use the term SIE. Other proposed

88   mechanisms explaining the SIE include low immigration rates on small islands (Morrison,

89   2011) and species-specific area requirements (Schrader et al., 2019), both of which prevent

90   the successful establishment of species on small islands. However, despite all of these various

91   proposed hypotheses, in general, studies attempting to identify the actual mechanisms driving

92   the SIE in a given system are scarce (e.g. Morrison, 2011; Schrader et al., 2019).

93

94   In addition to the mechanisms generating the SIE, many other aspects of the SIE are still

95   debated. For instance, there have been strong disagreements about the appropriate statistical

96   approach for identifying SIEs (Dengler, 2010; Triantis & Sfenthourakis, 2012; Morrison,

97   2014; Wang et al., 2016). These disagreements include debates on whether or not to include

98   islands with no species (Morrison, 2014; Wang et al., 2016), the most effective statistical

99   model to detect a SIE (Lomolino & Weiser, 2001; Toms & Lesperance, 2003; Triantis et al.,

100   2006; Matthews et al., 2014; Chisholm, Fung, Chimalakonda, & O'Dwyer, 2016; Gao, Cao,

101   Xu, & Perry, 2019), and the best variable transformation procedure to use when fitting SIE

102   models (Burns et al., 2009; Matthews et al., 2014). Furthermore, most SIE studies have

103   focused on documenting patterns in species numbers without considering that species

104   composition may help to elucidate the mechanisms underlying the SIE. For instance, islet

105   specialists (species that exclusively occur on cliffs, rocks and coastal habitats on typically

106   small, low, uninhabited islets; Rechinger & Rechinger-Moser, 1951) are an important

107   component of the flora of small islands in the Aegean Sea (Panitsa & Tzanoudakis, 2001; see

108   Morrison, 2014, for a similar example from the flora of the Bahamas), but are largely absent

109   from the larger islands, and thus their presence will likely influence any SIE patterns in these

110   islands. In addition, SIE studies have only considered the number of species as the response

111   variable, and thus assuming that all species are evolutionarily independent and ecologically

112   equivalent. Critical information about primary community assembly mechanisms on islands

113   might thus have been ignored.

114

115   Recently, ecologists and biogeographers have turned to phylogenetic and functional diversity-

116   based metrics to provide novel insights into their study systems. Phylogenetic diversity (PD)

117   incorporates the evolutionary relationships between species and provides a measure of the

118   amount of evolutionary history that is represented by the species in an assemblage (Webb,

119   Ackerly, McPeek, & Donoghue, 2002). Based on the assumption that phylogenetic
120   relatedness between species is positively correlated with functional / ecological relatedness
121   (an assumption that may not hold in all cases; Flynn, Mirotchnick, Jain, Palmer, & Naeem,
122   2011), PD provides an integrative measure encompassing the overall similarity that stems
123   from multiple traits. When used in combination with an appropriate null model, the
124   comparison of the observed PD value to the distribution of null values enables classification
125   of the observed value as being significantly larger than expected by chance (i.e.
126   overdispersed), significantly smaller than expected (i.e. clustered), or not significantly
127   different from random expectation. These different situations have been interpreted as being
128   indicative of different community assembly processes (e.g. habitat filtering resulting in
129   clustered patterns; see Münkemüller et al., 2020, for a review). PD approaches are now
130   widely utilized in community ecology, and they have provided compelling evidence for the
131   importance of integrating both ecological and evolutionary processes to test mechanisms of
132   community assembly (Emerson & Gillespie, 2008; Webb et al., 2002). However, the study of
133   PD on islands remains in its infancy (Weigelt et al., 2015), and to our knowledge PD has
134   never been applied in the study of the SIE.

135

136   The present study is based on the idea that re-approaching the SIE using a phylogenetic
137   community approach will be beneficial as it will allow for a better assessment of the
138   mechanisms driving the pattern. Based on previous work on the SIE, and on distinct studies
139   focused on PD, we can generate several predictions of how communities should assemble in
140   regard to island size and habitat diversity and how, in turn, this will result in a SIE. Overall, if
141   the dominant assembly processes differ between small and large islands, patterns of
142   phylogenetic community structure should shift across the island area gradient. More
143   specifically, it is known that smaller islands often tend to contain a limited number of
144   relatively less-complex habitats (e.g. sclerophyllous scrub vegetation in the Mediterranean)
145   (Snogerup & Snogerup, 2004; Sfenthourakis, Pafilis, Parmakelis, Poulakakis, & Triantis,
146   2018). If habitat availability is the main driver of small island richness (e.g. Triantis et al.,
147   2006; Chen et al., 2020), we would expect species occurring on small islands (in a given
148   archipelago) to be clustered in the phylogenetic tree, because only a subset of closely related
149   species (e.g. from a limited number of clades) will likely be adapted to the limited habitat
150   types on small islands. This should be reflected at the island scale with small islands being
151   less phylogenetically diverse than expected by chance (phylogenetic clustering). According to
152   this argument, and bBased on previous work (e.g. Niering, 1963), we would also predict that

153 most islands below a threshold area value will contain only limited habitat types (i.e. below
154 this point the relationship between island area and habitat diversity is decoupled). That is,
155 there will be a threshold point in area before habitat diversity and island area increase together
156 in the expected linear fashion (Triantis et al., 2006). Overall then, we predict that, if habitat
157 diversity is a primary driver of island richness, due to the decoupling of island area and
158 habitat diversity on small islands, the relationship between area and diversity (e.g. richness,
159 PD; Chen et al., ; Triantis et al.,) will differ from that on large islands (i.e. the small island
160 effect), and also that small islands are expected to support a limited subset of closely related
161 species (i.e. PD clustering). In addition, if small islands contain similar habitat types, we
162 should also expect the biota of these islands to be more phylogenetically similar (phylogenetic
163 convergence) than expected by chance (Graham & Fine,).
164
165 In contrast, larger islands should have higher habitat heterogeneity (Ricklefs & Lovette, 1999;
166 Whittaker & Fernández-Palacios, 2007; Sfenthourakis & Triantis, 2009), allowing more
167 distant lineages (that are adapted to the wider range of habitats) to establish, persist and
168 diversify, and thus the species on large islands should be more spread across the phylogenetic
169 tree than expected by chance (i.e. overdispersed). On larger islands, the biota may be either
170 more or less phylogenetically similar (phylogenetic convergence or divergence) than
171 expected, depending on, amongst other things, the level of habitat similarity between the
172 islands, and / or the prevalence of *in situ* speciation (Graham & Fine, 2008).
173
174 To test the aforementioned predictions, we used a large and well sampled dataset of 3262
175 vascular plant species distributed across 173 islands and islets of varying size in the Aegean
176 Sea (Greece). The phylogenetic community structure of each island was assessed using a
177 species-level plant phylogeny and null models were implemented to test whether islands were
178 more phylogenetically clustered or overdispersed than expected by chance. We used
179 piecewise regression models to evaluate whether there were breakpoints in the ISAR (i.e. the
180 traditional SIE), and in various types of island phylogenetic diversity–area relationships
181 (IPDAR; i.e. a SIE in phylogenetic diversity). For the purposes of this study we define a SIE
182 as a threshold point in island area where there is a statistically significant change in the slope
183 of the diversity–area relationship. We also assessed whether species turnover and
184 phylogenetic dissimilarity between islands were lower or higher than expected by chance
185 using null models, and tested whether there were differences between small and large islands.
186 Our study documents the first example of a SIE in the phylogenetic community structure of

187   islands. In addition, our findings provide evidence suggesting a role of habitat diversity in

188   driving the traditional SIE. Overall, we demonstrate how a phylogenetic community-based

189   approach can be used to provide insights into the potential mechanisms driving the SIE, and

190   community assembly on islands more generally.

191

192   ## 2.   MATERIALS AND METHODS

193   ### 2.1 Data collection and phylogeny construction

194   Building on the work of Kougioumoutzis et al. (2017), we used a comprehensive database of

195   the presence-absence of plant species on 173 Aegean islands and islets (Figure 1a). We

196   compiled our dataset from an extensive and detailed database of the Aegean islands

197   (Appendix S1). Systematic research on the flora of the large Aegean Islands started in the late

198   1920s and has been undertaken continuously up to the present day. For the larger islands, the

199   dataset comprises all available records with the exception of species that have been recorded

200   only once in the distant past (i.e. their occurrence has not been confirmed with more recent

201   surveys). Study of the floristic composition of the small islands (Figure 1b-d) was undertaken

202   between 1970 and 2000. For the vast majority of these small islands, a single exhaustive plant

203   sampling trip was conducted during this period, and these data are used in the database. Island

204   species richness ranged from two (Islet East Aspronisi) to 1751 species (Crete). In total, the

205   dataset contains 3262 native vascular plant species (all non-native taxa were excluded). The

206   chorological status of the species follows Dimopoulos et al. (2013) and Strid (2016). The area

207   (km^2) of each of the 173 islands and islets was sourced from the literature (Panitsa, 1997;

208   Sfenthourakis & Triantis, 2009) and ranged from 0.0006 km^2 for Tourleta islet to 8270 km^2

209   for Crete.

210

211   Coarse-scale habitat type information for each island was calculated using the CORINE Land

212   Cover (CLC) inventory (Copernicus Land Monitoring Service, 2018). The habitat diversity of

213   each island was calculated as the total number of CLC classes present (out of a total of 44

214   classes). The CLC has a resolution of 100 m, which is relatively coarse in comparison to the

215   size of the smallest islands / islets in our dataset. As such, for these smaller islands, we used a

216   combination of satellite imagery, photographs, expert knowledge and field data from previous

217   work (e.g. Panitsa, 1997; Panitsa & Tzanoudakis, 2001; Sfenthourakis et al., 2018) to

218   compliment and ground-truth the CLC data. In all cases, only CLC classes were assigned as

habitat types (Appendix S2). Overall, 27 habitats were identified across the 173 islands, with habitat diversity ranging from 1 (78 islands) to 23 (Evvia, 3660.74 km^2).

We used the dated phylogeny of vascular plants with 31,749 species (Zanne et al., 2014) as a backbone tree to create a phylogeny of all species in our dataset (Figure 1e). Missing genera and species were randomly grafted among tips within their respective family and genera (see Appendix S3).

2.2 Assessing island taxonomic and phylogenetic diversity

Our first metric of phylogenetic diversity was Faith's phylogenetic diversity (PD; Faith, 1992), which is the sum of the phylogenetic branch lengths connecting all species co-occurring on an island. PD captures the total amount of evolutionary history contained within a given island. We included PD as a phylogenetic richness metric as, although it is usually highly correlated with species richness, the corresponding island PD–area relationship (IPDAR) is easily comparable with the traditional ISAR (Morlon et al., 2011).

To assess phylogenetic community structure on each island independently of species richness, we calculated the effect size (ES) of three distinct phylogenetic community metrics (Tucker et al., 2017): (1) the ES of PD (ES$_{PD}$); (2) the ES of mean pairwise distance (ES$_{MPD}$), which estimates the average phylogenetic distance between all pairs of species in a community; and (3) the ES of mean nearest taxon distance (ES$_{MNTD}$), which measures the mean phylogenetic distance between each species and its nearest neighbour in the phylogenetic tree with which it co-occurs in the community (Webb et al., 2002). While ES$_{MPD}$ provides an indication of how species on an island are dispersed in the phylogenetic tree, ES$_{MNTD}$ provides a measure of how densely phylogenetically packed species are on an island. The ES values were obtained using the probit transformation of the probability P that the observed value is lower than expected by chance (see further details in Appendix S4 and in Lhotsky et al., 2016). For each island and each metric, we obtained the null values using a random assembly model, where species names were randomly shuffled across the tree tips to create 9,999 new random phylogenetic diversity values per island. This constrained randomization approach keeps the island species richness, species turnover, and the spatial structure of species distributions constant (Swenson et al., 2012). ES values represent a measure of departure from the null expectation; values greater than zero being larger than expected (i.e. phylogenetic overdispersion), and values smaller than zero being less than expected (i.e. phylogenetic

252 clustering). Values greater than 1.96 (probit of roughly 0.975) or less than -1.96 (probit of
253 roughly 0.025) were considered to be significantly greater or less than expected, respectively.
254 This calculation of effect sizes was preferred to the commonly-used standardized effect size
255 (SES) due to the non-Gaussian shape of some of our null distributions (Bernard-Verdier et al.,
256 2012).

257

258 **2.3 Detection of a small-island effect**

259 To test for a SIE, we selected two of the most widely used SIE piecewise models proposed by
260 Lomolino & Weiser (2001): the continuous one-threshold and the left-horizontal one-
261 threshold model. It has recently been demonstrated that the number of thresholds in the ISAR
262 often increases with the range in island area analysed (Gao et al., 2019). In studies that
263 contain a large range in island area (as is the case with our data), piecewise models with more
264 than one threshold should also be considered (Gao et al., 2019). As such, we also included
265 two models with two thresholds: the continuous and the left-horizontal continuous two-
266 threshold model (see Appendix S5 for a detailed description of the models). We didn't
267 include discontinuous threshold models that have been applied in previous studies (e.g.
268 Matthews et al., 2014; Gao et al., 2019) as the ecological logic of discontinuous relationships
269 in nature has been questioned (Yu et al., 2020). To test whether the relationships between
270 diversity and island area exhibited a SIE, we compared the four piecewise models with a
271 simple linear model. A model with just an intercept was also included in our analyses as a
272 null model to test for an absence of relationship between $\log_{10}$ area and diversity.

273 We fitted our six models to the relationships between area and (1) habitat diversity, as a
274 means to evaluate any decoupling of the relationship between habitat diversity and area, and
275 (2) our five diversity metrics (species richness, PD, ES_{PD}, ES_{MNTD} and ES_{MPD}). We used a
276 semi-log transformation ($\log_{10}$) approach, i.e. area was log transformed but diversity was not.
277 Although semi-log transformation has been widely used in SIE studies (Morrison, 2014),
278 Matthews et al. (2014) argued that the choice of log-transformation used when fitting
279 piecewise regression models should be an integral part of the study design, and that the choice
280 of transformation depends on the aim of the study. Here, a semi-log transformation was used
281 as there is no *a priori* reason to log-transform metrics such as PD or the three ES metrics. The
282 validity of this choice was tested using a sensitivity analysis (outlined below).

283 All models were fitted using ordinary least squares regression (OLS), and threshold values

284   were estimated using the minimum residual sum of squares (RSS) method (see electronic
285   supplementary material, Appendix S5). The functions for fitting and plotting the threshold
286   models, along with additional piecewise model functions, were integrated into the 'sars' R
287   package (Version 1.2.3; currently available on GitHub in the 'thresholds_2' branch)
288   (Matthews et al., 2019). For each diversity metric, we compared the different models using
289   the Bayesian Information Criterion (BIC; Burnham & Anderson, 2002). In each case, the best
290   model was considered to be that with the smallest BIC; however, all models with a ΔBIC
291   value < 2 were considered as having a similar degree of statistical support (Burnham &
292   Anderson, 2002). BIC was preferred to other information criteria because, as ES values tend
293   be more scattered than species richness values, spurious relationships might simply arise due
294   to random noise in the data and be incorrectly interpreted as legitimate relationships. BIC is
295   more effective in such situations due to the greater penalization in respect to the number of
296   model parameters (Link & Barker, 2006). We also tested the ability of the best model(s) to
297   account for potential spatial structure by estimating Moran's I spatial autocorrelation index
298   using the model residuals and the latitude and longitude of each island's centroid. Diagnostic
299   plots were inspected in order to detect residual deviations from normality and
300   homoscedasticity, and to identify outliers. Finally, we repeated the aforementioned model
301   fitting and model comparison analyses using habitat diversity instead of island area as the
302   predictor of the five diversity metrics.

303   **2.4 Assessing differences in species turnover and phylogenetic dissimilarity between**
304   **small and large islands**
305   We split islands into small and large using an average breakpoint calculated by averaging all
306   thresholds identified in the best models for all metrics (see results). We first estimated species
307   turnover between islands using the modified Simpson's index (β_{SIM}). β_{SIM} measures
308   compositional differences between communities independently of any species richness
309   gradients (Baselga, 2010). We investigated whether observed pairwise β_{SIM} values were larger
310   or smaller than expected under a stochastic model of community assembly. This was carried
311   out by generating 9,999 random assemblage matrices using the fixed-fixed independent swap
312   algorithm (Gotelli, 2000). The fixed-fixed independent swap algorithm keeps species
313   occurrence frequency and sample species richness constant while shuffling species presences
314   across sites. The effect size (ES) of each observed pairwise β_{SIM} value was then calculated to
315   quantify the deviation from the null expectation with ES values greater than 1.96 or less than -
316   1.96 considered to be significantly more dissimilar or dissimilar than expected, respectively.

317   Both pairwise β_{SIM} and the associated ES values were calculated using all islands and also

318   using the small and large island subsets.

319

320   Phylogenetic dissimilarity was measured in an analogous way to the measurement of

321   phylogenetic community structure. For a given pair of islands A and B, we calculated the

322   mean of all phylogenetic distances separating species occurring in A from species occurring

323   in B (MPD_{DISS}), and the average of the phylogenetic distance between each species in A and

324   its closest relative in B ($MNTD_{DISS}$). We computed the ES of MPD_{DISS} and $MNTD_{DISS}$ values

325   using the same approach as for ES_{PD}, ES_{MPD} and ES_{MNTD}.

326

327   We further tested for differences in mean β_{SIM} and ES between small and large islands using

328   Wilcoxon Rank Sum tests, and whether the distribution of ES values within small and large

329   islands significantly deviated from the null expectation (median = 0) using a one-sample

330   Wilcoxon test (Swenson & Enquist, 2009). Since islands in the Aegean are spread across a

331   large geographical region i.e. the Aegean Sea, differences in turnover and phylogenetic

332   dissimilarity between small and large islands could also be driven simply by geographical

333   distance (i.e. dispersal limitation). For example, if higher turnover is found between small

334   islands compared to large islands, this simply could be due to the fact that small islands are

335   relatively far away from each other and not because they differ in terms of the habitat types

336   they contain. β_{SIM}, ES β_{SIM}, ES MPD_{DISS} and ES $MNTD_{DISS}$ dissimilarity matrices were

337   therefore correlated with the geographical distance between islands using a Mantel

338   permutation test, based on 9,999 permutations. This was implemented for all islands and for

339   small and large islands separately.

340   **2.5 Complementary analyses**

341   We undertook three complementary analyses to ensure our results were robust. First, we

342   checked the sensitivity of our results to the method used for placing unresolved genera and

343   species in the phylogeny. Second, we re-ran the above analyses on two smaller subsets of

344   islands, namely North-Eastern Aegean (83 islands and islets) and the Central Aegean (60

345   islands and islets) (Kougioumoutzis et al., 2017; Appendix S6). Third, we performed a

346   sensitivity test to ensure that we did not generate spurious small-island effects because of our

347   decisions regarding data transformation using the analytical procedure of Burns et al. (2009)

348   (Appendix S7). All analyses were undertaken using R (version 3.6; R Core Team, 2019).

349

3   RESULTS

3.1 The small-island effect: habitat diversity, the ISAR, IPDAR and phylogenetic community structure

There was some model uncertainty regarding the relationship between habitat diversity and $\log_{10}$ Area, with the continuous and left-horizontal continuous one-threshold models and the more complex continuous two-threshold models all having similar support (Figure 2a and Table 1). However, all of these best models explained 94% of the variation in habitat diversity and all showed at least one threshold between 3.45 and 6.28 km^2 (Figure 2a and Table 1). On small islands there was a slow, or an absence of an, increase in habitat diversity with $\log_{10}$ Area (most of the small islands containing no more than two habitats), while for the large islands a strong positive relationship was detected.

As expected, species richness and PD were highly correlated (Pearson's correlation 0.97, P<0.001, Figure S1). For both metrics, SIE thresholds were detected (Figure 2b, c and Table 1). For species richness, the left-horizontal two-threshold model was identified as the single best model (Figure 2b). For PD, the continuous one-threshold model and the continuous two-threshold model were identified as the best and the second-best model respectively (Figure 2c and Table 1). For both the ISAR and IPDAR, the best models explained 95% of the variation (Table 1).

For all three ES metrics, the continuous one-threshold model was found to be the best model (Table 1) explaining overall less variation than for the ISAR and IPDAR with R^2 values ranging from 37% for ES_{PD} to 22% ES_{MNTD} (Table 1). For the three ES metrics, the slope of the first segment of the piecewise regression was always negative, with islands switching from neutral assembly (-1.96 < ES < 1.96) to phylogenetic clustering (ES < -1.96). The slope of the second segment was always positive, with islands switching from phylogenetic clustering to neutral assembly and, to a lesser extent, to phylogenetic overdispersion (ES > 1.96) (Figure 2d,e,f). For ES_{MNTD}, the left-horizontal two-threshold model was found to be the second best model (Figure 2f and Table 1). Overall, no spatial autocorrelation was detected in the residuals of the best models (Table S1 in Appendix S8), and no specific deviations from the assumptions of regression (e.g. normality, presence of outliers) were detected.

The relationship between species richness and habitat diversity was significantly positive and

383 linear, with habitat diversity explaining 93% of the variation in species richness (Table S2 and

384 Figure S2a in Appendix S8). A similar pattern was found for PD although the relationship

385 included a first phase of faster increase (Table S2 and Figure S2b). For the ES_{PD} and ES_{MPD}

386 metrics, the continuous one-threshold model was found to be the best model (Table S2 and

387 Figure S2c,d), with the slope of the first segment negative and the slope of the second

388 segment positive; the threshold being at 3.9 and 2.9 habitats respectively. For ES_{MNTD}, the

389 left-horizontal one-threshold model was found to be best with a threshold identified at 6.7

390 habitats (Table S2 and Figure S2e).

391

392 **3.2 Comparing species turnover and phylogenetic dissimilarity between small and large**

393 **islands**

394 We classified islands as small or large using a common threshold of 2.57 km^2 (log10-

395 tranformed value = 0.41; corresponding to 92 small and 81 large islands), calculated by

396 averaging all thresholds identified in the best models across all metrics (range in threshold

397 values across the best models: 0.14 km^2 – 34.52 km^2). Although this range of threshold values

398 was quite large (mainly due to the two-threshold models), the average resulted in quite a clear

399 split between small and large islands (see Figure S3 in Appendix S8). Overall, among the

400 3262 species recorded in our dataset, 640 occurred on small islands with three being found

401 exclusively on small islands. Differences in habitat composition between small islands was

402 very low, with most small islands containing either natural grasslands and/or sclerophyllous

403 vegetation (Appendix S2). In contrast, large islands displayed substantial differences in

404 habitat composition, with a mean habitat diversity of 8.1 (SD = 4.7) out of a total of 27

405 habitats present on large islands.

406

407 Observed species turnover (β_{SIM}) was significantly higher between small (mean β_{SIM} = 0.66)

408 than between large (mean β_{SIM} = 0.24) islands (Wilcoxon test P < 0.001, Figure 3a). ES β_{SIM}

409 was significantly lower between small than between large islands (mean ES β_{SIM} = -0.20 and -

410 0.09 respectively, Wilcoxon test P < 0.001, Figure 3b). For small and large islands, mean ES

411 β_{SIM} was significantly lower than 0 (one-sample Wilcoxon test P < 0.001). However, most of

412 the pairwise ES β_{SIM} values for small islands did not show any significant deviation from

413 random (99.9% of the pairwise ES β_{SIM} values were neutral, Figure 3), while for large islands

414 a few pairs were identified as being significantly phylogenetically convergent or divergent

415 (7.2% and 5.9% respectively, Figure 3b). For β_{SIM}, a significant and positive correlation with

416 geographical distance was found only for large islands, while for ES β_{SIM}, a significant and

417　positive correlation with geographical distance was found for all islands and for small and

418　large islands separately (Table 2).

419

420　In regard to phylogenetic composition, both small and large islands exhibited significant

421　average phylogenetic convergence for ES MPD_{DISS} (one-sample Wilcoxon test P < 0.001;

422　mean ES MPD_{DISS} = -0.91 and -0.87 respectively, Figure 3c), and ES MPD_{DISS} values did not

423　differ between small and large islands (Wilcoxon test P = 0.64). For ES $MNTD_{DISS}$, small

424　islands exhibited significant average convergence, while large islands exhibited significant

425　average divergence (Wilcoxon test p < 0.001; one-sample Wilcoxon test P < 0.001 for both

426　size-groups; ES $MNTD_{DISS}$ = -0.63 and 1.54, respectively) (Figure 3d). There was no

427　significant correlation between either ES MPD_{DISS} or ES $MNTD_{DISS}$ and geographical

428　distance, for either small or large islands (Table 2).

429

430　**3.3 Complementary analyses**

431　We found that the values of phylogenetic diversity and dissimilarity were not sensitive to the

432　method used for placing unresolved genera and species in the phylogeny (Table S3 in

433　Appendix S8). Results obtained for the North-Eastern Aegean and Central Aegean subsets are

434　given in Appendix S6 and Tables S4-S5 and Figures S4-S8 in Appendix S8. Overall, results

435　obtained for the North-Eastern Aegean and the Central Aegean were very similar to those

436　obtained using all islands; the main exception was for the Central Aegean, where no SIE was

437　detected when using the phylogenetic ES_{PD} and ES_{MNTD} metrics. Using the approach of Burns

438　et al. (2009) we found that the presence of a SIE in the relationship between island area and

439　species richness is not simply an artefact of our use of a semi-log transformation (Figure S9 in

440　Appendix S8).

441

442　**4　DISCUSSION**

443　**4.1 Phylogenetic community structure and the small-island effect**

444　We have found evidence of a SIE in the island species–area relationship, the island

445　phylogenetic diversity–area relationship and when using three phylogenetic community

446　metrics (ES_{PD}, ES_{MPD} and ES_{MNTD}). Interestingly, although there was some degree of model

447　uncertainty and the thresholds estimated from the one and two-threshold models were slightly

448　different, all thresholds (across all response variables) fall within a range of 34.38 km^2 (from

449　0.14 to 34.52 km^2) representing no more than 0.4% of the total area range found in our data

450   (8270 km^2). This suggests that the abrupt shifts in the various diversity metrics were more or

451   less synchronised, and could reflect that it is around this area range that the diversity of

452   habitats starts to increase with area in a linear fashion.

453

454   In relation to ES$_{PD}$, ES$_{MPD}$ and the ES$_{MNTD}$, we found that, for the smaller islands (i.e. those to

455   the left of the breakpoint), island communities tended to be more phylogenetic clustered as

456   area increased (i.e. a negative slope – discussed in the next paragraph, Figure 2d,e,f). We also

457   observed that within the SIE range, area and habitat diversity are decoupled (see also Triantis

458   et al., 2006), with habitat diversity being almost stable throughout the SIE range. If habitat

459   diversity is the main driver of diversity rather than area per se, one possible explanation for

460   our observation of phylogenetic clustering on smaller islands is greater habitat filtering, where

461   only species being able to establish in the limited habitat types accumulate, leading to

462   phylogenetic redundancy (species drawn from a limited array of clades) within the species

463   assemblages. It is worth noting that small islands mostly contain one or both of natural

464   grasslands and sclerophyllous vegetation as land-cover classes (Appendix S2). This result

465   corroborates an earlier study assessing the habitats of a number of small islets in the Aegean

466   (including various islands included in our dataset) where it was found that many major habitat

467   types, such as mature forest and freshwater, were missing (Snogerup & Snogerup, 2004; see

468   also Panitsa, 1997; Sfenthourakis et al., 2018). However, these findings were not discussed in

469   the context of the ISAR or the SIE (Snogerup & Snogerup, 2004).

470

471   Interestingly, we did not identify phylogenetic clustering when focusing just on the very small

472   islands (i.e. the smallest of the islands we have classified here as small) (see Figure 2d,e,f).

473   The above habitat filtering explanation may explain this observation. However, this could

474   also be due to the species present on the smallest islands being adapted (i.e. possessing salt

475   tolerance) to littoral habitats, that are prevalent on very small islands (Snogerup & Snogerup,

476   2004). If salt tolerance is distributed across several families, this would explain the random

477   patterns observed in the phylogenetic structure of these islets (Flowers, Galal, & Bromham,

478   2010). As small island area increases, 'phrygana' communities primarily comprising annual

479   grasses and members of few other families start to dominate, leading to phylogenetic

480   clustering. In addition, it could be that on the very smallest islets natural disturbance (e.g.

481   intense storms and/or waves) has more of an impact (Whittaker & Fernández-Palacios, 2007),

482   driving community assembly to be essentially random with regard to phylogenetic structure.

483   Alternatively, it could be that, for the small islands, another variable co-varies with area, such

484 as human disturbance. For example, the smallest islets may be unsuitable for human activities

485 (e.g. tourism or land for grazing), meaning that there is in fact a greater amount of certain

486 types of disturbance (and reduction in habitat heterogeneity) on the larger of the smaller

487 islands, leading to the patterns observed.

488 As island area increased, we observed a gradual shift from phylogenetic clustering toward

489 phylogenetic overdispersion on individual islands, which mirrored a clear increase in habitat

490 diversity. One interpretation of this finding is that, on larger islands, the greater range of

491 habitats, such as freshwater systems, sub-alpine ecosystems and mature forest, enables a

492 much wider variety of functional forms drawn from a broader range of clades to persist. This

493 will result in greater phylogenetic diversity (and its ES) on the larger islands.

494

495 **4.2 Differences in species and phylogenetic composition between small and large islands**

496 A shortcoming of previous SIE studies has been the focus on species numbers, without any

497 consideration of species composition (Morrison, 2014). We found striking differences

498 between small and large islands in regard to species turnover: species spatial turnover

499 between small islands was high (Figure 3a) and mostly random (ES β_{SIM}, Figure 3b). The

500 strong correlation between ES β_{SIM} and geographical distance also suggests that neutral

501 processes such as dispersal limitation might have played an important role in shaping species

502 composition on small islands in the Aegean (see Runemark, 1971). In contrast, phylogenetic

503 composition tended to be, on average, more similar on small islands than expected by chance,

504 and independent of geographical distance, supporting our hypothesis of possible deterministic

505 convergence regarding the main plant clades present on small islands. This is also reinforced

506 by the small differences in habitat composition between the small islands, with almost all

507 small islands containing only natural grasslands and/or sclerophyllous vegetation. Altogether,

508 this indicates that, while only certain clades are able to establish and survive on the smaller

509 islands, the identity of these taxa (i.e. the identity of the taxa within these clades) is more

510 randomly distributed. This is likely due, at least in part, to stochastic extinction / immigration

511 and / or dispersal. That is, both neutral and niche-based processes might operate in

512 combination in this system to drive the SIE.

513

514 On the larger islands, much lower spatial turnover was observed, implying that species

515 composition on the larger islands was more ordered. Patterns of phylogenetic composition on

516 the large islands were more complex, with our two metrics showing opposite results: ES

517   MPD$_{DISS}$ identifying mostly convergence and ES MNTD$_{DISS}$ mostly divergence. ES MPD$_{DISS}$

518   is a measure of phylogenetic dissimilarity that captures variation associated with basal nodes

519   (e.g. substitutions of a given order or even higher taxonomic ranks; Kembel et al., 2010) and

520   thus showed an overall similarity between large islands due to the fact most of the higher

521   taxonomic groups represented in the regional island species pool occur on most of the large

522   islands. In contrast, ES MNTD$_{DISS}$ focuses on variation associated with terminal nodes and is

523   more sensitive to lower-level taxonomic substitutions (i.e. changes in representation of

524   families/genera) among communities (Kembel et al., 2010). Therefore, the divergence

525   reported between large islands with ES MNTD$_{DISS}$ could be due to differences in the source

526   pools (Kougioumoutzis et al., 2017), or *in-situ* speciation.

527

528   **4.3 Caveats and future directions**

529   This study is not the first to find breakpoints in the ISAR (Matthews et al., 2014; Niering,

530   1963; Wang et al., 2016), or to postulate that the mechanisms driving the ISAR (and SIE)

531   may change across breakpoints in the relationship (e.g. see Triantis et al. 2006 and Losos &

532   Parent, 2010, for a discussion on the speciation–area relationship). However, to our

533   knowledge this is the first assessment of the SIE undertaken through the lens of phylogenetic

534   community ecology, and the approach has proven useful for shining light on the potential

535   mechanisms driving the SIE. However, as always, in the absence of controlled experiments

536   and trait data (discussed below) and more explicit fine-scale measurements of habitat

537   diversity, it is impossible to be certain that habitat availability is the primary driver of the SIE

538   in this system. Nonetheless, based on our results we would argue that the approach applied

539   here (i.e. combining evaluation of the ISAR, IPDAR and species and phylogenetic

540   composition) represents a useful framework for analysing the SIE and the underlying

541   mechanisms, and assembly processes on islands more generally.

542

543   Going forward it would be interesting to expand this approach to include functional diversity

544   (e.g. see Schrader et al., 2020, for an interesting start in this direction). An assumption of the

545   use of phylogenetic diversity in these types of studies is that there is a correlation between

546   phylogenetic diversity and functional diversity. However, while a correlation between

547   phylogenetic diversity and functional diversity may exist in many systems, it is unlikely ever

548   to be a perfect association (Flynn et al., 2011). While many traits have been shown to have

549   strong phylogenetic signal, such that related species have similar traits, some have not

550   (Srivastava et al., 2012), and arguably functional trait data are better than phylogenetic data in

551 community ecology studies of this nature (Paine, Deasey, Duthie, & Thompson, 2018). Thus,

552 the expansion of the framework presented here to include functional diversity will no doubt

553 increase our understanding of the SIE.

554

559

560 **DATA ACCESSIBILITY**

561 The data are provided in the Supporting information.

562

563 **REFERENCES**

564 Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity.
565   *Global Ecology and Biogeography*, *19*, 134–143.

566 Bernard-Verdier, M., Navas, M. L., Vellend, M., Violle, C., Fayolle, A., & Garnier, E.
567   (2012). Community assembly along a soil depth gradient: contrasting patterns of plant
568   trait convergence and divergence in a Mediterranean rangeland. *Journal of Ecology*,
569   *100*, 1422–1433.

570 Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multi-model inference: a*
571   *practical information-theoretic approach*, 2nd edn. New-York: Springer.

572 Burns, K. C., Paul McHardy, R., & Pledger, S. (2009). The small-island effect: fact or
573   artefact? *Ecography*, *32*, 269–276.

574 Chen, C., Yang, X., Tan, X., & Wang, Y. (2020). The role of habitat diversity in generating
575   the small-island effect. *Ecography*, DOI: 10.1111/ecog.05092.

576 Chisholm, R. A., Fung, T., Chimalakonda, D., & O'Dwyer, J. P. (2016). Maintenance of
577   biodiversity on islands. *Proceedings of the Royal Society B: Biological Sciences*, *283*,
578   20160102.

579 Copernicus Land Monitoring Service (2018) *CLC 2018*, version v.20_20u1. URL:
580   https://land.copernicus.eu/pan-european/corine-land-cover/clc2018?tab=download. ©
581   European Union, Copernicus Land Monitoring Service 2018, European Environment
582   Agency (EEA).

583 Dengler, J. (2010). Robust methods for detecting a small island effect. *Diversity and*
584   *Distributions*, *16*, 256–266.

585 Dimopoulos, P., Raus, T., Bergmeier, E., Constantinidis, T., Iatrou, G., Kokkini, S., Strid, A.,
586   & Tzanoudakis, D. (2013). *Vascular plants of Greece. An annotated checklist*. Berlin:
587   Botanic Garden and Botanical Museum Berlin-Dahlem.

588 Emerson, B. C., & Gillespie, R. G. (2008). Phylogenetic analysis of community assembly and
589   structure over space and time. *Trends in Ecology & Evolution*, *23*, 619–630.

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

592 Flowers, T. J., Galal, H. K., & Bromham, L. (2010). Evolution of halophytes: multiple origins
593   of salt tolerance in land plants. *Functional Plant Biology*, *37*, 604–612.

594 Flynn, D. F. B., Mirotchnick, N., Jain, M., Palmer, M. I., & Naeem, S. (2011). Functional and
595   phylogenetic diversity as predictors of biodiversity–ecosystem-function relationships.
596   *Ecology*, *92*, 1573–1581.

597 Gao, D., Cao, Z., Xu, P., & Perry, G. (2019). On piecewise models and species–area patterns.
598   *Ecology and Evolution*, *9*, 8351–8361.

599 Gotelli, N. J. (2000). Null model analysis of species co-occurrence patterns. *Ecology*, *81*,
600   2606–2621.

601 Graham, C. H., & Fine, P. V. A. (2008). Phylogenetic beta diversity: linking ecological and
602   evolutionary processes across space in time. *Ecology Letters*, *11*, 1265–1277.

603 Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D.,
604   Blomberg, S. P., & Webb, C. O. (2010). Picante: R tools for integrating phylogenies
605   and ecology. *Bioinformatics*, *26*, 1463–1464.

606 Kougioumoutzis, K., Valli, A. T., Georgopoulou, E., Simaiakis, S. M., Triantis, K. A., &
607   Trigas, P. (2017). Network biogeography of a complex island system: the Aegean
608   Archipelago revisited. *Journal of Biogeography*, *44*, 651–660.

609 Lhotsky, B., Kovács, B., Ónodi, G., Csecserits, A., Rédei, T., Lengyel, A., Miklos K., &
610   Botta-Dukát, Z. (2016). Changes in assembly rules along a stress gradient from open
611   dry grasslands to wetlands. *Journal of Ecology*, *104*, 507-517.

612 Link, W. A., & Barker, R. J. (2006). Model weights and the foundations of multimodel
613   inference. *Ecology*, *87*, 2626–2635.

614 Lomolino, M. V., & Weiser, M. D. (2001). Towards a more general species–area relationship:
615   diversity on all islands, great and small. *Journal of Biogeography*, *28*, 431–445.

616 Losos, J. B., & Parent, C. E. (2010). The speciation–area relationship. In J. B. Losos & R. E.
617   Ricklefs (Eds.), *The theory of island biogeography revisited* (pp. 415–438). Princeton:
618   Princeton University Press.

619 MacArthur, R. H., & Wilson, E. O. (1967). *The theory of island biogeography*. Princeton:
620   Princeton University Press.

621 Matthews, T. J., Steinbauer, M. J., Tzirkalli, E., Triantis, K. A., & Whittaker, R. J. (2014).
622   Thresholds and the species–area relationship: a synthetic analysis of habitat island
623   datasets. *Journal of Biogeography*, *41*, 1018–1028.

624 Matthews, T. J., Triantis, K., Whittaker, R. J., & Guilhaumon, F. (2019). sars: an R package
625 for fitting, evaluating and comparing species–area relationship models. *Ecography,*
626 *42*, 1446–1455.

627 Morlon, H., Schwilk, D. W., Bryant, J. A., Marquet, P. A., Rebelo, A. G., Tauss, C.,
628 Bohannan, B. J. M., & Green, J. L. (2011). Spatial patterns of phylogenetic diversity.
629 *Ecology Letters*, *14*, 141–149.

630 Morrison, L.W. (2011) Why do some small islands lack vegetation? Evidence from long-term
631 introduction experiments. *Ecography*, *34*, 384–391.

632 Morrison, L. W. (2014). The small-island effect: empty islands, temporal variability and the
633 importance of species composition. *Journal of Biogeography*, *41*, 1007–1017.

634 Münkemüller, T., Gallien, L., Pollock, L.J., Barros, C., Carboni, M., Chalmandrier, L., Mazel,
635 F., Mokany, K., Roquet, C., Smyčka, J., Talluto, M.V., & Thuiller, W. (2020). Dos
636 and don'ts when inferring assembly rules from diversity patterns. *Global Ecology and*
637 *Biogeography*, DOI: 10.1111/geb.13098

638 Niering, W. A. (1956). Bioecology of Kapingamarangi Atoll. Caroline Islands: terrestrial
639 aspects. *Atoll Research Bulletin*, *49*, 1–32.

640 Niering, W. A. (1963). Terrestrial ecology of Kapingamarangi Atoll, Caroline Islands.
641 *Ecological Monographs*, *33*, 131–160.

642 Paine, C. E. T., Deasey, A., Duthie, A. B., & Thompson, K. (2018). Towards the general
643 mechanistic prediction of community dynamics. *Functional Ecology*, *32*, 1681–1692.

644 Panitsa, M. (1997). *Contribution to the knowledge of the flora and vegetation of the east*
645 *Aegean islets (Greece)*. Patras, Greece: University of Patras.

646 Panitsa, M., & Tzanoudakis, D. (2001). Contribution to the study of the Greek flora: flora and
647 phytogeography of Lipsos and Arki islet groups (East Aegean area, Greece). *Folia*
648 *Geobotanica*, *36*, 265–279.

649 R Core Team (2019). *R: a language and environment for statistical computing (Version 3.6)*.
650 Vienna, Austria: R foundation for statistical computing. Retrieved from
651 https://www.R-project.org/

652 Rechinger, K. H., & Rechinger-Moser, F. (1951). Phytogeographia Aegaea. In: *Akademie der*
653 *Wissenschaften Wien, Mathematische-Naturwissenschaftliche Klasse, Denkschrift,*
654 *105,* Wien: Springer.

655 Ricklefs, R. E., & Lovette, I. J. (1999). The roles of island area per se and habitat diversity in
656 the species–area relationships of four Lesser Antillean faunal groups. *Journal of*
657 *Animal Ecology, 68*, 1142–1160.

658 Runemark, H. (1971). The phytogeography of the central Aegean. *Opera Botanica*, *30*, 20–
659 28.

660 Schrader, J., Moeljono, S., Keppel, G., & Kreft, H. (2019). Plants on small islands revisited:
661   the effects of spatial scale and habitat quality on the species–area relationship.
662   *Ecography*, *42*, 1405–1414.

663 Schrader, J., König, C., Triantis, K. A., Trigas, P., Kreft, H., & Weigelt, P. (2020). Species–
664   area relationships on small islands differ among plant growth forms. *Global Ecology*
665   *and Biogeography, 29*, 814–829.

666 Sfenthourakis, S., & Triantis, K. A. (2009). Habitat diversity, ecological requirements of
667   species and the Small Island Effect. *Diversity and Distributions*, *15*, 131–140.

668 Sfenthourakis, S., Pafilis, P., Parmakelis, A., Poulakakis, N., & Triantis, K. (eds.) (2018).
669   *Biogeography and biodiversity of the Aegean*. Nikosia: Broken Hill Publishers.

670 Snogerup, S., & Snogerup, B. (2004). Changes in the flora of some Aegean islets 1968–2000.
671   *Plant Systematics and Evolution*, *245*, 169–213.

672 Srivastava, D. S., Cadotte, M. W., MacDonald, A. A. M., Marushia, R. G., Mirotchnick, N.,
673   & Mooers, A. (2012). Phylogenetic diversity and the functioning of ecosystems.
674   *Ecology Letters*, *15*, 637–648.

675 Strid, A. (2016). Atlas of the Aegean flora. *Englera*, *33*, 1–1578.

676 Swenson, N. G., & Enquist, B. J. (2009). Opposing assembly mechanisms in a Neotropical
677   dry forest: implications for phylogenetic and functional community ecology. *Ecology*,
678   *90*, 2161–2170.

679 Swenson, N. G., Enquist, B. J., Pither, J., Kerkhoff, A. J., Boyle, B., Weiser, M. D., et al.
680   (2012). The biogeography and filtering of woody plant functional diversity in North
681   and South America. *Global Ecology and Biogeography*, *21*, 798–808.

682 Toms, J. D., & Lesperance, M. L. (2003). Piecewise regression: a tool for identifying
683   ecological thresholds. *Ecology*, *84*, 2034–2041.

684 Triantis, K. A., & Sfenthourakis, S. (2012). Island biogeography is not a single-variable
685   discipline: the small island effect debate. *Diversity and Distributions*, *18*, 92–96.

686 Triantis, K. A., Mylonas, M., Weiser, M. D., Lika, K., & Vardinoyannis, K. (2005). Species
687   richness, environmental heterogeneity and area: a case study based on land snails in
688   Skyros archipelago (Aegean Sea, Greece). *Journal of Biogeography*, *32*, 1727–1735.

689 Triantis, K. A., Vardinoyannis, K., Tsolaki, E. P., Botsaris, I., Lika, K., & Mylonas, M.
690   (2006). Re-approaching the small island effect. *Journal of Biogeography*, *33*, 914–
691   923.

692 Tucker, C. M., Cadotte, M. W., Carvalho, S. B., Davies, T. J., Ferrier, S., Fritz, S. A.,
693   Grenyer, R., Helmus, M. R., Jin, L. S., Mooers, A. O., Pavoine, S., Purschke, O.,
694   Redding, D. W., Rosauer, D. F., Winter, M., & Mazel, F. (2017). A guide to
695   phylogenetic metrics for conservation, community ecology and macroecology.
696   *Biological Reviews*, *92*, 698–715.

697 Wang, Y., Millien, V., & Ding, P. (2016). On empty islands and the small-island effect. *Global Ecology and Biogeography*, *25*, 1333–1345.
698

699 Webb, C. O., Ackerly, D. D., McPeek, M. A., & Donoghue, M. J. (2002). Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, *33*, 475–505.
700

701 Weigelt, P., Daniel Kissling, W., Kisel, Y., Fritz, S. A., Karger, D. N., Kessler, M., Lehtonen, S., Svenning, J.-C., & Kreft, H. (2015). Global patterns and drivers of phylogenetic structure in island floras. *Scientific Reports*, *5*, 12213.
702
703

704 Whitehead, D. R., & Jones, C. E. (1969). Small islands and the equilibrium theory of insular biogeography. *Evolution*, *23*, 171–179.
705

706 Whittaker, R. J., & Fernández-Palacios, J. M. (2007). *Island biogeography: ecology, evolution, and conservation*, 2nd edn. Oxford: Oxford University Press.
707

708 Yu, J., Li, D., Zhang, Z., & Guo, S. (2020). Species–area relationship and small-island effect of bryophytes on the Zhoushan Archipelago, China. *Journal of Biogeography*, *47*, 978-992.
709
710

711 Zanne, A. E., Tank, D. C., Cornwell, W. K., Eastman, J. M., Smith, S. A., FitzJohn, R. G., et al. (2014). Three keys to the radiation of angiosperms into freezing environments. *Nature*, *506*, 89–92.
712
713

714

715 **BIOSKETCHES**

716 **Tom Matthews** is a macroecologist and biogeographer at the University of Birmingham, UK.

717 He is interested in the application of macroecological methods to global environmental

718 change questions, and his previous work has focused on the impacts of habitat fragmentation

719 and the form of the species–area relationship in fragmented landscapes.

720

721 **Francois Rigal** is an assistant professor at Université de Pau et des Pays de l'Adour. His

722 research interests include island biogeography, macroecology and functional ecology.

723

724 **Author contributions:** KAT, TJM, and FR designed the analytical study. KK and PT

725 collected the data. FR and TJM wrote the code, and FR analyzed the data. TJM and FR wrote

726 the first draft of the manuscript. All authors contributed to the final manuscript.

Table 1. Summary of the BIC-based model selection for habitat diversity, species richness, PD and the three effect sizes ES_{PD}, ES_{MPD} and ES_{MNTD} (with $\log_{10}$ Area as the predictor). Only results for the best model(s) are given. For each best model and each metric, the likelihood (LL), number of parameters (k), BIC value, ΔBIC, BIC weight (w_i), R^2 and the area threshold T_1 and, for the two breakpoint cases, T_2, are given. All best models were significant ($P<0.001$).

| Metrics	Best models	LL	k	BIC	ΔBIC	wi	R^2	T1 (km^2)	T2 (km^2)
Habitat diversity									
	Continuous one-threshold	-267.24	5	560.25	0	0.31	0.94	6.28	-
	Left-horizontal continuous one-threshold	-269.87	4	560.35	0.10	0.29	0.94	4.55	-
	Left-horizontal continuous two-thresholds	-264.88	6	560.68	0.43	0.25	0.94	3.45	143.88
	Continuous two-thresholds	-262.83	7	561.74	1.49	0.15	0.94	4.55	150.66
Species richness									
	Left-horizontal continuous two-thresholds	-1004.01	6	2038.94	0	0.77	0.95	0.67	34.52
PD									
	Continuous one-threshold	-1499.69	5	3025.14	0	0.53	0.95	2.28	-
	Continuous two-thresholds	-1494.76	7	3025.60	0.46	0.42	0.95	0.66	17.30
ES_{PD}									
	Continuous one-threshold	-250.28	5	526.39	0	0.89	0.38	13.43	-
ES_{MPD}									
	Continuous one-threshold	-223.54	5	472.85	0	0.91	0.34	6.14	-
ES_{MNTD}									
	Continuous one-threshold	-222.07	5	469.91	0	0.62	0.22	3.22	-
	Left-horizontal continuous two-thresholds	-220.24	6	471.39	1.48	0.30	0.24	0.14	0.74

731

732 **Table 2**. Results of the Mantel permutation tests implemented between the β_{SIM}, ES β_{SIM}, ES

733 MPD$_{DISS}$ and ES MNTD$_{DISS}$ dissimilarity matrices and the geographical distances between

734 islands. The analysis was implemented using all islands and for small and large islands

735 separately. The Pearson's correlation coefficients (r) and associated P-values (*P*) are given.

736 Significant results are highlighted in bold.

737

| Dissimilarity	All islands		Small islands		Large islands	
	r	*P*	*r*	*P*	*r*	*P*
β_{SIM}	-0.132	0.997	-0.012	0.427	**0.310**	**<0.001**
ES β_{SIM}	**0.442**	**<0.001**	**0.303**	**<0.001**	**0.615**	**<0.001**
ES MPD$_{DISS}$	0.018	0.368	-0.075	0.809	0.099	0.097
ES MNTD$_{DISS}$	**0.190**	**<0.001**	-0.055	0.834	0.052	0.140

738

739
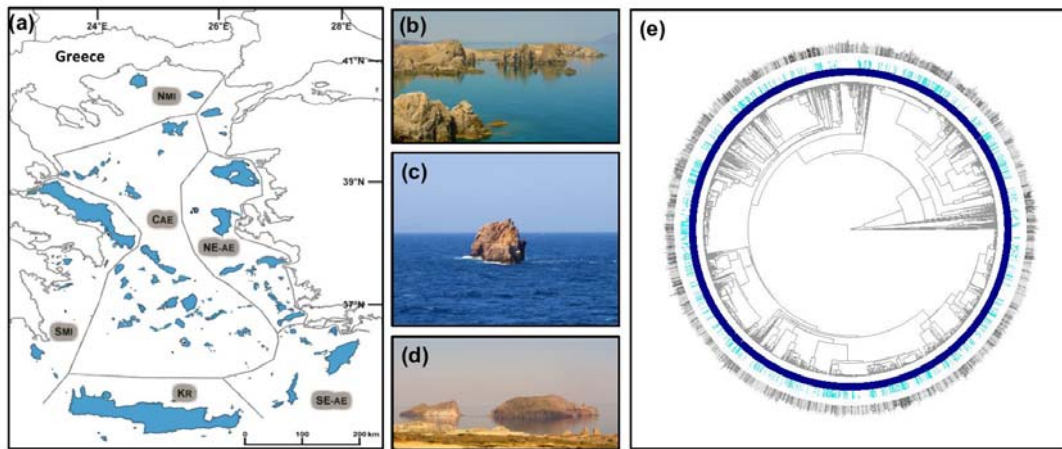
740

741 **FIGURE 1** (a) Map of the Aegean Sea islands. Major island groups are highlighted: NMI =

742 North Islands; CAE = Central Aegean; NE-AE = North East Aegean; SMI = South Islands;

743 KR = Kriti; and SE-AE = South Eastern Aegean. (b) Kampanes, Milos, (c) Portes, Paros, and

744 (d) Akradies, Milos. (e) Community phylogeny of the vascular plants of the 173 Aegean

745 Islands in our dataset. The occurrence of each species in large and small islands is indicated

746 by colours along the tips: light blue for small islands and dark blue for large islands. Tip

747 labels coloured black highlight species grafted randomly on the tree.

748

749

750
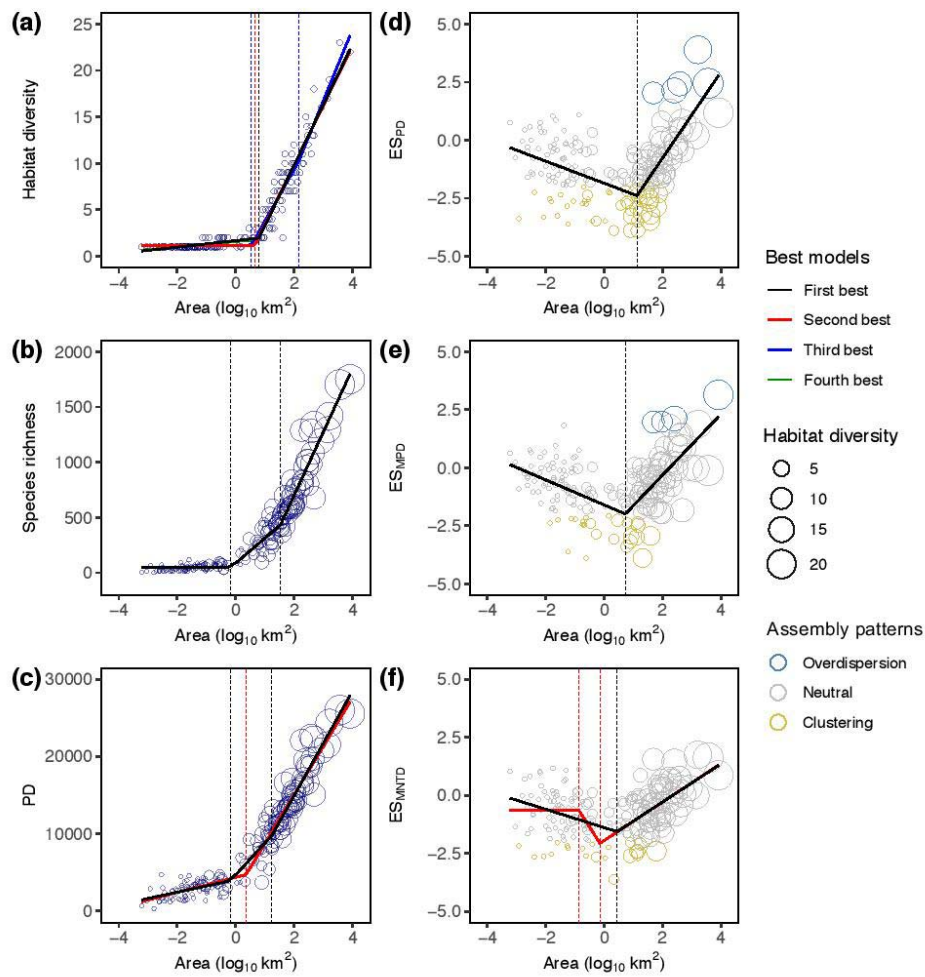
FIGURE 2 Relationships between $\log_{10}$ Area and habitat diversity (a), species richness (b), PD (c), ES_{PD} (d), ES_{MPD} (e) and ES_{MNTD} (f). Coloured circles indicate islands. Thick solid lines represent the predicted values of the best model(s) and dashed vertical lines represent the threshold(s) of the best model(s). Best models were selected using BIC-based model selection comparing four SIE piecewise models, a simple linear model and a null model. To illustrate the link between $\log_{10}$ Area and habitat diversity (HD) and the diversity relationships (b to f), the size of the dots were made proportional to the number of habitats. For the ES metrics, overdispersion indicates that values are significantly higher than expected by chance (ES >1.96) and clustering indicates that phylogenetic diversity values are significantly lower than expected by chance (ES < -1.96). Neutral indicates that values are neither lower nor higher than expected by chance (-1.96 < ES <1.96).

762

763
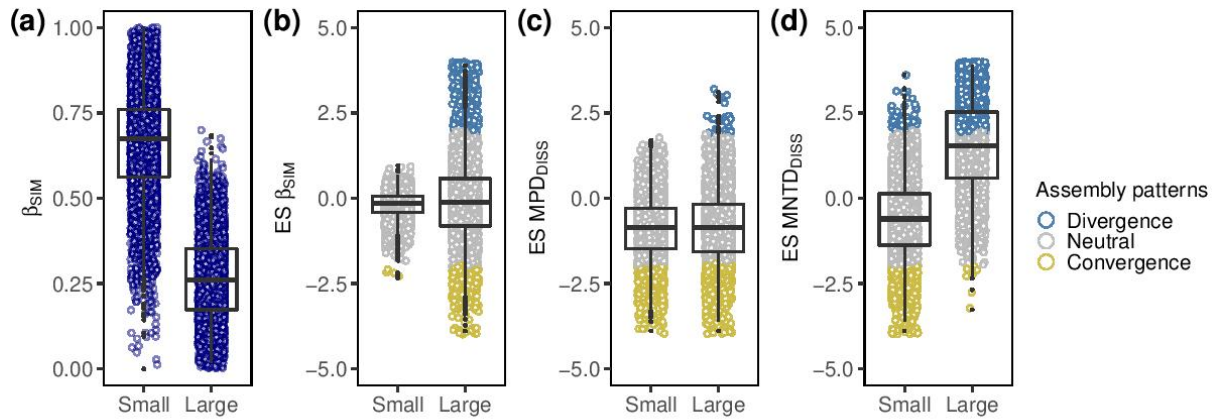
FIGURE 3 Differences in species turnover (a, b) and phylogenetic dissimilarity (c, d) between small and large islands. Islands were split into small and large using a breakpoint value of 2.57 km^2, calculated by averaging all breakpoints identified in the best models across all diversity metrics (see Figure 2b-f). Dots indicate the dissimilarity between pairs of islands. Species turnover was calculated using the index β_{SIM} (a) as well as its effect size ES β_{SIM} (b). Phylogenetic dissimilarity was calculated using ES MPD$_{DISS}$ (c) and ES MNTD$_{DISS}$ (d). Convergence indicates that species turnover / phylogenetic dissimilarity values are significantly lower than expected by chance (ES < -1.96), and divergence indicates values that are significantly higher than expected by chance (ES >1.96). Neutral indicates values that are neither lower nor higher than expected by chance (-1.96 < ES <1.96).