

Journal of Biogeography

Unravelling the small-island effect through phylogenetic community ecology

Journal:	<i>Journal of Biogeography</i>
Manuscript ID:	JB1-20-0040.R1
Manuscript Type:	Research Paper
Date Submitted by the Author:	15-Jun-2020
Complete List of Authors:	Matthews, Thomas; University of Birmingham, School of Geography, Earth and Environmental Sciences and the Birmingham Institute of Forest Research; Univ. dos Açores , Azorean Biodiversity Group Rigal, François; Environment and Microbiology Team, Université de Pau et des Pays de l'Amour, Université de Pau et des Pays de l'Amour Kougioumoutzis, Konstantinos; National and Kapodistrian University of Athens, Ecology and Systematics; Agricultural University of Athens, Agricultural University of Athens Trigas, Panayiotis; Agricultural University of Athens, Faculty of Crop Science Triantis, Kostas; National and Kapodistrian University of Athens, Department of Ecology and Taxonomy
Key Words:	Breakpoints, island biogeography, phylogenetic diversity, small-island effect, species–area relationship, beta-diversity

Unravelling the small-island effect through phylogenetic community ecology

Thomas J. Matthews* ^{1, 2+}, François Rigal ^{3, 2+}, Kostas Kougiumoutzis⁴, Panayiotis Trigas⁵ & Kostas A. Triantis*⁴

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

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

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

⁴Department of Ecology and Taxonomy, Faculty of Biology, National and Kapodistrian University of Athens, Athens GR-15701, Greece.

⁵Faculty 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

Abstract

Aim: The small-island effect (SIE) describes a different relationship between island area and species richness on smaller compared to larger islands. The pattern has recently gained widespread support. However, few studies have attempted to identify the actual mechanisms driving the SIE. Here, we use a phylogenetic community framework to study the SIE, based on the assumption that if the dominant assembly processes differ between small and large islands, patterns of phylogenetic community structure should shift across the area and habitat diversity gradient.

Location: The Aegean Archipelago, Greece.

Methods: We used a large dataset of 3262 vascular plant species distributed across 173 islands, in combination with a species-level phylogeny. The phylogenetic community structure of each island was calculated using a null modelling framework and was quantified using effect sizes (ES); negative values indicating phylogenetic clustering and positive values overdispersion. Habitat diversity, species richness, phylogenetic diversity (PD) and ES values were regressed against \log_{10} -transformed area and we tested for a SIE using piecewise regression models. We also assessed differences in taxonomic and phylogenetic composition between small and large islands using a beta-diversity framework.

Results: We found evidence of a SIE using species richness, PD and phylogenetic community structure (ES values). Small islands displayed low variation in habitat diversity and tended to be more phylogenetically clustered, while large islands shifted from phylogenetic clustering towards phylogenetic overdispersion with increasing area and habitat diversity. In addition, we showed that phylogenetic composition tended to be more similar between small islands than expected.

Main conclusion: Overall, our results provide an example of a SIE in the analysis of island phylogenetic community structure, and point to a role of habitat diversity in driving the SIE more generally.

Keywords: Breakpoints, island biogeography, phylogenetic diversity, small-island effect, species–area relationship, beta-diversity

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 Whittaker & 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., 2017; 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 driver of richness (Sfenthourakis & Triantis, 2009; Chen et al., 2020). For example, Niering (1963) postulated that the lack of mature soils on islands

smaller than 3.5 acres explained the SIE pattern he observed; it should be noted that Niering did not explicitly use the term SIE. Other proposed mechanisms explaining the SIE include low immigration rates on small islands (Morrison, 2011) and species-specific area requirements (Schrader et al., 2019), both of which prevent the successful establishment of species on small islands. However, despite all of these various proposed hypotheses, in general, studies attempting to identify the actual mechanisms driving the SIE in a given system are scarce (e.g. Morrison, 2011; Schrader et al., 2019).

In addition to the mechanisms generating the SIE, many other aspects of the SIE are still debated. For instance, there have been strong disagreements about the appropriate statistical approach for identifying SIEs (Dengler, 2010; Triantis & Sfenthourakis, 2012; Morrison, 2014; Wang et al., 2016). These disagreements include debates on whether or not to include islands with no species (Morrison, 2014; Wang et al., 2015), the most effective statistical model to detect a SIE (Lomolino & Weiser, 2001; Toms & Jaspers, 2003; Triantis et al., 2006; Matthews et al., 2014; Chisholm, Chung, Chimalakonda, & O'Dwyer, 2016; Gao, Cao, Xu, & Perry, 2019), and the best variable transformation procedure to use when fitting SIE models (Burns et al., 2009; Matthews et al., 2014). Furthermore, most SIE studies have focused on documenting patterns in species numbers without considering that species composition may help to elucidate the mechanisms underlying the SIE. For instance, island specialists (species that exclusively occur on cliffs, rocks and coastal habitats on typically small, low, uninhabited islets; Rechinger & Rechinger-Moser, 1951) are an important component of the flora of small islands in the Aegean Sea (Panitsa & Tzanoudakis, 2001; see Morrison, 2014, for a similar example from the flora of the Bahamas), but are largely absent from the larger islands, and thus their presence will likely influence any SIE patterns in these islands. In addition, SIE studies have only considered the number of species as the response variable, and thus assuming that all species are evolutionarily independent and ecologically equivalent. Critical information about primary community assembly mechanisms on islands might thus have been ignored.

Recently, ecologists and biogeographers have turned to phylogenetic and functional diversity-based metrics to provide novel insights into their study systems. Phylogenetic diversity (PD) incorporates the evolutionary relationships between species and provides a measure of the amount of evolutionary history that is represented by the species in an assemblage (Webb, Ackerly, McPeck, & Donoghue, 2002). Based on the assumption that phylogenetic relatedness between species is positively correlated with functional / ecological relatedness (an assumption

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

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

Overall then, we predict that, if habitat diversity is a primary driver of island richness, small islands are expected to support a limited subset of closely related species (i.e. PD clustering) and, due to the decoupling of island area and habitat diversity on small islands, the relationship between area and diversity (e.g. richness, PD; Triantis et al., 2006; Chen et al., 2020) will differ from that on large islands (i.e. the small island effect). In addition, if small islands contain similar habitat types, we should also expect the biota of these islands to be more phylogenetically similar (phylogenetic convergence) than expected by chance (Graham & Fine, 2008).

In contrast, larger islands should have higher habitat heterogeneity (Ricklefs & Lovette, 1999; Whittaker & Fernández-Palacios, 2007; Menchouakis & Triantis, 2009), allowing more distant lineages (that are adapted to the wider range of habitats) to establish, persist and diversify, and thus the species on large islands should be more spread across the phylogenetic tree than expected by chance (i.e. overdispersion). On larger islands, the biota may be either more or less phylogenetically similar (phylogenetic convergence or divergence) than expected, depending on, amongst other things, the level of habitat similarity between the islands, and / or the prevalence of *in situ* speciation (Graham & Fine, 2008).

To test the aforementioned predictions, we used a large and well sampled dataset of 3262 vascular plant species distributed across 173 islands and islets of varying size in the Aegean Sea (Greece). The phylogenetic community structure of each island was assessed using a species-level plant phylogeny and null models were implemented to test whether islands were more phylogenetically clustered or overdispersed than expected by chance. We used piecewise regression models to evaluate whether there were breakpoints in the ISAR (i.e. the traditional SIE), and in various types of island phylogenetic diversity–area relationships (IPDAR; i.e. a SIE in phylogenetic diversity). For the purposes of this study we define a SIE as a threshold point in island area where there is a statistically significant change in the slope of the diversity–area relationship. We also assessed whether species turnover and phylogenetic dissimilarity between islands were lower or higher than expected by chance using null models, and tested whether there were differences between small and large islands. Our study documents the first example of a SIE in the phylogenetic community structure of islands. In addition, our findings provide evidence suggesting a role of habitat diversity in driving the traditional SIE. Overall, we demonstrate how a phylogenetic community-based approach can be used to provide insights

into the potential mechanisms driving the SIE, and community assembly on islands more generally.

2. MATERIALS AND METHODS

2.1 Data collection and phylogeny construction

Building on the work of Kougiumoutzis et al. (2017), we used a comprehensive database of the presence-absence of plant species on 173 Aegean islands and islets (Figure 1a). We compiled our dataset from an extensive and detailed database for the Aegean islands (Appendix S1). Systematic research on the flora of the large Aegean Islands started in the late 1920s and has been undertaken continuously up to the present day. For the larger islands, the dataset comprises all available records with the exception of species that have been recorded only once in the distant past (i.e. their occurrence has not been confirmed with more recent surveys). Study of the floristic composition of the small islands (Figure 1b-d) was undertaken between 1970 and 2000. For the vast majority of the small islands, a single exhaustive plant sampling trip was conducted during this period, and these data are used in the database. Island species richness ranged from two (Islet East Aspruni) to 1751 species (Crete). In total, the dataset contains 3262 native vascular plant species (all non-native taxa were excluded). The chorological status of the species follows Dimopoulos et al. (2012) and Strid (2016). The area (km²) of each of the 173 islands and islets was sourced from the literature (Panitsa, 1997; Sfenthourakis & Triantis, 2009) and ranged from 0.0006 km² for Touzeta islet to 8270 km² for Crete.

Coarse-scale habitat type information for each island was calculated using the CORINE Land Cover (CLC) inventory (Copernicus Land Monitoring Service, 2018). The habitat diversity of each island was calculated as the total number of CLC classes present (out of a total of 44 classes). The CLC has a resolution of 100 m, which is relatively coarse in comparison to the size of the smallest islands / islets in our dataset. As such, for these smaller islands, we used a combination of satellite imagery, photographs, expert knowledge and field data from previous work (e.g. Panitsa, 1997; Panitsa & Tzanoudakis, 2001; Sfenthourakis et al., 2018) to 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 km²).

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 clustering). Values greater than 1.96 (probit of roughly 0.975) or less than -1.96 (probit of roughly 0.025) were considered to be significantly greater or less than expected, respectively. This calculation of effect sizes was preferred to the commonly-used standardized effect size (SES) due to the non-Gaussian shape of some of our

252 null distributions (Bernard-Verdier et al., 2012).

253

254 2.3 Detection of a small-island effect

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

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

279 All models were fitted using ordinary least square regression (OLS), and threshold values were
 280 estimated using the minimum residual sum of squares (RSS) method (see electronic
 281 supplementary material, Appendix S5). The functions for fitting and plotting the threshold
 282 models, along with additional piecewise model functions, were integrated into the 'sars' R
 283 package (Version *; available on GitHub) (Matthews et al., 2019). For each diversity metric,

we compared the different models using the Bayesian Information Criterion (BIC; Burnham & Anderson, 2002). In each case, the best model was considered to be that with the smallest BIC; however, all models with a ΔBIC value < 2 were considered as having a similar degree of statistical support (Burnham & Anderson, 2002). BIC was preferred to other information criteria because, as ES values tend to be more scattered than species richness values, spurious relationships might simply arise due to random noise in the data and be incorrectly interpreted as legitimate relationships. BIC is more effective in such situations due to the greater penalization in respect to the number of model parameters (Link & Barker, 2006). We also tested the ability of the best model(s) to account for potential spatial structure by estimating Moran's I spatial autocorrelation index using the model residuals and the latitude and longitude of each island's centroid. Diagnostic plots were inspected in order to detect residual deviations from normality and homoscedasticity, and to identify outliers. Finally, we repeated the aforementioned model fitting and model comparison analyses using habitat diversity instead of island area as the predictor of the five diversity metrics.

2.4 Assessing differences in species turnover and phylogenetic dissimilarity between small and large islands

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

Phylogenetic dissimilarity was measured in an analogous way to the measurement of phylogenetic community structure. For a given pair of islands A and B, we calculated the mean

1
2
3 317 of all phylogenetic distances separating species occurring in A from species occurring in B
4 318 (MPD_{DISS}), and the average of the phylogenetic distance between each species in A and its
5
6 319 closest relative in B ($MNTD_{DISS}$). We computed the ES of MPD_{DISS} and $MNTD_{DISS}$ values
7
8 320 using the same approach as for ES_{PD} , ES_{MPD} and ES_{MNTD} .
9

10 321
11 322 We further tested for differences in mean β_{SIM} and ES between small and large islands using
12
13 323 Wilcoxon Rank Sum tests, and whether the distribution of ES values within small and large
14
15 324 islands significantly deviated from the null expectation (median = 0) using a one-sample
16
17 325 Wilcoxon test (Swenson & Enquist, 2009). Since islands in the Aegean are spread across a large
18
19 326 geographical region i.e. the Aegean Sea, differences in turnover and phylogenetic dissimilarity
20
21 327 between small and large islands could also be driven simply by geographical distance (i.e.
22
23 328 dispersal limitation). For example, if higher turnover is found between small islands compared
24
25 329 to large islands, this simply could be due to the fact that small islands are relatively far away
26
27 330 from each other and not because they differ in terms of the habitat types they contain. β_{SIM} , ES
28
29 331 β_{SIM} , ES MPD_{DISS} and ES $MNTD_{DISS}$ dissimilarity matrices were therefore correlated with
30
31 332 geographical distance between islands using a Mantel permutation test, based on 9,999
32
33 333 permutations. This was implemented for all islands and for small and large islands separately.

33 334 **2.5 Complementary analyses**

34
35 335 We undertook three complementary analyses to ensure our results were robust. First, we
36
37 336 checked the sensitivity of our results to the method used for placing unresolved genera and
38
39 337 species in the phylogeny. Second, we re-ran the above analyses on two smaller subsets of
40
41 338 islands, namely North-Eastern Aegean (83 islands and islets) and the Central Aegean (60
42
43 339 islands and islets) (Kougioumoutzis et al., 2017; Appendix S6). Third, we performed a
44
45 340 sensitivity test to ensure that we did not generate spurious small-island effects because of our
46
47 341 decisions regarding data transformation using the analytical procedure of Burns et al. (2009)
48
49 342 (Appendix S7). All analyses were undertaken using R (version 3.6; R Core Team, 2019).
50

51
52 344 **3 RESULTS**

53
54
55 345 **3.1 The small-island effect: habitat diversity, ISAR, the IPDAR and phylogenetic**
56 346 **community structure**

57
58 347 There was some model uncertainty regarding the relationship between habitat diversity and
59
60 348 \log_{10} Area, with the left-horizontal continuous one-threshold and more complex continuous

two-threshold models 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 km² (Figure 2a and Table 1). On small islands there was a slow, or an absence of an, increase in habitat diversity with log₁₀ 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 MPDAR, 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 MPDAR with R^2 values ranging from 38% for SES_{PD} to 22% SES_{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 linear, with habitat diversity explaining 93% of the variation in species richness (Table S2 and Figure S2a in Appendix S8). A similar pattern was found for PD although the relationship included a first phase of faster increase (Table S2 and Figure S2b). For the ES_{PD} and ES_{MPD} metrics, the continuous one-threshold model was found to be the best model (Table S2 and Figure S2a), with the slope of the first segment negative and the slope of the second segment positive; the threshold being at 2.9 and 3.9 habitats respectively. For ES_{MNTD} , the left-horizontal

one-threshold model was found to be best with a threshold identified at 6.7 habitats (Table S2 and Figure S2a).

3.2 Comparing species turnover and phylogenetic dissimilarity between small and large islands

We classified islands as small or large using a common threshold of 2.57 km² (log10-transformed value = 0.41; corresponding to 92 small and 81 large islands), calculated by averaging all thresholds identified in the best models across all metrics (range in threshold values across the best models: 0.14 km²–34.52 km²). This resulted in quite a clear split between small and large islands (see Figure S3 in Appendix S8). Overall, among the 3262 species recorded in our dataset, 640 occurred on small islands with three being found exclusively on small islands. Differences in habitat composition between small islands was very low, with most small islands containing either natural grasslands and/or sclerophyllous vegetation (Appendix S2). In contrast, large islands displayed substantial differences in habitat composition, with a mean habitat diversity of 8.1 (SD = 4.7) out of a total of 27 habitats present on large islands.

Observed species turnover (β_{SIM}) was significantly higher between small (mean $\beta_{\text{SIM}} = 0.66$) than between large (mean $\beta_{\text{SIM}} = 0.24$) islands (Wilcoxon test $P < 0.001$, Figure 3a). ES β_{SIM} was significantly lower between small than between large islands (mean ES $\beta_{\text{SIM}} = -0.20$ and -0.09 respectively, Wilcoxon test $P < 0.001$, Figure 3b). For small and large islands, mean ES β_{SIM} was significantly lower than 0 (one-sample Wilcoxon test $P < 0.001$). However, most of the pairwise ES β_{SIM} values for small islands did not show any significant deviation from random (99.9% of the pairwise SES β_{SIM} values were neutral, Figure 3), while for large islands a few pairs were identified as being significantly phylogenetically convergent or divergent (7.2% and 5.9% respectively, Figure 3b). For β_{SIM} , a significant and positive correlation with geographical distance was found only for large islands, while for ES β_{SIM} , a significant and positive correlation with geographical distance was found for all islands and for small and large islands separately (Table 2).

In regard to phylogenetic composition, both small and large islands exhibited significant average phylogenetic convergence for ES MPD_{DISS} (one-sample Wilcoxon test $P < 0.001$; mean ES MPD_{DISS} = -0.91 and -0.87 respectively, Figure 3c), and ES MPD_{DISS} values did not differ between small and large islands (Wilcoxon test $P = 0.64$). For ES MNTD_{DISS}, small islands

exhibited significant average convergence, while large islands exhibited significant average divergence (one-sample Wilcoxon test $P < 0.001$ for both size-groups; $ES_{MNTD_{DISS}} = -0.63$ and 1.54 , respectively) (Figure 3d). There was no significant correlation between either $ES_{MPD_{DISS}}$ or $ES_{MNTD_{DISS}}$ and geographical distance, for either small or large islands (Table 2).

3.3 Complementary analyses

We found that the values of phylogenetic diversity and dissimilarity were not sensitive to the method used for placing unresolved genera and species in the phylogeny (Table S3 in Appendix S8). Results obtained for the North-Eastern Aegean and Central Aegean subsets are given in Appendix S6 and Tables S4-S5 and Figures S4-S8 in Appendix S8. Overall, results obtained for the North-Eastern Aegean and the Central Aegean were very similar to those obtained using all islands; the main exception was for the Central Aegean, where no SIE was detected when using the phylogenetic ES_{PD} and ES_{MNTD} metrics. Using the approach of Burns et al. (2009) we found that the presence of a SIE in the relationship between island area and species richness is not simply an artefact of our use of a semi-log transformation (Figure S9 in Appendix S8).

4 DISCUSSION

4.1 Phylogenetic community structure and the spatial island effect

We have found evidence of a SIE in the island species–area relationship, the island phylogenetic diversity–area relationship and when using three phylogenetic community metrics (ES_{PD} , ES_{MPD} and ES_{MNTD}). Interestingly, although there was some degree of model uncertainty and the thresholds estimated from the one and two-threshold models were slightly different, all thresholds (across all response variables) fall within a range of 34.38 km^2 (from 0.14 to 34.52 km^2) representing no more than 0.4% of the total area range found in our data (8270 km^2). This suggests that the abrupt shifts in the various diversity metrics were more or less synchronised, and could reflect that it is around this area range that the diversity of habitats starts to increase with area in a linear fashion.

In relation to ES_{PD} , ES_{MPD} and the ES_{MNTD} , we found that, for the smaller islands (i.e. those to the left of the breakpoint), island communities tended to be more phylogenetic clustered as area increased (i.e. a negative slope – discussed in the next paragraph, Figure 2d,e,f). We also observed that within the SIE range, area and habitat diversity are decoupled (see also Triantis

et al., 2006), with habitat diversity being almost stable throughout the SIE range. If habitat diversity is the main driver of diversity rather than area per se, one possible explanation for our observation of phylogenetic clustering on smaller islands is greater habitat filtering, where only species being able to establish in the limited habitat types accumulate, leading to phylogenetic redundancy (species drawn from a limited array of clades) within the species assemblages. It is worth noting that small islands mostly contain one or both of natural grasslands and sclerophyllous vegetation as land-cover classes (Appendix S2). This result corroborates an earlier study assessing the habitats of a number of small islets in the Aegean (including various islands included in our dataset) where it was found that many major habitat types, such as mature forest and freshwater, were missing (Snogerup & Snogerup, 2004; see also Panitsa, 1997; Sfenthourakis et al., 2018). However, these findings were not discussed in the context of the ISAR or the SIE (Snogerup & Snogerup, 2004).

Interestingly, we did not identify phylogenetic clustering when focusing just on the very small islands (i.e. the smallest of the islands we have classified here as small) (see Figure 2d,e,f). The above habitat filtering explanation may explain this observation. However, this could also be due to the species present on the smallest islands being adapted (e.g. possessing salt tolerance) to littoral habitats, that are prevalent on very small islands (Snogerup & Snogerup, 2004). If salt tolerance is distributed across several families, this would explain the random patterns observed in the phylogenetic structure of these islets (Flowers, Galal & Eromham, 2010). As small island area increases, 'phrygana' communities primarily comprising annual grasses and members of few other families start to dominate, leading to phylogenetic clustering. In addition, it could be that on the very smallest islets natural disturbance (e.g. intense storms and/or waves) have more of an impact (Whittaker & Fernández-Palacios, 2007), driving community assembly to be essentially random with regard to phylogenetic structure. Alternatively, it could be that, for the small islands, another variable co-varies with area, such as human disturbance. For example, the smallest islets may be unsuitable for human activities (e.g. tourism or land for grazing), meaning that there is in fact a greater amount of certain types of disturbance (and reduction in habitat heterogeneity) on the larger of the smaller islands, leading to the patterns observed.

As island area increased, we observed a gradual shift from phylogenetic clustering toward phylogenetic overdispersion on individual islands, which mirrored a clear increase in habitat diversity. One interpretation of this finding is that, on larger islands, the greater range of

habitats, such as freshwater systems, sub-alpine ecosystems and mature forest, enables a much wider variety of functional forms drawn from a broader range of clades to persist. This will result in greater phylogenetic diversity (and its ES) on the larger islands.

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

A shortcoming of previous SIE studies has been the focus on species numbers, without any consideration of species composition (Morrison, 2014). We found striking differences between small and large islands in regard to species turnover: species spatial turnover between small islands was high (Figure 3a) and mostly random (ES β_{SIM} , Figure 3b). The strong correlation between ES β_{SIM} and geographical distance also suggests that neutral processes such as dispersal limitation might have played an important role in shaping species composition on small islands in the Aegean (see Runemark, 1971). In contrast, phylogenetic composition tended to be, on average, more similar on small islands than expected by chance, and independent of geographical distance, supporting our hypothesis of possible deterministic convergence regarding the main plant clades present on small islands. **This is also reinforced by the small differences in habitat composition between the small islands, almost all comprising only natural grasslands and/or sclerophyllous vegetation.** Altogether, this indicates that, while only certain clades are able to establish and survive on the smaller islands, the identity of these taxa (i.e. the identity of the taxa within these clades) is more randomly distributed. This is likely due, at least in part, to stochastic extinction / immigration and / or dispersal. That is, both neutral and niche-based processes might operate in combination in this system to drive the SIE.

On the larger islands, much lower spatial turnover was observed, implying that species composition on the larger islands was more ordered. Patterns of phylogenetic composition on the large islands were more complex, with our two metrics showing opposite results: ES MPD_{DISS} identifying mostly convergence and ES $MNTD_{DISS}$ mostly divergence. ES MPD_{DISS} is a measure of phylogenetic dissimilarity that captures variation associated with basal nodes (e.g. substitutions of a given order or even higher taxonomic ranks; Kembel et al., 2010) and thus showed an overall similarity between large islands due to the fact most of the higher taxonomic groups represented in the regional island species pool occur on most of the large islands. In contrast, ES $MNTD_{DISS}$ focuses on variation associated with terminal nodes and is more sensitive to lower-level taxonomic substitutions (i.e. changes in representation of families/genera) among communities (Kembel et al., 2010). Therefore, the divergence reported between large islands with ES $MNTD_{DISS}$ could be due to differences in the source pools

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

4.3 Caveats and future directions

This study is not the first to find breakpoints in the ISAR (Matthews et al., 2014; Niering, 1963; Wang et al., 2016), or to postulate that the mechanisms driving the ISAR (and SIE) may change across breakpoints in the relationship (e.g. see Triantis et al. 2006 and Losos & Parent, 2010, for a discussion on the speciation–area relationship). However, to our knowledge this is the first assessment of the SIE undertaken through the lens of phylogenetic community ecology, and the approach has proven useful for shining light on the potential mechanisms driving the SIE. However, as always, in the absence of controlled experiments and trait data (discussed below) and more explicit fine-scale measurements of habitat diversity, it is impossible to be certain that habitat availability is the primary driver of the SIE in this system. Nonetheless, based on our results we would argue that the approach applied here (i.e. combining evaluation of the ISAR, IPDAR and species and phylogenetic composition) represents a useful framework for analysing the SIE and the underlying mechanisms, and assembly processes on islands more generally.

Going forward it would be interesting to expand this approach to include functional diversity (e.g. see Schrader et al., 2020, for an interesting start in this direction.). An assumption of the use of phylogenetic diversity in these types of studies is that there is a correlation between phylogenetic diversity and functional diversity. However, while a correlation between phylogenetic diversity and functional diversity may exist in many systems, it is unlikely ever to be a perfect association (Flynn et al., 2011). While many traits have been shown to have strong phylogenetic signal, such that related species have similar traits, some have not (Srivastava et al., 2012), and arguably functional trait data are better than phylogenetic data in community ecology studies of this nature (Paine, Deasey, Duthie, & Thompson, 2018). Thus, the expansion of the framework presented here to include functional diversity will no doubt increase our understanding of the SIE.

ACKNOWLEDGEMENTS

An International Travel Fund Grant from the College of Life and Environmental Sciences (University of Birmingham) allowed TJM to visit the University of Athens to work on the study, and F. Guilhaumon kindly hosted TJM during the writing of the paper.

DATA ACCESSIBILITY

The data are provided in the Supporting information.

REFERENCES

- Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, *19*, 134–143.
- Bernard-Verdier, M., Navas, M. L., Vellend, M., Violle, C., Fayolle, A., & Garnier, E. (2012). Community assembly along a soil depth gradient: contrasting patterns of plant trait convergence and divergence in a Mediterranean rangeland. *Journal of Ecology*, *100*, 1422–1433.
- Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multi-model inference: a practical information-theoretic approach*, 2nd edn. New-York: Springer.
- Burns, K. C., Paul McHardy, R., & Pledger, S. (2009). The small-island effect: fact or artefact? *Ecography*, *32*, 269–276.
- Chen, C., Yang, X., Tan, X., & Wang, L. (2020). The role of habitat diversity in generating the small-island effect. *Ecography*, DOI: 10.1111/ecog.05092.
- Chisholm, R. A., Fung, T., Chimala-Anderson, D., & O'Dwyer, J. P. (2016). Maintenance of biodiversity on islands. *Proceedings of the Royal Society B: Biological Sciences*, *283*, 20160102.
- Copernicus Land Monitoring Service (2018) *CLC 2018*, version v.20_20u1. URL: <https://land.copernicus.eu/pan-european/corine-land-cover/clc2018?tab=download>. © European Union, Copernicus Land Monitoring Service 2018 European Environment Agency (EEA).
- Dengler, J. (2010). Robust methods for detecting a small island effect. *Diversity and Distributions*, *16*, 256–266.
- Dimopoulos, P., Raus, T., Bergmeier, E., Constantinidis, T., Iatrou, G., Kokkini, S., Strid, A., & Tzanoudakis, D. (2013). *Vascular plants of Greece. An annotated checklist*. Berlin: Botanic Garden and Botanical Museum Berlin-Dahlem.
- Emerson, B. C., & Gillespie, R. G. (2008). Phylogenetic analysis of community assembly and structure over space and time. *Trends in Ecology & Evolution*, *23*, 619–630.
- Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation*, *61*, 1–10.
- Flowers, T. J., Galal, H. K., & Bromham, L. (2010). Evolution of halophytes: multiple origins of salt tolerance in land plants. *Functional Plant Biology*, *37*, 604–612.
- Flynn, D. F. B., Mirotchnick, N., Jain, M., Palmer, M. I., & Naeem, S. (2011). Functional and phylogenetic diversity as predictors of biodiversity–ecosystem-function relationships. *Ecology*, *92*, 1573–1581.

- Gao, D., Cao, Z., Xu, P., & Perry, G. (2019). On piecewise models and species–area patterns. *Ecology and Evolution*, 9, 8351–8361.
- Gotelli, N. J. (2000). Null model analysis of species co-occurrence patterns. *Ecology*, 81, 2606–2621.
- Graham, C. H., & Fine, P. V. A. (2008). Phylogenetic beta diversity: linking ecological and evolutionary processes across space in time. *Ecology Letters*, 11, 1265–1277.
- Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., Blomberg, S. P., & Webb, C. O. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, 26, 1463–1464.
- Kougioumoutzis, K., Valli, A. T., Georgopoulou, E., Simaiakis, S. M., Triantis, K. A., & Trigas, P. (2017). Network biogeography of a complex island system: the Aegean Archipelago revisited. *Journal of Biogeography*, 44, 651–660.
- Lhotsky, B., Kovács, B., Ónodi, G., Csecserits, A., Rédei, T., Lengyel, A., Miklos K., & Botta-Dukát, Z. (2016). Changes in assembly rules along a stress gradient from open dry grasslands to wetlands. *Journal of Ecology*, 104, 507–517.
- Link, W. A., & Barker, R. J. (2006). Model weights and the foundations of multimodel inference. *Ecology*, 87, 2625–2635.
- Lomolino, M. V., & Weiser, M. D. (2001). Towards a more general species–area relationship: diversity on all islands, great and small. *Journal of Biogeography*, 28, 431–445.
- Losos, J. B., & Parent, C. E. (2010). The speciation–area relationship. In J. B. Losos & R. E. Ricklefs (Eds.), *The theory of island biogeography revisited* (pp. 415–438). Princeton: Princeton University Press.
- MacArthur, R. H., & Wilson, E. O. (1967). *The theory of island biogeography*. Princeton: Princeton University Press.
- Matthews, T. J., Steinbauer, M. J., Tzirkalli, E., Triantis, K. A., & Whittaker, R. J. (2014). Thresholds and the species–area relationship: a synthetic analysis of habitat island datasets. *Journal of Biogeography*, 41, 1018–1028.
- Matthews, T. J., Triantis, K., Whittaker, R. J., & Guilhaumon, F. (2019). sars: an R package for fitting, evaluating and comparing species–area relationship models. *Ecography*, 42, 1446–1455.
- Morlon, H., Schwilk, D. W., Bryant, J. A., Marquet, P. A., Rebelo, A. G., Tauss, C., Bohannan, B. J. M., & Green, J. L. (2011). Spatial patterns of phylogenetic diversity. *Ecology Letters*, 14, 141–149.
- Morrison, L.W. (2011) Why do some small islands lack vegetation? Evidence from long-term introduction experiments. *Ecography*, 34, 384–391.
- Morrison, L. W. (2014). The small-island effect: empty islands, temporal variability and the importance of species composition. *Journal of Biogeography*, 41, 1007–1017.

- Münkemüller, T., Gallien, L., Pollock, L.J., Barros, C., Carboni, M., Chalmandrier, L., Mazel, F., Mokany, K., Roquet, C., Smyčka, J., Talluto, M.V., & Thuiller, W. (2020). Dos and don'ts when inferring assembly rules from diversity patterns. *Global Ecology and Biogeography*, DOI: 10.1111/geb.13098
- Niering, W. A. (1956). Bioecology of Kapingamarangi Atoll. Caroline Islands: terrestrial aspects. *Atoll Research Bulletin*, 49, 1–32.
- Niering, W. A. (1963). Terrestrial ecology of Kapingamarangi Atoll, Caroline Islands. *Ecological Monographs*, 33, 131–160.
- Paine, C. E. T., Deasey, A., Duthie, A. B., & Thompson, K. (2018). Towards the general mechanistic prediction of community dynamics. *Functional Ecology*, 32, 1681–1692.
- Panitsa, M. (1997). *Contribution to the knowledge of the flora and vegetation of the east Aegean islets (Greece)*. Patras, Greece: University of Patras.
- Panitsa, M., & Tzanoudakis, D. (2001). Contribution to the study of the Greek flora: flora and phytogeography of Lipsos and Arki islet groups (East Aegean area, Greece). *Folia Geobotanica*, 36, 265–271.
- R Core Team (2019). *R: a language and environment for statistical computing (Version 3.6)*. Vienna, Austria: R foundation for statistical computing. Retrieved from <https://www.R-project.org/>
- Rechinger, K. H., & Rechinger-Moser, F. (1951). Phytogeography of Aegaea. In: *Akademie der Wissenschaften Wien, Mathematisch-Naturwissenschaftliche Klasse, Denkschrift*, 105, Wien: Springer.
- Ricklefs, R. E., & Lovette, I. J. (1999). The roles of island area per se and habitat diversity in the species–area relationships of four Lesser Antillean faunal groups. *Journal of Animal Ecology*, 68, 1142–1160.
- Runemark, H. (1971). The phytogeography of the central Aegean. *Opera Botanica*, 30, 20–28.
- Schrader, J., Moeljono, S., Keppel, G., & Kreft, H. (2019). Plants on small islands revisited: the effects of spatial scale and habitat quality on the species–area relationship. *Ecography*, 42, 1405–1414.
- Schrader, J., König, C., Triantis, K. A., Trigas, P., Kreft, H., & Weigelt, P. (2020). Species–area relationships on small islands differ among plant growth forms. *Global Ecology and Biogeography*, 29, 814–829.
- Sfenthourakis, S., & Triantis, K. A. (2009). Habitat diversity, ecological requirements of species and the Small Island Effect. *Diversity and Distributions*, 15, 131–140.
- Sfenthourakis, S., Pafilis, P., Parmakelis, A., Poulakakis, N., & Triantis, K. (eds.) (2018). *Biogeography and biodiversity of the Aegean*. Nikosia: Broken Hill Publishers.
- Snogerup, S., & Snogerup, B. (2004). Changes in the flora of some Aegean islets 1968–2000. *Plant Systematics and Evolution*, 245, 169–213.

- 662 Srivastava, D. S., Cadotte, M. W., MacDonald, A. A. M., Marushia, R. G., Mirotchnick, N., &
663 Mooers, A. (2012). Phylogenetic diversity and the functioning of ecosystems. *Ecology*
664 *Letters*, 15, 637–648.
- 665 Strid, A. (2016). Atlas of the Aegean flora. *Englera*, 33, 1–1578.
- 666 Swenson, N. G., & Enquist, B. J. (2009). Opposing assembly mechanisms in a Neotropical dry
667 forest: implications for phylogenetic and functional community ecology. *Ecology*, 90,
668 2161–2170.
- 669 Swenson, N. G., Enquist, B. J., Pither, J., Kerkhoff, A. J., Boyle, B., Weiser, M. D., et al. (2012).
670 The biogeography and filtering of woody plant functional diversity in North and South
671 America. *Global Ecology and Biogeography*, 21, 798–808.
- 672 Toms, J. D., & Lesperance, M. L. (2003). Piecewise regression: a tool for identifying ecological
673 thresholds. *Ecology*, 84, 2034–2041.
- 674 Triantis, K. A., & Sfenthourakis, S. (2012). Island biogeography is not a single-variable
675 discipline: the small island effect debate. *Diversity and Distributions*, 18, 92–96.
- 676 Triantis, K. A., Mylonas, M., Weiser, M. D., Lika, K., & Vardinoyannis, K. (2005). Species
677 richness, environmental heterogeneity and area: a case study based on land snails in
678 Skyros archipelago (Aegean Sea, Greece). *Journal of Biogeography*, 32, 1727–1735.
- 679 Triantis, K. A., Vardinoyannis, K., Tselaki, E. P., Botsaris, J., Lika, K., & Mylonas, M. (2006).
680 Re-approaching the small island effect. *Journal of Biogeography*, 33, 914–923.
- 681 Tucker, C. M., Cadotte, M. W., Carvalho, S. B., Davies, T. J., Fennel, S., Fritz, S. A., Grenyer,
682 R., Helmus, M. R., Jin, L. S., Mooers, A. O., Pavoine, S., Pärtel, O., Pärtel, O., Redding,
683 D. W., Rosauer, D. F., Winter, M., & Mazel, F. (2017). A guide to phylogenetic metrics
684 for conservation, community ecology and macroecology. *Biological Reviews*, 92, 698–
685 715.
- 686 Wang, Y., Millien, V., & Ding, P. (2016). On empty islands and the small-island effect. *Global*
687 *Ecology and Biogeography*, 25, 1333–1345.
- 688 Webb, C. O., Ackerly, D. D., McPeck, M. A., & Donoghue, M. J. (2002). Phylogenies and
689 community ecology. *Annual Review of Ecology and Systematics*, 33, 475–505.
- 690 Weigelt, P., Daniel Kissling, W., Kisel, Y., Fritz, S. A., Karger, D. N., Kessler, M., Lehtonen,
691 S., Svenning, J.-C., & Kreft, H. (2015). Global patterns and drivers of phylogenetic
692 structure in island floras. *Scientific Reports*, 5, 12213.
- 693 Whitehead, D. R., & Jones, C. E. (1969). Small islands and the equilibrium theory of insular
694 biogeography. *Evolution*, 23, 171–179.
- 695 Whittaker, R. J., & Fernández-Palacios, J. M. (2007). *Island biogeography: ecology, evolution,*
696 *and conservation*, 2nd edn. Oxford: Oxford University Press.
- 697 Yu, J., Li, D., Zhang, Z., & Guo, S. (2020). Species–area relationship and small-island effect
698 of bryophytes on the Zhoushan Archipelago, China. *Journal of Biogeography*, 47, 978–
699 992.

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.

BIOSKETCHES

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

Francois Rigal is an assistant professor at Université de Pau et des Pays de l'Adour. His research interests include island biogeography, macroecology and functional ecology.

Author contributions: KAT, TJM and FR designed the analytical study. KK and PT collected the data. FR and TJM wrote the code, and FR analyzed the data. TJM and FR wrote 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	w_i	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	-262.55	7	561.74	1.49	0.15	0.94	4.55	150.66
Species richness									
	Left-horizontal continuous two thresholds	-1004.01	6	2035.94	0	0.77	0.95	0.67	34.52
PD									
	Continuous one-threshold	-1499.59	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	526.89	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

720

For Peer
Review

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 2. Results of the Mantel permutation tests implemented between the β_{SIM} , ES β_{SIM} , ES MPD_{DISS} and ES MNTD_{DISS} dissimilarity matrices and the geographical distances between islands. The analysis was implemented using all islands and for small and large islands separately. The Pearson’s correlation coefficients (r) and associated P-values (P) are given. Significant results are highlighted in bold.

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

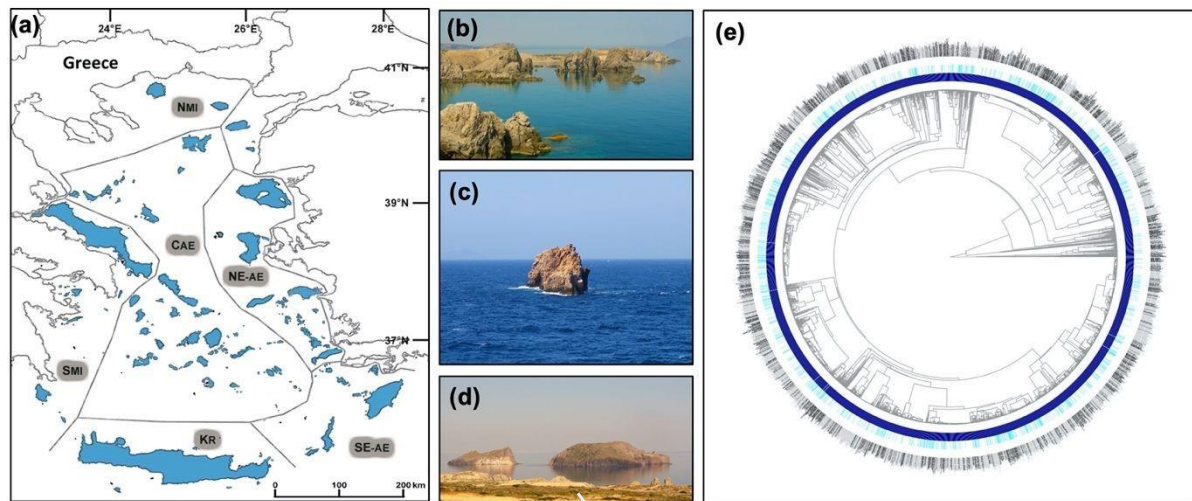


FIGURE 1 (a) Map of the Aegean Sea islands. Major island groups are highlighted: NMI = North Islands; CAE = Central Aegean; NE-AE = North East Aegean; SMI = South Islands; KR = Kriti; and SE-AE = South Eastern Aegean. (b) Kampanes, Milos, (c) Portes, Paros, and (d) Akradies, Milos. (e) Community phylogeny of the vascular plants of the 173 Aegean Islands in our dataset. The occurrence of each species in large and small islands is indicated by colours along the tips: light blue for small islands and dark blue for large islands. Tip labels coloured black highlight species grafted randomly on the tree.

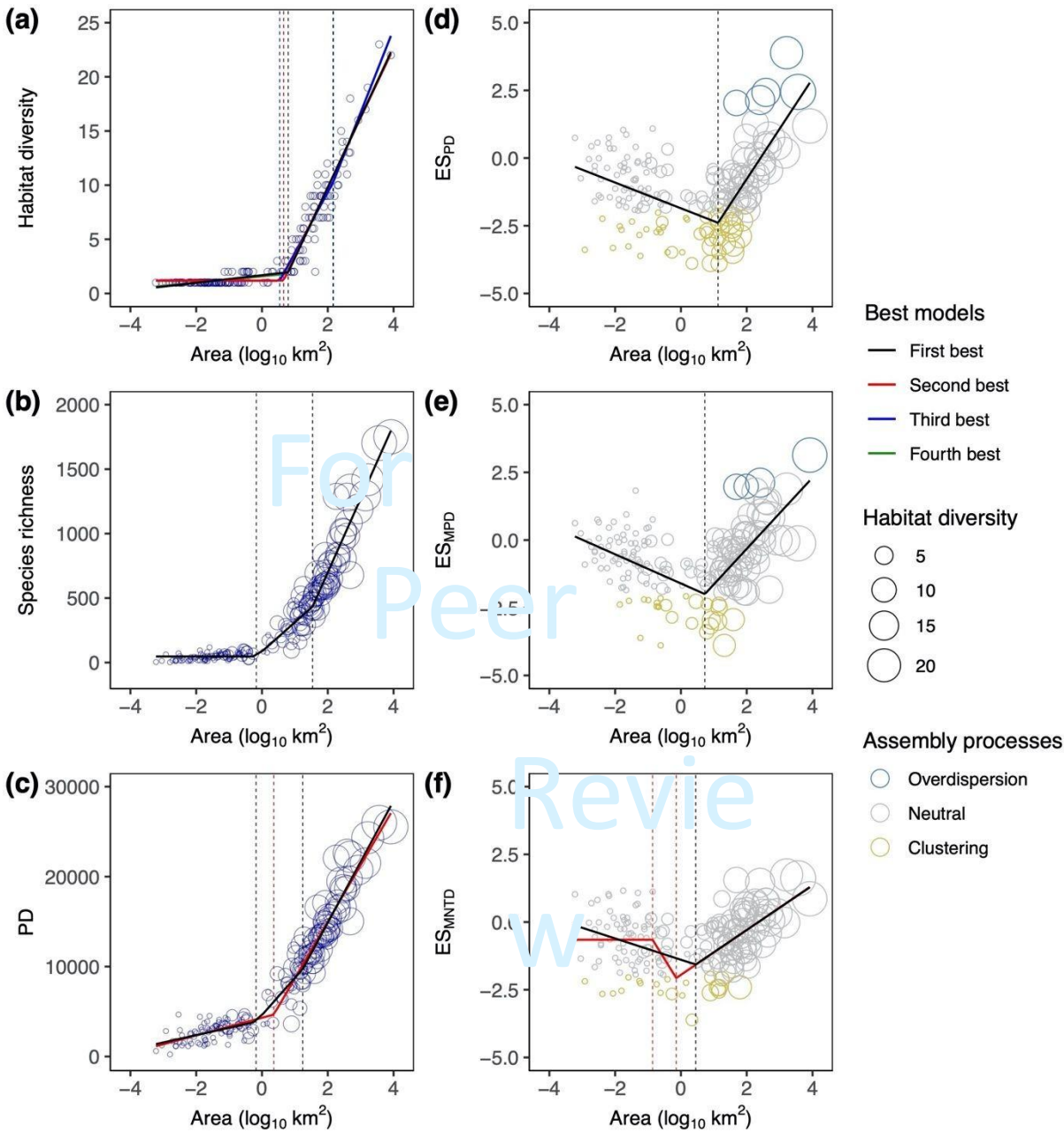


FIGURE 2 Relationships between \log_{10} area and habitat diversity (a), species richness (b), PD (c), ES_{PD} (d), ES_{MPD} (e) and ES_{MNTD} (f). Coloured circles indicate islands. Thick solid lines represent the predicted values of the best models and dashed vertical lines represent the threshold(s) of the best models. 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 ($\text{ES} > 1.96$) and clustering indicates that phylogenetic diversity values are significantly lower than expected

748 by chance ($SES < -1.96$). Neutral indicates that values are neither lower nor higher than
749 expected by chance ($-1.96 < SES < 1.96$).

750

For Review

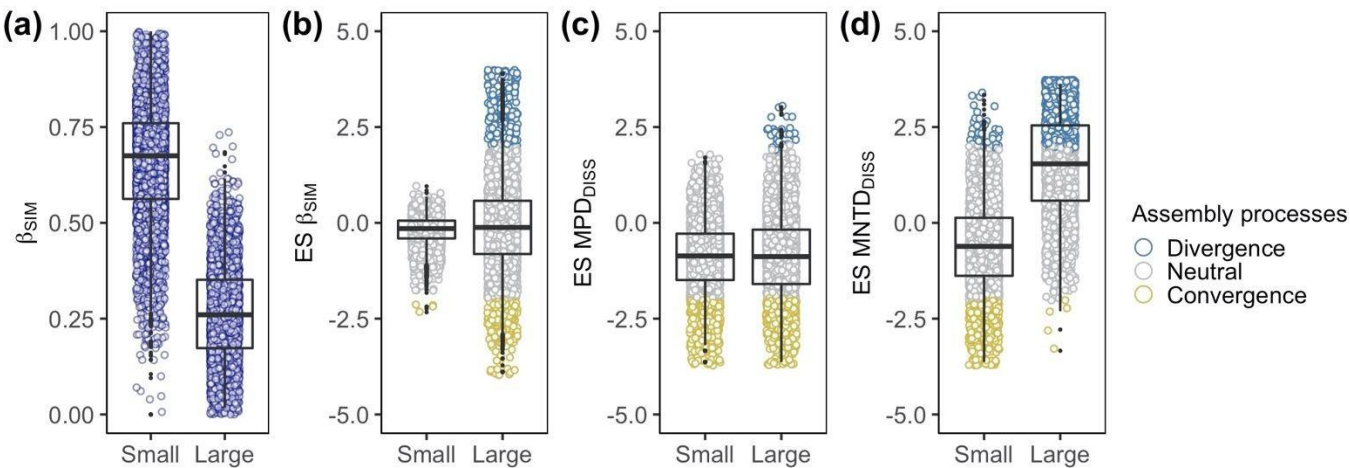


FIGURE 3 Differences in species turnover (a, b) and phylogenetic dissimilarity (c, d) between small and large islands. Islands were split into small and large using a breakpoint value of 2.57 km², calculated by averaging all breakpoints identified in the best models across all diversity metrics (see 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 $ES \beta_{SIM}$ (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 ($SES < -1.96$), and divergence indicates values that are significantly higher than expected by chance ($SES > 1.96$). Neutral indicates values that are neither lower nor higher than expected by chance ($-1.96 < SES < 1.96$).

