1	Article: Beta diversity patterns derived from island biogeography theory
2	Muyang Lu ^{*1} , David Vasseur ¹ , Walter Jetz ^{1, 2}
3	
4	1. Department of Ecology and Evolutionary Biology, Yale University, 165 Prospect
5	Street, New Haven, CT, 06520, USA.
6	2. Department of Life Sciences, Imperial College London, Silwood Park Campus,
7	Buckhurst Road, SL5 7PY Ascot, Berks, UK.
8	
9	Running head: beta diversity of islands
10	Words in the abstract: 164
11	Words in main body: 3820
12	References: 85
13	Table: 1
14	Figure: 3
15	* Correspondence: Muyang Lu (email: muyang.lu@yale.edu)

16 Abstract

17	The Theory of Island Biogeography (TIB) has been successful in predicting alpha
18	diversity patterns such as species-area relationships and species-abundance
19	distributions. Although beta diversity (i.e. the dissimilarity of community composition)
20	has long been recognized as an important element of the TIB and is crucial for
21	understanding community assembly processes, it has never been formally
22	incorporated into the theory. Here we derive theoretical predictions for the expected
23	pairwise beta diversity values under a species-level neutral scenario where all species
24	have equal colonization and extinction rates. We test these predictions for the avian
25	community composition of 42 islands (and 93 species) in the Thousand Island Lake,
26	China. We find that alpha diversity patterns alone do not distinguish a species-level
27	neutral model from a non-neutral model. In contrast, beta diversity patterns clearly
28	reject a species-level neutral model. We suggest that the presented theoretical
29	integration beta diversity offers a powerful path for testing the presence of neutral
30	processes in ecology and biogeography.
31	
32	Keywords: neutral theory; null model; beta diversity; meta-community; incidence

33 function; community assembly; regional species pool

34 INTRODUCTION

35 Island biogeography is undergoing a renaissance (Warren et al. 2015; Santos et al. 36 2016; Patiño et al. 2017), at the heart of which lies the equilibrium theory of island biogeography (TIB) formulated by MacArthur and Wilson 50 years ago (1967). The 37 38 TIB is most celebrated for elegantly linking species richness with colonization and 39 extinction rates, which in turn are influenced by island area and distance to the 40 mainland. Equilibrium TIB makes two key predictions: 1) equilibrium species richness depends on colonization and extinction rates; 2) community composition of 41 42 an island is undergoing constant turnover when at equilibrium. While early empirical 43 studies have focused on testing the equilibrium and temporal turnover predictions (Diamond 1969; Simberloff 1969; Simberloff and Wilson 1969, 1970; Diamond and 44 45 May 1977), recent interest in TIB is stimulated by the success of Hubbell's unified 46 neutral theory (a descendant of the TIB; Hubbell 2001), the advance of food web 47 theories (Holt 2009), the developments of phylogenetic analyses (Valente et al. 2015; 48 Lim and Marshall 2017), and the accumulation of island datasets (Weigelt et al. 2013, 49 2016). Motivated by the desire for a general island biogeography theory, more and 50 more complexities have been added to the original framework, including allometric scaling (Jacquet et al. 2017), trophic interactions (Gravel et al. 2011; Cazelles et al. 51 52 2016), speciation (Chen and He 2009; Rosindell and Phillimore 2011; Rosindell and 53 Harmon 2013), habitat heterogeneity (Kadmon and Allouche 2007), and island 54 ontogenies (Whittaker et al. 2008; Borregaard et al. 2016). While the complexity of

55	the theory has increased, examinations to date have focused on alpha diversity
56	patterns such as species richness (MacArthur and Wilson 1967), species-abundance
57	relationships (Rosindell and Phillimore 2011; Rosindell and Harmon 2013; Kessler
58	and Shnerb 2015), functional diversity (Jacquet et al. 2017), and endemism (Chen and
59	He 2009).
60	One particular omission in the theory and tests of the TIB to date is beta diversity,
61	i.e. the compositional dissimilarities among communities, a widely used metric in
62	biodiversity studies (Leprieur et al. 2011; Stegen et al. 2013; Segre et al. 2014; Si et al.
63	2016). The integration of beta diversity patterns into the TIB has the potential to
64	unlock a range of uses in community assembly studies and the development of a
65	unified meta-community theory (Leibold et al. 2004).
66	
67	The concept of beta diversity was first introduced by Whittaker (1960, 1965) and
68	
00	was quickly linked to an array of important mechanisms in community assembly:
69	was quickly linked to an array of important mechanisms in community assembly: biotic interaction (Graham and Fine 2008), environment filtering (Veech and Crist
69	biotic interaction (Graham and Fine 2008), environment filtering (Veech and Crist
69 70	biotic interaction (Graham and Fine 2008), environment filtering (Veech and Crist 2007; Buckley and Jetz 2008), dispersal limitation (Ojima and Jiang 2017; Wu et al.
69 70 71	 biotic interaction (Graham and Fine 2008), environment filtering (Veech and Crist 2007; Buckley and Jetz 2008), dispersal limitation (Ojima and Jiang 2017; Wu et al. 2017), and historical contingency (Fukami and Nakajima 2011). However, despite the
69 70 71 72	 biotic interaction (Graham and Fine 2008), environment filtering (Veech and Crist 2007; Buckley and Jetz 2008), dispersal limitation (Ojima and Jiang 2017; Wu et al. 2017), and historical contingency (Fukami and Nakajima 2011). However, despite the widespread recognition of the beta diversity concept, there is little consensus about

76	White 1999; Legendre and De Cáceres 2013); based on pairwise or multiple-sites
77	comparisons (Hui and McGeoch 2014; Arita 2017; Latombe et al. 2017; Marion et al.
78	2017); decomposed in additive or multiplicative framework (with respect to alpha and
79	gamma diversity; Jost 2007; Cabral et al. 2014); calculated by incidence
80	(presence-absence) or abundance data (Chao et al. 2016). Even for the commonly
81	used pairwise beta diversity, there is a heated debate about what is the best way to
82	partition it into richness difference and replacement (or nestedness and turnover)
83	components (Baselga 2010; Podani and Schmera 2011; Legendre 2014; Baselga and
84	Leprieur 2015). It is generally agreed that different measures quantify different
85	aspects of communities and that their use should be guided by the question asked
86	(Anderson et al. 2011; Legendre 2014). Although beta-diversity measures have been
87	increasingly synthesized under more general frameworks such as Hill's number and
88	variance of community matrix (Chao et al. 2016), only a few of them have been
89	examined analytically from process-based theories (Chave and Leigh 2002; Condit et
90	al. 2002; Zillio et al. 2005; Connolly et al. 2017). Moreover, theoretical
91	considerations of beta-diversity patterns to date have been abundance-based and
92	hence limited to more restricted datasets, such as vegetation plots (Chave and Leigh
93	2002; Condit et al. 2002; Zillio et al. 2005). In summary, the link between beta
94	diversity patterns and fundamental processes such as colonization and extinction in
95	the TIB remain ill developed.
96	

97	In this study, we set out to address this shortcoming and provide a theoretical and
98	example empirical integration of beta diversity into TIB. Specifically, we present the
99	first theoretical predictions for the expected pairwise beta diversity patterns for both
100	Jaccard and Sørensen indices in the classic MacArthur and Wilson framework. This
101	framework constitutes in essence a species-level neutral model (as opposed to the
102	individual-level neutral theory; Hubbell 2009; Rosindell and Phillimore 2011). The
103	species-level neutral model assumes that all species have the same colonization and
104	extinction rates, while the individual-level neutral model assumes that all individuals
105	have the same colonization and extinction rates, which allows species-level
106	asymmetry to arise because of different abundance (Rosindell and Harmon 2013). We
107	first derived the probability mass function and expected value of pairwise beta
108	diversity in a general case where the equilibrium assumption is not required. We then
109	examined our results under the classic MacArthur-Wilson equilibrium framework.
110	Finally, we test the derived predictions for a dataset of 93 avian species occurring
111	across 42 islands of the Thousand Island Lake Region in China (Appendix, Wang et
112	<i>al.</i> 2010).

114 MATERIAL AND METHODS

To derive expected beta diversity, we need to consider the joint statistical distribution of multiple islands. We here focus on pairwise community comparisons, because average pairwise dissimilarity is shown to be the only unbiased and consistent

118	estimator when applied to empirical data (Marion et al. 2017). We also calculated the
119	expected value of the partitioned components of pairwise beta dissimilarity for both
120	Baselga and Podani families (Baselga 2010; Podani and Schmera 2011). To test our
121	predictions, we applied an "Incidence Function" approach pioneered by Diamond
122	(1975) and used by Connor and Simberloff (1978, 1979) during early debates of null
123	models.

125 **Probability mass function of pairwise beta diversity**

All pairwise beta diversity indices share the same probability mass function to describe island community composition because they are calculated from four quantities: the total number of species of two islands *N* (regional species richness), the number of species shared by both islands *k*, the number of species unique to the first island *i*, and the number of species unique to the second island *j* (thus N = k + i + j). Then the number of species present on the first island is *i*+*k*, and the number of species present on the second island is *j*+*k*.

Denoting the size of the mainland species pool as M, the probability mass function of pairwise beta diversity, conditioning on N, k and i, can be derived as follows: let the occurrence probability of the i^{th} species on the first island be p, and the j^{th} species on the second island be q. Furthermore, assuming that the occurrence probabilities are properties of the island rather than of species, the probability of obtaining exactly iunique species and k shared species from a regional pool of size M is given by 139 $p^{k+i}(1-p)^{M-(k+i)}$. Similarly, the probability of obtaining *j* unique and *k* shared species on 140 the second island is $q^{k+j}(1-q)^{M-(k+j)}$, which yields the probability of obtaining the 141 regional community specified by *i*, *j*, *k* and *M*, as $p^{k+i}(1-p)^{M-(k+i)}q^{k+j}(1-q)^{M-(k+j)}$.

142 Assuming that species are neutral with respect to their occurrence probabilities 143 within an island, we then count how many different combinations of island 144 communities could be obtained with regional richness N, k shared species, and i 145 unique species to the first island (the number of unique species to the second island is then fixed as k-i). There are exactly $\binom{M}{N}$ ways of choosing N out of M species to be 146 147 present in either or both islands from the mainland species pool. Similarly, there are $\binom{N}{k}$ ways to choose k out of N species to be shared by both islands and for the unique 148 *N-k* species, there are $\binom{N-k}{i}$ ways of assigning *i* species to the first island. After 149 150 simplification, the probability of pairwise beta diversity conditioning on the total 151 number of species of both islands N, the number of species shared by both islands k, 152 and the number of species unique to the first island *i* is:

153
$$\Pr(regional.community \mid N, k, i) = \begin{pmatrix} M \\ N \end{pmatrix} \begin{pmatrix} N \\ k, i \end{pmatrix} p^{k+i} q^{k+j} (1-p)^{M-(k+i)} (1-q)^{M-(k+j)}$$

154 (1)

155

Because beta diversity measures (e.g. Jaccard and Sørensen indices) are usually undefined when N = 0, the expected pairwise beta diversity should be normalized by the term 1- $(1-p)^{M}(1-q)^{M}$ to exclude "double-absence" scenarios (Anderson et al. 2011). Only the expected Jaccard dissimilarity (Jaccard dissimilarity is defined as

160
$$1 - \frac{\kappa}{N}$$
) has a simple analytical form that is independent of the size of mainland
161 species pool *M* (see Appendix S1 in Supporting Information for derivation; we do not
162 show the analytical form of the expected Sørensen dissimilarity because the specific
163 form depends on the size of mainland species pool):

164
$$E(Jaccard.dissimilarity) = \frac{p+q-2pq}{p+q-pq}$$
(2)

166 It has a form that is similar to the expected pairwise Jaccard derived by Chase et al. 167 (2011) but allows two islands to have different occurrence probabilities. p and q could 168 also be interpreted as probabilities of the same island at different times, and in that 169 case (2) becomes the expected temporal Jaccard dissimilarity. This quantity does not 170 require any equilibrium assumptions of alpha diversity and hence could be seen as a 171 generalized version of the apparent turnover derived by Diamond and May (1977) 172 which is essentially a pairwise Sørensen dissimilarity index in equilibrium (Morrison 173 2017). We also calculated the expected pairwise beta diversity for Jaccard and Sørensen 174 175 families as well as their partitioned components (for detailed formulas of the Baselga 176 and Podani families see Baselga & Leprieur 2015) conditioning on both islands 177 having species (i+k > 0 and i+k > 0) because empty islands are sometimes excluded 178 from analysis (Wang et al. 2016). In Baselga's family, pairwise dissimilarity is 179 partitioned into the nestedness component and the turnover component. The 180 nestedness component measures the extent to which species in a smaller community

181	are a subset of species in a larger community, while the turnover component measures
182	how much of the dissimilarity is caused by species replacement (Baselga 2010). The
183	turnover component and the nestedness component in Baselga's family, respectively,
184	correspond to the replacement component and the richness difference component in
185	Podani's family (Legendre 2014).
186	
187	Island biogeography theory
188	Following the stochastic version of the TIB (MacArthur and Wilson 1967), the
189	occurrence probability of a species on an island is a function of colonization rate c
190	and extinction rate <i>e</i> :
191	$\frac{dp}{dt} = c(1-p) - ep \tag{3}$
192	When occurrence probability of a species is at stochastic equilibrium (stationary
193	distribution), the occurrence probability is:
194	$p_{i,eq} = \frac{c}{c+e} \tag{4}$
195	Let $\theta = \frac{e}{c}$, which is the relative extinction rates, equation (4) becomes:
196	$p_{i,eq} = \frac{1}{1+\theta} \tag{5}$
197	Substitute $p = q = \frac{1}{1+\theta}$ into equation (2), the expected Jaccard dissimilarity when
198	two islands have equal colonization and extinction probabilities is:
199	$E(Jaccard.dissimilarity) = \frac{2\theta}{1+2\theta} $ (6)
200	To take into account asymmetry between islands (e.g. the effect of area, isolation or
201	habitat types), let $p = \frac{1}{1+\theta}$ and $q = \frac{1}{1+z\theta}$, so that the relative extinction rate on the

second island is *z* times the relative extinction rate on the first island. Equation (2)

204
$$E(Jaccard.dissimilarity) = \frac{(1+z)\theta}{1+(1+z)\theta}$$
 (7)

205

206 Empirical tests

207 Data

208 We use a published dataset of avian community composition for 93 birds and 42 209 islands in The Thousand Island Lake, China to test our predictions (see Appendix in 210 Wang et al. 2010). The region (29°22" – 29°50" N, 118°34"– 119°15" E) consists of 211 an inundated lake with more than 1000 islands created by dam constructions in 1959. 212 Because the islands were formed recently, there was no in situ speciation in this region. The relative small area of the region (573 km^2) ensures that the islands are 213 accessible to most of the bird species (Si et al. 2016). Bird occupancies from 2006 to 214 215 2009 on 42 islands were documented using line-transect method. Island variables measured in the dataset include area, distance to mainland, habitat diversity and plant 216 217 species richness.

218

219 Incidence function and statistical analysis

220 Because extinction rates and colonization rates are difficult to measure directly, an

- alternative way is to use an "Incidence function" approach to estimate parameters
- from a snapshot of occupancy patterns (Diamond 1975; Hanski 2009). To test the

223 predictions of island biogeography theory, equation (4) is modeled as a function of

- isolation and area. Parameters are fitted by maximum likelihood with binomial
- distribution. Three neutral models and one non-neutral model are examined in this

study:

227 Neutral model 1: Inverse ratio

228 Colonization rates are modeled as an inverse ratio function of isolation: $c = \frac{a}{I}$, and

229 extinction rates are modeled as an inverse ratio of area: $e = \frac{b}{A}$. *a* and *b* are fitted

230 parameters.

- 231 Neutral model 2: Exponential
- 232 Colonization rates are modeled as an exponential function of isolation:
- 233 $c = \exp(-aI)$, and extinction rates are modeled as an exponential function of area:
- 234 $e = \exp(-bA)$. a and b are fitted parameters. Substitute $c = \exp(-aI)$ and
- 235 $e = \exp(-bA)$ into equation (4), the occurrence probability becomes

236
$$p = \frac{1}{1 + \exp(-bA + aI)}$$
. Thus the exponential neutral model is equivalent to a logistic

- regression without intercepts, which could be directly fitted by 'glm' in R with
- binomial distribution and logit link.
- 239 Neutral model 3: GLM
- 240 This model adds habitat diversity and plant richness to the predictors in addition to
- area and isolation in Neutral Model 2. AIC is used to select the best model. Area and
- isolation are log-transformed to ensure better linearity.
- 243 Non-neutral model: Aggregate species-level GLM

244	To take into account species-level non-neutrality, species identity is included as a
245	fixed effect categorical variable in the GLM, which allows each species to have a
246	different intercept (baseline occurrence probability) but share the same response to
247	island area, isolation, plant richness and habitat types. This is essentially a stacked
248	species distribution modeling approach (Calabrese et al. 2014; Ko et al. 2016), which
249	increases the number of parameters by the number of species minus one (92
250	parameters in this case). This procedure aims to capture the observed
251	presence-absence variation among species but does not pin down the causes of
252	species-level non-neutrality such as traits and abundance differences. More realistic
253	assumptions of non-neutrality such as different responses to area and isolation could
254	be made but are not the main focus of this paper. AUC and AIC values are calculated
255	to compare the overall performance of the models. AUC is calculated with a
256	Mann-Whitney U statistic using R package 'PresenceAbsence'. AIC is calculated
257	from the best-fitted likelihood function. ΔAIC are calculated by subtracting the AIC
258	of the null model (GLM with only one intercept) from the AIC of the fitted model.
259	
260	Predictions of alpha diversity and beta diversity
0(1	

The predicted species richness of each island is given by the summed fitted

- 262 occurrence probabilities of all species as used in conventional stacked species
- distribution modeling (Calabrese et al. 2014; Ko et al. 2016). While predicted
- 264 pairwise Jaccard dissimilarity can be calculated analytically from equation (2), the

265	partitioned nestedness and turnover components can only be estimated by simulations
266	(we only present the results of Baselga's family because its partitioned components
267	are independent with each other when gamma diversity is fixed; Baselga & Leprieur
268	2015). We therefore estimate predicted pairwise beta diversity as the mean of 1000
269	community matrices simulated from the fitted occurrence probabilities. Observed
270	values are regressed against predicted values using ordinary least square. R^2 is used as
271	a measure of predictive power for alpha and beta diversity patterns. If the model
272	predicts the observed patterns well, the fitted regression line should be close to the 1:1
273	ratio line when observed values are plotted against predicted values. All statistical
274	analyses are performed in R 3.3.2.

276 **RESULTS**

277 Symmetric islands (same colonization and extinction rates for both islands)

278 The expected Jaccard dissimilarity conditional on both islands having non-zero 279 species richness increases monotonically with relative extinction rate θ and the size (number of species) of the mainland species pool M. As M increases, the expected 280 Jaccard dissimilarity converges to $\frac{2\theta}{1+2\theta}$ (Fig. 1a). When decomposed into turnover 281 282 and nestedness (or replacement and richness difference) components, contrasting 283 patterns are observed: while the turnover component and the replacement component 284 increase monotonically with θ and M (Fig. 1b, e), the nestedness component and richness difference component are both unimodal functions of θ with maximum 285

286 values less than 0.3 (Fig. 1c, f). The maximum nestedness decreases as M increases 287 (Fig. 1c), while the maximum richness difference changes little with the increase of M288 (Fig. 1f). The relative importance of turnover increases monotonically with θ and M 289 (Fig. 1d). In a special case of 2 mainland species, the ratio of expected turnover and 290 expected nestedness equals θ (Fig. 1d). The Sørensen family indices have similar 291 quantitative behaviors as the Jaccard family indices (see Fig. S1 in Appendix). 292 293 Asymmetric islands (different colonization and extinction rates for two islands) 294 The expected Jaccard dissimilarity conditional on both islands having non-zero species richness converges to $\frac{(1+z)\theta}{1+(1+z)\theta}$ as *M* increases. But the deviation from 295 $\frac{(1+z)\theta}{1+(1+z)\theta}$ also increases with islands asymmetry z (Fig. 2a, e, i). The turnover 296 297 component increases monotonically with θ and M, and decreases with z (Fig. 2b, f, j). 298 The nestedness component decreases with M when z is small (Fig. 2c), but this 299 relationship with M is reversed when z gets larger (Fig. 2g, k). The ratio of expected turnover and expected nestedness also increases with M when z is small (Fig. 2d), and 300 301 decreases with M when z is large (Fig. 2h, 1). 302 303 **Empirical test**

We tested the above predictions for an inland lake island system using 93 bird species

and 42 islands. We used an incidence function approach to fit the observed occupancy

306	patterns and calculated predicted alpha and beta diversity patterns from the fitted
307	models. The non-neutral GLM has the lowest AIC and highest AUC in all models
308	(Table 1). Among three neutral models, the neutral GLM model has the lowest AIC
309	and the highest AUC. All models are better at predicting alpha diversity patterns than
310	at predicting beta diversity patterns (Fig. 3). The neutral exponential model and GLM
311	systematically underestimate the nestedness component, but overestimate the turnover
312	component (Fig. 3f, j). In contrast, the neutral inverse ratio model systematically
313	underestimates all observed pairwise beta diversity at the lower range of the
314	predictions, yet it overestimates pairwise beta diversity at the higher range of the
315	predictions (Fig. 3a-d). Both the neutral GLM and the non-neutral GLM successfully
316	predict the observed alpha diversity pattern (Fig. 3i, m), but only the non-neutral
317	GLM successfully predicts the observed beta diversity patterns (Fig. 3n-p).

319

320 **DISCUSSION**

321 Theoretical results

We derived a set of novel predictions of beta diversity patterns from the island biogeography theory. Under the classic MacArthur and Wilson framework, pairwise beta diversity patterns are determined by three factors: the size of the mainland species pool, extinction rates and colonization rates. Conditional on at least one island having species (N>0), only the expected Jaccard dissimilarity is independent of the size of mainland species pool. When conditioning on both islands having non-zero

328	richness ($i+k > 0$ and $j+k > 0$), all indices are dependent on the size of mainland
329	species pool. Ignoring empty islands has a huge influence on beta diversity patterns
330	when the size of mainland species pool is small (e.g. less than 10 species), but the
331	effect becomes negligible toward larger mainland species pool sizes, because of the
332	rapidly decreasing probability of empty islands (Fig. 1a). This result stresses the need
333	to include empty islands in empirical tests of the island biogeography theory (Wang et
334	al. 2016).

336 Pairwise Jaccard dissimilarity increases with relative extinction rates (the ratio between extinction rates and colonization rates). This is because when extinction 337 events become more frequent the chance of forming "checkerboard" patterns 338 339 (Diamond 1975) grows, as shown by the increasing turnover component (Fig. 1b). 340 The nestedness component first increases then decreases with relative extinction rates 341 (Fig. 1c). The maximum nestedness is achieved when extinction rates are less than 342 colonization rates (Fig. 1c). When relative extinction rates or the size of mainland 343 species pool gets larger, the increase of Jaccard dissimilarity is mainly driven by the 344 increasing turnover component (Fig. 1d). Indices in Podani's family (the replacement 345 and richness difference components) have similar qualitative behaviors, consistent 346 with previous finding that the partitioned components in Baslega's family are 347 correlated with the partitioned components in Podani's family (Legendre 2014; 348 Baselga and Leprieur 2015). Island asymmetry, which takes into account the effect of 349 isolation and area, does not change the qualitative behaviors of the indices (Fig 2), but

- increases the level of nestedness, because species in a small (distant) island are more
- 351 likely to be a subset of species in a large (near) island.
- 352

353 Neutral theory and null models

354 Our theoretical results are derived from a species-level neutral theory. We show that 355 both the neutral GLM and the non-neutral GLM are successful in predicting species 356 richness (alpha diversity) of birds in the Thousand Islands Lake (Fig. 3i, m). But only 357 the non-neutral model predicts the observed beta diversity patterns. The patterns of 358 partitioned components further reveal that the neutral GLM fails to predict the 359 observed pairwise Jaccard dissimilarity because it overestimates turnover and 360 underestimates nestedness (Fig. 3j-l). Our analysis does not include biological factors 361 that may cause species-level non-neutrality, but the presented framework is flexible to 362 include such information if available.

Null models have become one of the most important tools in ecology. Despite their increasing popularity, it has also been recognized that their assumptions (e.g. the widely used random shuffling approaches) can have serious limitations, especially in the presence of multiple processes (Gotelli and Ulrich 2012; Pigot and Etienne 2015; Miller et al. 2017; O'Dwyer et al. 2017). In terms of beta diversity, the entangled links among alpha, beta and gamma diversity are known to reduce the statistical power of randomization tests and generate ambiguity in their interpretations (Chase et

370	al. 2011; Kraft et al. 2011; Qian et al. 2012, 2013; Ulrich et al. 2017). Our results
371	support the use of mechanistic null models such as those based on neutral theory
372	(O'Dwyer et al. 2009; Rosindell et al. 2012), maximum entropy theory (Xiao et al.
373	2015, 2016; O'Dwyer et al. 2017), and incidence functions (Hanski et al. 1996; Helm
374	et al. 2006; Hanski 2009) to improve upon random-shuffling null models in
375	hypothesis testing. These mechanistic models could generate multiple diagnostic
376	patterns and allows stronger test for ecological theories.

378 Future directions

379 Our results are derived from a mainland-island model where the colonization and

380 extinction events on islands are independent of one another. More realistic

381 modifications should be considered in order to predict biodiversity changes in a real

382 landscape (Hanski 2009). For example, environmental heterogeneity and local

383 dispersal are found to be important drivers of beta diversity in experimental studies

- 384 (Grainger and Gilbert 2016; Gianuca et al. 2017; Ojima and Jiang 2017; Rodrigues
- and Diniz-Filho 2017). Moreover, in oceanic islands, high endemism arises from

386 speciation events (Chen and He 2009; Gascuel et al. 2016; Steinbauer et al. 2016),

387 which are likely to increase the level of species turnover. Future studies should also

- 388 extend presence-based measures to abundance-based beta diversity measures which
- may provide additional insights into the dynamics of meta-community (Bashan et al.
- 390 2016; Kalyuzhny and Shnerb 2017). We believe that the integration of beta diversity

- 391 patterns into the Theory of Island Biogeography offers new opportunities to infer
- 392 community assembly processes.
- 393

394 ACKNOWLEDGEMENTS

- 395 We thank Dr. James Rosindell for providing helpful comments for earlier drafts. We
- thank members in Jetz Lab and Vasseur Lab for feedbacks and dicussions.

397 AUTHORSHIP

- 398 ML and DV conceived of the idea. The idea was further developed with DV and WJ's
- input. ML conducted the analyses and wrote the manuscript. All authors assisted with
- 400 revisions.

401 **REFERENCE**

- 402 Anderson, M. J., T. O. Crist, J. M. Chase, M. Vellend, B. D. Inouye, A. L. Freestone,
- 403 N. J. Sanders, et al. 2011. Navigating the multiple meanings of β diversity: a
- 404 roadmap for the practicing ecologist. Ecology Letters 14:19–28.
- 405 Arita, H. T. 2017. Multisite and multispecies measures of overlap, co-occurrence, and
- 406 co-diversity. Ecography 40:709–718.
- 407 Baselga, A. 2010. Partitioning the turnover and nestedness components of beta
- 408 diversity. Global Ecology and Biogeography 19:134–143.
- 409 Baselga, A., and F. Leprieur. 2015. Comparing methods to separate components of
- 410 beta diversity. Methods in Ecology and Evolution 6:1069–1079.
- 411 Bashan, A., T. E. Gibson, J. Friedman, V. J. Carey, S. T. Weiss, E. L. Hohmann, and

412	YY. Liu. 2016.	Universality	v of human	microbial	dvnamics.	Nature 534:259-

- 413 262.
- 414 Borregaard, M. K., T. J. Matthews, and R. J. Whittaker. 2016. The general dynamic
- 415 model: towards a unified theory of island biogeography? Global Ecology and
- 416 Biogeography 25:805–816.
- 417 Buckley, L. B., and W. Jetz. 2008. Linking global turnover of species and
- 418 environments. Proceedings of the National Academy of Sciences 105:17836–
- 419 17841.
- 420 Cabral, J. S., P. Weigelt, W. D. Kissling, and H. Kreft. 2014. Biogeographic, climatic
- 421 and spatial drivers differentially affect α -, β and γ -diversities on oceanic

422 archipelagos. Proceedings. Biological sciences / The Royal Society

- 423 281:20133246.
- 424 Calabrese, J. M., G. Certain, C. Kraan, and C. F. Dormann. 2014. Stacking species
- 425 distribution models and adjusting bias by linking them to macroecological
- 426 models. Global Ecology and Biogeography 23:99–112.
- 427 Cazelles, K., N. Mouquet, D. Mouillot, and D. Gravel. 2016. On the integration of
- 428 biotic interaction and environmental constraints at the biogeographical scale.
- 429 Ecography 39:921–931.
- 430 Chao, A., C. H. Chiu, and D. Warton. 2016. Bridging the variance and diversity
- 431 decomposition approaches to beta diversity via similarity and differentiation
- 432 measures. Methods in Ecology and Evolution 7:919–928.

- 433 Chase, J. M., N. J. B. Kraft, K. G. Smith, M. Vellend, and B. D. Inouye. 2011. Using
- null models to disentangle variation in community dissimilarity from variation in
- 435 α -diversity. Ecosphere 2:art24.
- 436 Chave, J., and E. G. Leigh. 2002. A spatially explicit neutral model of β-diversity in
- 437 tropical forests. Theoretical Population Biology 62:153–168.
- Chen, X.-Y., and F. He. 2009. Speciation and endemism under the model of island
 biogeography. Ecology 90:39–45.
- 440 Condit, R., N. Pitman, E. G. Leigh Jr., J. Chave, J. Terborgh, R. B. Foster, P. Nunez
- 441 V., et al. 2002. Beta-diversity in tropical forest trees. Science 295:666–669.
- 442 Connolly, S. R., T. P. Hughes, and D. R. Bellwood. 2017. A unified model explains

443 commonness and rarity on coral reefs. Ecology Letters 20:477–486.

- 444 Connor, E. F., and D. Simberloff. 1978. Species number and compositional similarity
- 445 of the Galapagos Flora and Avifauna. Ecological Monographs 48:219–248.
- 446 ——. 1979. The assembly of species communities: chance or competition?
- 447 Ecology 60:1132.
- 448 Diamond, J. M. 1969. Avifaunal equilibria and species turnover rates on the Channel
- 449 Islands of California. Proceedings of the National Academy of Sciences 64:57–
- 450 63.
- 453 Diamond, J. M., and R. M. May. 1977. Species turnover rates on islands: dependence

454 on census interval. Science 197:266–270.

455	Fukami, T., and M. Nakajima. 2011. Community assembly: alternative stable states or
456	alternative transient states? Ecology Letters 14:973-984.
457	Gascuel, F., F. Laroche, A. S. Bonnet-Lebrun, and A. S. L. Rodrigues. 2016. The
458	effects of archipelago spatial structure on island diversity and endemism:
459	predictions from a spatially-structured neutral model. Evolution 70:2657–2666.
460	Gianuca, A. T., S. A. J. Declerck, P. Lemmens, and L. De Meester. 2017. Effects of
461	dispersal and environmental heterogeneity on the replacement and nestedness
462	components of β-diversity. Ecology 98:525–533.
463	Gotelli, N. J., and W. Ulrich. 2012. Statistical challenges in null model analysis.
464	Oikos 121:171–180.
465	Graham, C. H., and P. V. A. Fine. 2008. Phylogenetic beta diversity: linking
466	ecological and evolutionary processes across space in time. Ecology Letters
467	11:1265–1277.
468	Grainger, T. N., and B. Gilbert. 2016. Dispersal and diversity in experimental
469	metacommunities: linking theory and practice. Oikos 125:1213-1223.
470	Gravel, D., F. Massol, E. Canard, D. Mouillot, and N. Mouquet. 2011. Trophic theory
471	of island biogeography. Ecology Letters 14:1010-1016.
472	Hanski, I. 2009. The theories of island biogeography and metapopulation dynamics.

- 473 Pages 186–213 *in* J. B. Losos and R. E. Ricklefs, eds. The theory of island
- 474 biogeography revisited. Princeton University Press, Princeton.

- 475 Hanski, I., A. Moilanen, T. Pakkala, and M. Kuussaari. 1996. The quantitative
- 476 incidence function model and persistence of an endangered butterfly
- 477 metapopulation. Conservation Biology 10:578–590.
- 478 Helm, A., I. Hanski, and M. Pärtel. 2006. Slow response of plant species richness to
- habitat loss and fragmentation. Ecology Letters 9:72–77.
- 480 Holt, R. D. 2009. Toward a trophic island biogeography. Pages 143–185 in J. B.
- 481 Losos and R. E. Ricklefs, eds. The theory of island biogeography revisited.
- 482 Princeton University Press, Princeton.
- 483 Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography.
- 484 Princeton University Press, Princeton.
- 485 ——. 2009. Neutral theory and the theory of island biogeography. Pages 263–292
- 486 *in* J. B. Losos and R. E. Ricklefs, eds. The theory of island biogeography
- 487 revisited. Princeton University Press, Princeton.
- 488 Hui, C., and M. A. McGeoch. 2014. Zeta diversity as a concept and metric that unifies
- 489 Incidence-based biodiversity patterns. The American Naturalist 184:684–694.
- 490 Jacquet, C., D. Mouillot, M. Kulbicki, and D. Gravel. 2017. Extensions of Island
- 491 Biogeography Theory predict the scaling of functional trait composition with
- habitat area and isolation. Ecology Letters 20:135–146.
- 493 Jost, L. 2007. Partitioning diversity into independent alpha and beta components.
- 494 Ecology 88:2427–2439.
- 495 Kadmon, R., and O. Allouche. 2007. Integrating the effects of area, isolation, and

496	habitat hetero	geneity on s	pecies o	liversity: a	unification	of island	biogeography
-----	----------------	--------------	----------	--------------	-------------	-----------	--------------

- and niche theory. The American Naturalist 170:443–454.
- 498 Kalyuzhny, M., and N. M. Shnerb. 2017. Dissimilarity-overlap analysis of community
- 499 dynamics: opportunities and pitfalls. Methods in Ecology and Evolution 0:1–10.
- 500 Kessler, D. A., and N. M. Shnerb. 2015. Generalized model of island biodiversity.
- 501 Physical Review E Statistical, Nonlinear, and Soft Matter Physics 91:1–11.
- 502 Ko, C. Y., O. J. Schmitz, and W. Jetz. 2016. The limits of direct community modeling
- 503 approaches for broad-scale predictions of ecological assemblage structure.
- 504 Biological Conservation 201:396–404.
- 505 Kraft, N. J. B., L. S. Comita, J. M. Chase, N. J. Sanders, N. G. Swenson, T. O. Crist, J.
- 506 C. Stegen, et al. 2011. Disentangling the drivers of diversity along latitudinal and
 507 elevational gradients. Science 333:1755–1758.
- 508 Latombe, G., C. Hui, and M. A. McGeoch. 2017. Multi-site generalised dissimilarity
- 509 modelling: using zeta diversity to differentiate drivers of turnover in rare and
- 510 widespread species. Methods in Ecology and Evolution 8:431–442.
- 511 Legendre, P. 2014. Interpreting the replacement and richness difference components

of beta diversity. Global Ecology and Biogeography 23:1324–1334.

- 513 Legendre, P., and M. De Cáceres. 2013. Beta diversity as the variance of community
- 514 data: dissimilarity coefficients and partitioning. Ecology Letters 16:951–963.
- 515 Leibold, M. A., M. Holyoak, N. Mouquet, P. Amarasekare, J. M. Chase, M. F.
- 516 Hoopes, R. D. Holt, et al. 2004. The metacommunity concept: a framework for

517 multi-scale community ecology. Ecology Letters 7:601–613.	517	multi-scale commu	nity ecology.	Ecology L	etters 7:601–613.
---	-----	-------------------	---------------	-----------	-------------------

- 518 Leprieur, F., P. A. Tedesco, B. Hugueny, O. Beauchard, H. H. Dürr, S. Brosse, and T.
- 519 Oberdorff. 2011. Partitioning global patterns of freshwater fish beta diversity
- 520 reveals contrasting signatures of past climate changes. Ecology Letters 14:325–
- 521 334.
- 522 Lim, J. Y., and C. R. Marshall. 2017. The true tempo of evolutionary radiation and
- 523 decline revealed on the Hawaiian archipelago. Nature 543:710–713.
- 524 MacArthur, R. H., and E. O. Wilson. 1967. The theory of island biogeography.
- 525 Princeton University Press, Princeton.
- 526 Marion, Z. H., J. A. Fordyce, and B. M. Fitzpatrick. 2017. Pairwise beta diversity
- 527 resolves an underappreciated source of confusion in calculating species turnover.
- 528 Ecology 98:933–939.
- 529 Miller, E. T., D. R. Farine, and C. H. Trisos. 2017. Phylogenetic community structure
- 530 metrics and null models: a review with new methods and software. Ecography
- 531 40:461–477.
- 532 Morrison, L. W. 2017. Insular plant turnover across a 22-year interval: a critical
- retrospective of the roles of pseudoturnover and cryptoturnover. Journal of
- 534 Biogeography 44:1007–1017.
- Nekola, J. C., and P. S. White. 1999. The distance decay of similarity in biogeography
 and ecology. Journal of Biogeography 26:867–878.
- 537 O'Dwyer, J. P., J. K. Lake, A. Ostling, V. M. Savage, and J. L. Green. 2009. An

- 538 integrative framework for stochastic, size-structured community assembly.
- 539 Proceedings of the National Academy of Sciences 106:6170–6175.
- 540 O'Dwyer, J. P., A. Rominger, and X. Xiao. 2017. Reinterpreting maximum entropy in
- 541 ecology: a null hypothesis constrained by ecological mechanism. Ecology
- 542 Letters 20:832–841.
- 543 Ojima, M. N., and L. Jiang. 2017. Interactive effects of disturbance and dispersal on
 544 community assembly. Oikos 126:682–691.
- 545 Patiño, J., R. J. Whittaker, P. A. V. Borges, J. M. Fernández-Palacios, C. Ah-Peng, M.
- 546 B. Araújo, S. P. Ávila, et al. 2017. A roadmap for island biology: 50 fundamental
- questions after 50 years of The Theory of Island Biogeography. Journal of
- 548 Biogeography 44:963–983.
- 549 Pigot, A. L., and R. S. Etienne. 2015. A new dynamic null model for phylogenetic
- community structure. Ecology Letters 18:153–163.
- 551 Podani, J., and D. Schmera. 2011. A new conceptual and methodological framework
- for exploring and explaining pattern in presence absence data. Oikos 120:1625–
 1638.
- 554 Qian, H., S. Chen, L. Mao, and Z. Ouyang. 2013. Drivers of β-diversity along
- latitudinal gradients revisited. Global Ecology and Biogeography 22:659–670.
- 556 Qian, H., X. Wang, and Y. Zhang. 2012. Comment on "Disentangling the drivers of
- diversity along latitudinal and elevational gradients." Science 335:1573–1573.
- 558 Rodrigues, J. F. M., and J. A. F. Diniz-Filho. 2017. Dispersal is more important than

559	climate in	structuring	turtle	communities acros	s different	t biogeogra	phical	realms.

- 560 Journal of Biogeography 44:2109–2120.
- 561 Rosindell, J., and L. J. Harmon. 2013. A unified model of species immigration,
- extinction and abundance on islands. Journal of Biogeography 40:1107–1118.
- 563 Rosindell, J., S. P. Hubbell, F. He, L. J. Harmon, and R. S. Etienne. 2012. The case
- for ecological neutral theory. Trends in Ecology and Evolution 27:203–208.
- 565 Rosindell, J., and A. B. Phillimore. 2011. A unified model of island biogeography
- sheds light on the zone of radiation. Ecology Letters 14:552–560.
- 567 Santos, A. M. C., R. Field, and R. E. Ricklefs. 2016. New directions in island
- 568 biogeography. Global Ecology and Biogeography 25:751–768.
- 569 Segre, H., R. Ron, N. De Malach, Z. Henkin, M. Mandel, and R. Kadmon. 2014.
- 570 Competitive exclusion, beta diversity, and deterministic vs. stochastic drivers of
- 571 community assembly. Ecology Letters 17:1400–1408.
- 572 Si, X., A. Baselga, F. Leprieur, X. Song, and P. Ding. 2016. Selective extinction
- 573 drives taxonomic and functional alpha and beta diversities in island bird
- assemblages. Journal of Animal Ecology 85:409–418.
- 575 Simberloff, D. S. 1969. Experimental zoogeography of islands: a model for insular
- 576 colonization. Ecology 50:296–314.
- 577 Simberloff, D. S., and E. O. Wilson. 1969. Experimental zoogeography of islands: the
- 578 colonization of empty Islands. Ecology 50:278–296.
- 579 Simberloff, D. S., and E. O. Wilson. 1970. Experimental zoogeography of islands: a

580	two-year record	of colonization.	Ecology 51:934–937.
-----	-----------------	------------------	---------------------

- 581 Stegen, J. C., A. L. Freestone, T. O. Crist, M. J. Anderson, J. M. Chase, L. S. Comita,
- H. V. Cornell, et al. 2013. Stochastic and deterministic drivers of spatial and
- temporal turnover in breeding bird communities. Global Ecology and
- 584 Biogeography 22:202–212.
- 585 Steinbauer, M. J., R. Field, J. Grytnes, P. Trigas, C. Ah-Peng, F. Attorre, H. J. B.
- 586 Birks, et al. 2016. Topography-driven isolation, speciation and a global increase
- of endemism with elevation. Global Ecology and Biogeography 25:1097–1107.
- 588 Tuomisto, H. 2010*a*. A diversity of beta diversities: straightening up a concept gone
- awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity.
- 590 Ecography 33:2–22.
- 591 ——. 2010b. A diversity of beta diversities: straightening up a concept gone awry.
- 592 Part 2. Quantifying beta diversity and related phenomena. Ecography 33:23–45.
- 593 Ulrich, W., A. Baselga, B. Kusumoto, T. Shiono, H. Tuomisto, and Y. Kubota. 2017.
- 594 The tangled link between β and γ -diversity: a Narcissus effect weakens
- statistical inferences in null model analyses of diversity patterns. Global Ecologyand Biogeography 26:1–5.
- 597 Valente, L. M., A. B. Phillimore, and R. S. Etienne. 2015. Equilibrium and
- non-equilibrium dynamics simultaneously operate in the Galápagos islands.
- 599 Ecology Letters 18:844–852.
- 600 Veech, J. A., and T. O. Crist. 2007. Habitat and climate heterogeneity maintain

- 601 beta-diversity of birds among landscapes within ecoregions. Global Ecology and
- 602 Biogeography 16:650–656.
- 603 Wang, Y., Y. Bao, M. Yu, G. Xu, and P. Ding. 2010. Nestedness for different reasons:
- the distributions of birds, lizards and small mammals on islands of an inundated
- lake. Diversity and Distributions 16:862–873.
- 606 Wang, Y., V. Millien, P. Ding, and A. Algar. 2016. On empty islands and the
- 607 small-island effect. Global Ecology and Biogeography 25:1333–1345.
- 608 Warren, B. H., D. Simberloff, R. E. Ricklefs, R. Aguilée, F. L. Condamine, D. Gravel,
- H. Morlon, et al. 2015. Islands as model systems in ecology and evolution:
- 610 prospects fifty years after MacArthur-Wilson. Ecology Letters 18:200–217.
- 611 Weigelt, P., W. Jetz, and H. Kreft. 2013. Bioclimatic and physical characterization of
- 612 the world's islands. Proceedings of the National Academy of Sciences
- 613 110:15307–15312.
- 614 Weigelt, P., M. J. Steinbauer, J. S. Cabral, and H. Kreft. 2016. Late Quaternary
- 615 climate change shapes island biodiversity. Nature 532:99–102.
- 616 Whittaker, R. H. 1960. Vegetation of the Sisiyou Mountains, Orgeon and California.
- 617 Ecological Monographs 30:279–338.
- 618 Whittaker, R. H. 1965. Dominance and diversity in land plant communities. Science
- 619 147:250–260.
- 620 Whittaker, R. J., K. A. Triantis, and R. J. Ladle. 2008. A general dynamic theory of
- 621 oceanic island biogeography. Journal of Biogeography 35:977–994.

622	Wu, L., X.	Si, R. K.	Didham, D.	Ge, and P. Ding.	2017. Dis	persal modality
-----	------------	-----------	------------	------------------	-----------	-----------------

- 623 determines the relative partitioning of beta diversity in spider assemblages on
- 624 subtropical land-bridge islands. Journal of Biogeography 44:2121–2131.
- 625 Xiao, X., D. J. McGlinn, and E. P. White. 2015. A strong test of the Maximum
- Entropy Theory of Ecology. The American Naturalist 185:E70–E80.
- 627 Xiao, X., J. P. O'Dwyer, and E. P. White. 2016. Comparing process-based and
- 628 constraint-based approaches for modeling macroecological patterns. Ecology
- 629 97:1228–1238.
- 630 Zillio, T., I. Volkov, J. R. Banavar, S. P. Hubbell, and A. Maritan. 2005. Spatial
- 631 scaling in model plant communities. Physical Review Letters 95:1–4.

633 Table 1. Models fitted to the presence-absence patterns of 93 bird species in 42

			Number of		R	2
Model	ΔΑΙϹ	AUC	parameters	Richness	Dissimilarity	Turno
non-neutral.GLM	-2627	0.95	97	0.91	0.37	0.53
neutral.GLM	-172	0.61	5	0.90	0.01	0.45
neutral.exponential	-8	0.57	2	0.41	0.01	0.17
neutral.inverse ratio	1587	0.60	1	0.69	0.04	0.46
Total observation $= 390$	6					

634 islands in the Thousand Island Lake, China.

635 Note: For the neutral exponential model: $c = \exp(-aI)$ and $e = \exp(-bA)$. For the

636 neutral inverse ratio model: $c = \frac{a}{I}$ and $e = \frac{b}{A}$. Dissimilarity is the sum of turnover

637 component and nestedness component. ΔAIC are calculated by subtracting the AIC of

the null model (GLM with only one intercept) from the AIC of the fitted model. AUC

639 are calculated with a Mann-Whitney U statistic.

640 Figure legends

Figure 1. Expected pairwise beta diversity of two identical islands (same colonization and extinction rates), conditioning on both islands having species (i > 0, j > 0). Results are derived from the joint presence-absence distribution of two islands and shown for different mainland pool sizes. Blue dashed line in panel (a) is the analytical solution $E(Jaccard.dissimilarity) = \frac{2\theta}{1+2\theta}$ for the case of at least one island having species (N > 0). Relative extinction rate is the ratio of extinction rate and colonization rate.

649 Figure 2. Expected pairwise Jaccard beta diversities of two islands differing in 650 colonization and/or extinction rates, conditioning on both islands having species (i > 0, j > 0). The blue dashed line in panel (a), (e) and (i) are the analytical solutions of 651 $E(Jaccard.dissimilarity) = \frac{(1+z)\theta}{1+(1+z)\theta}$, conditional on at least one island having 652 653 nonzero species. Expectations are provided for different levels of z, i.e. the number of 654 times relative extinction rate on the second island exceeds that on the first. The black 655 dashed lines in panel (d), (h) and (l) represent a 1:1 ratio. Different lines in a graph 656 represent different mainland pool sizes (for color legend see Fig. 1). 657

Figure 3. Empirical evaluation of the theoretically fitted alpha and beta diversity
patterns for the Thousand Islands Lake bird dataset. (a) - (l) are predictions of three
neutral models, (m) - (p) are predictions of a non-neutral model. Solid black lines

- 661 represent a 1:1 relationship and dashed red lines are fits of ordinary linear regressions,
- 662 Point densities from high to low are color-coded from red to blue.

Figure 1

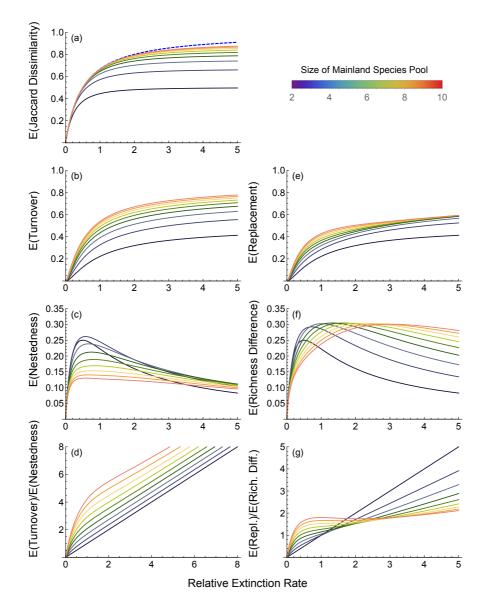


Figure 2

