University of Groningen

Can we ignore trait-dependent colonization and diversification in island biogeography?

Xie, Shu; Valente, Luis; Etienne, Rampal S

Published in:
Evolution

DOI:
10.1093/evolut/qpad006

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2023

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Xie, S., Valente, L., & Etienne, R. S. (2023). Can we ignore trait-dependent colonization and diversification in island biogeography? *Evolution*, *77*(3), 670–681. Advance online publication. https://doi.org/10.1093/evolut/qpad006

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 01-11-2023

Evolution, 2023, **77(3)**, 670–681
https://doi.org/10.1093/evolut/qpad006
Advance access publication 13 January 2023
Brief Communication

Can we ignore trait-dependent colonization and diversification in island biogeography?

Shu Xie[1], **Luis Valente**[1,2], **Rampal S. Etienne**[1]

[1]Groningen Institute for Evolutionary Life Sciences, University of Groningen, Groningen, the Netherlands
[2]Naturalis Biodiversity Center, Leiden, the Netherlands

Corresponding author: Groningen Institute for Evolutionary Life Sciences, University of Groningen, 9700CC, Groningen, the Netherlands. Email: s.xie@rug.nl

Abstract

The application of state-dependent speciation and extinction models to phylogenetic trees has shown an important role for traits in diversification. However, this role remains comparatively unexplored on islands, which can include multiple independent clades resulting from different colonization events. To explore whether assuming no dependence on traits leads to bias in inference on island dynamics, we extend an island biodiversity model, DAISIE (Dynamic Assembly of Islands through Speciation, Immigration, and Extinction) to include trait-dependent diversification simulations, and evaluate the robustness of the inference model which ignores this trait-dependence. Our results indicate that when the differences between colonization, extinction, and speciation rates between trait states are moderate, the model shows negligible error for a variety of island diversity metrics, suggesting that island diversity dynamics can be accurately estimated without the need to explicitly model trait dependence. We conclude that for many biologically realistic scenarios with trait-dependent diversification and colonization, this simple trait-less inference model is informative and robust to trait effects on colonization, speciation, and extinction. Nonetheless, our new simulation model may provide a useful tool for studying patterns of trait variation.

Keywords: island biogeography, trait-dependent diversification, phylogenetic inference, robustness analysis

Introduction

Understanding the mechanisms underlying species richness variation among clades is a key challenge in evolutionary biology. One of the hypotheses to explain diversity differences between clades and regions is variation in diversification rate, where higher net diversification rates are likely to lead to species-rich clades (McPeek & Brown, 2007). Recent studies support this hypothesis, suggesting that diversification rate variation can explain much of the richness variation among major clades under different taxonomic ranks (Kozak & Wiens, 2016; Li & Wiens, 2022; Scholl & Wiens, 2016). However, this begs the question how this variation in diversification rates arises. Diversification rates might be affected by environmental, clade-specific, or historical factors (Kisel et al., 2011; Tietje et al., 2022). Comparative analyses have suggested that traits play an important role in diversification rates (Chevin, 2016; Cracraft, 1985; Jablonski, 2008; Rabosky & McCune, 2010; Simpson, 2013; Stanley, 1975), but this can be difficult to test. Typically, when a certain trait state occurs more frequently in a speciose clade than in a species-poor clade, we may be tempted to conclude that this state promotes speciation or reduces extinction. However, such a pattern may be due to the evolutionary conservation of the trait (Rabosky & Goldberg, 2015). An excess of species with a particular state may also be due to asymmetrical transition rates between states (Burin et al., 2016; Goldberg & Igić, 2012).

A geographical setting where traits have long been proposed to influence diversity is islands (Cowie & Holland, 2008; Parent et al., 2008). Biogeographical and environmental characteristics of oceanic islands may lead to strong phenotypic divergence from the mainland, often resulting in island syndromes (Auffret et al., 2017; Cássia-Silva et al., 2020; García-Verdugo et al., 2020), such as flightlessness in birds, dwarfism and gigantism in vertebrates, woodiness in plants, and small size and inconspicuous coloring of flowers (Hetherington-Rauth & Johnson, 2020). At the same time, oceanic islands are home to some of the most extraordinary radiations, such as Darwin's finches (Losos & Ricklefs, 2009), which may lead to large richness variation among clades within an insular system (Patiño et al., 2017). A key question is whether certain traits or trait states have played a role in the presence or absence of rapid radiations. While insular radiations are thought to be mostly driven by increased ecological opportunity on islands, it has long been hypothesized that some traits may trigger, facilitate, or hinder diversification in an insular setting (García-Verdugo et al., 2014; Zhu et al., 2020). Some characteristics of species, such as seed size in plants and flight ability in birds, evidently affect the chances of species colonizing an island or an isolated habitat (Onstein et al., 2017). Furthermore, after the successful colonization of an island, changes in morphological characteristics occur, facilitated by ecological release (Losos et al., 1997; Millien, 2006). These character changes have been shown to affect *in situ* diversification rates of species, which are important to address the evolutionary assembly on islands (Aleixandre et al., 2013; Biddick et al., 2019; Burns, 2016).

Received January 10, 2022; revisions received December 24, 2022; accepted January 11, 2023

The increasing availability of molecular phylogenies has stimulated the development of statistical methods to detect how traits are associated with diversification (Mitter et al., 1988). A likelihood-based framework, based on the binary-state-dependent speciation and extinction model (BiSSE), resolved the shortcomings of the sister clade comparison method which could only address the variation in net diversification rates (Maddison et al., 2007). In addition, BiSSE inspired a large number of state-dependent diversification models, which are known as the state-dependent speciation and extinction (SSE) model family. These models extend BiSSE in various ways to enable state-dependent analyses to infer state-dependent diversification under a variety of scenarios or to model more complex phenotypic traits (Fitzjohn, 2010; Fitzjohn et al., 2009; Goldberg & Igić, 2012; Goldberg et al., 2011). Furthermore, models incorporating hidden traits were developed to overcome high type I errors with BiSSE (Beaulieu & O'Meara, 2016; Herrera-Alsina et al., 2019). However, despite the fundamental role of islands in the conceptual development of trait diversification theory, investigations of trait diversification dynamics on islands in a phylogenetic context are rare compared to those for continental phylogenies.

Current island biodiversity models focus mostly on the associations between island characteristics (e.g., area and isolation) and diversification rates (Algar & Losos, 2011). These models help to understand the species richness variation between different islands or between islands and the mainland (Leidinger & Cabral, 2017). However, mostly for practical reasons, they tend to ignore traits, even though traits can be important factors that affect diversification rates, and in turn lead to richness variation between clades within the same island (Supsup et al., 2020; Zheng & Zheng, 2022). Currently, there is no likelihood-based inference model focusing on phylogenetic data from islands that incorporates both trait dynamics and their effects on diversification rates (an island version of an SSE model, or an SSE version of an island biogeography model). The current practice of using models that ignore traits may lead to erroneous estimates of parameters (colonization, speciation, and extinction) if species traits help shape the phylogenetic trees on islands. But perhaps such models still do a reasonable job in describing diversification dynamics. While trait dependence can be relatively easily incorporated in simulations, deriving a likelihood of a model that includes trait-dependent and diversity-dependent diversification does not seem feasible, and hence it is all the more important to explore to what extent this is problematic for inference.

Here, we aim to determine to what extent we can trust existing island biodiversity models, by answering the following questions. In the absence of a method to estimate trait state-dependent colonization, extinction, and speciation rates (CES rates) from island communities, can we still obtain meaningful results regarding island diversification using current island biogeography models? And under what conditions can trait-less biodiversity models still be used to make accurate predictions of island diversity, distribution of island clade sizes, and diversity changes through time? To address these questions, we extend a phylogeny-based island biogeography simulation framework, DAISIE (Dynamic Assembly of Islands through Speciation, Immigration, and Extinction) to include trait-dependent diversification rates, and evaluate the robustness of the trait-independent inference model data. DAISIE (Box 1) is the first island biodiversity model that uses phylogenetically informed data on island assembly in a whole-community framework. It allows estimation of colonization and diversification, as well as simulation of the dynamic assembly of insular communities, which can include multiple phylogenetic trees resulting from several colonization events of an island (e.g., all mammals on an island) (Valente et al., 2015, 2018; 2020).

In this study, we assess whether this simple model can accurately reconstruct diversity and phylogenies on islands without considering trait effects on colonization, speciation, and extinction, that is, whether we can ignore trait dependence in phylogenetic studies of island biogeography. We investigate whether, in the presence of trait dependence in the model generating the data, the inference model that assumes no trait dependence can still accurately infer island diversity dynamics through time. If so, this will suggest that simple models do a good job of explaining island diversity and that reliable analyses can be performed even without trait data of island species (which are often absent, incomplete, or difficult to obtain), i.e., the model is robust to trait dependence. If, instead, we find that under certain conditions the existence of trait dependence substantially alters the predictions of the model, this will suggest that traits cannot be ignored in these cases and that a new estimation approach is needed.

Methods

State-dependent and state-independent simulation models

The DAISIE inference framework uses maximum likelihood to estimate colonization and diversification rates of insular biota from phylogenetic information. The core version of DAISIE assumes that all island species share the same CES rates, and the model is essentially neutral at the species level (Valente et al., 2015). However, the dynamics of trait-dependent diversification and colonization within island clades are not modeled.

Here, we introduce a trait state-dependent island biodiversity simulation model, an extension of the DAISIE simulation model combining it with features of the BiSSE model (Figure 1). To distinguish the two simulation models, the new simulation model is termed the *state-dependent simulation* (SDS) model (Figure Box 1B), and the original trait-less DAISIE simulation model is termed the *state-independent simulation* (SIS) model (Figure Box 1A). Likewise, we will call the standard DAISIE inference model the *state-independent inference* (SII) model. A state-dependent inference model does not yet exist. In the SDS model, the rates of all evolutionary processes are trait-state-dependent. For simplicity, we consider a binary trait with two states, 1 and 2. Species in the same state have the same CES rates, while species with different states may differ in one or more rates. Mainland species can be regarded as forming two assemblages according to their trait states. Immigration of species in each assemblage to the island is determined by the number of mainland species in each state (M_1 and M_2) and their colonization rates (γ_1 and γ_2). Once an immigrant species (which inherits the trait state from its mainland ancestor) successfully colonizes the island, it can undergo population divergence from the mainland population (via anagenesis λ_1^a and λ_2^a), *in situ* speciation (via cladogenesis λ_1^c and λ_2^c) or extinction (μ_1 and μ_2). Island species can shift between trait states at a certain rate of transition from state 1 to state 2 (q_{12}) or from state 2 to state 1 (q_{21}). The transition rates can be equal or

Box 1. Description of the DAISIE framework.

DAISIE (Dynamic Assembly of Island biota through Speciation, Immigration, and Extinction) is a framework to study biological communities on islands in a phylogenetic context. It is available as an R package.

DAISIE simulation model. DAISIE can simulate phylogenetic datasets of island communities using a set of colonization, speciation, and extinction (CES rates). The island can be colonized from a mainland pool of species. The model assumes that each mainland species can colonize the island at a rate γ, and undergo speciation at a rate λ or extinction at a rate μ. Two processes of speciation are considered: anagenesis (λ^a, island population diverges from the mainland ancestral due to long-term geographical isolation) and cladogenesis (λ^c, one island species split into two new endemic species). Cladogenesis and colonization rates can be diversity-dependent, declining linearly with the number of species accumulating on the island. The diversity-dependence parameter K' can be regarded as the maximum number of species niches in each clade (species from the same mainland ancestor) when there is no extinction. Thus, in the DAISIE framework, the carrying capacity can be regarded as the strength of the effect of crowding on reducing colonization and diversification rates. The original DAISIE simulation model is state-independent (SIS). In this study, we add an additional state-dependent simulation (SDS) model with state-dependent CES rates (Figure Box 1B).

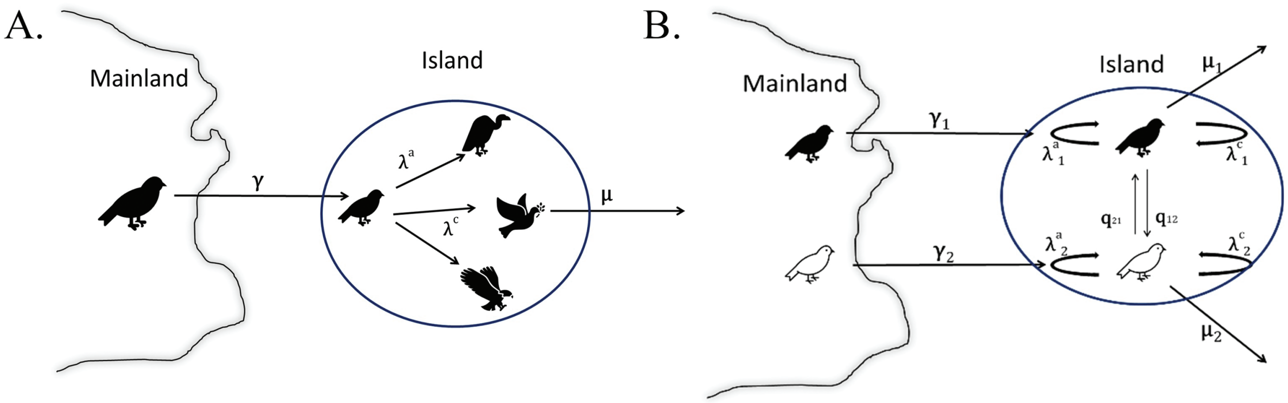

Figure Box 1. Graphical visualization of (A) state-independent simulation (SIS) model and (B) state-dependent simulation (SDS) model.

DAISIE inference model. DAISIE's maximum likelihood inference method uses information from colonization and branching times extracted from time-calibrated phylogenetic trees (resulting from different colonization events of an island). The approach has been shown to accurately estimate island colonization, speciation, and extinction rates, as well as the clade-level carrying capacity.

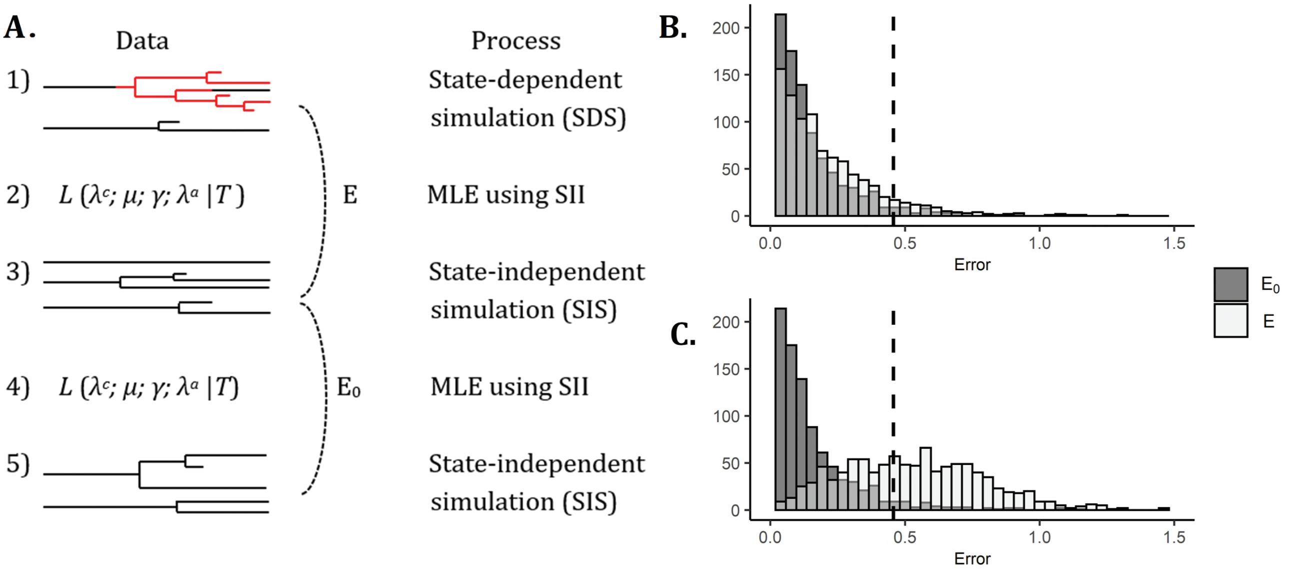

Figure 1. (A) Schematic representation of the robustness pipeline. (1) Simulate phylogenetic data with the SDS model. The binary states are represented by two different colors (red and black). (2) Use the data obtained from step 1 to estimate parameters with the SII model. (3) Simulate data using the SIS model with parameters estimated in step 2. (4) Use the SII model again to estimate parameters. (5) Simulate data using the SIS model with the estimated parameters from step 4. E_0—baseline error when simulation and inference model are identical; E—error when simulation and inference model differ. (B and C) Visualization of E and E_0 distributions. The light gray and the dark gray bars correspond to the distribution of E and E_0, respectively. The black dashed line is the 95th percentile of the baseline error distribution. ED_{95} is the percentage of the distribution E that lies on the right side of the dashed line. In (B), the E and E_0 distributions almost overlap, and the ED_{95} in (B) is lower than that of (C).

different. In speciation via anagenesis or cladogenesis, daughter lineages are assumed to inherit the trait state of their parent species, which means no state shifts occur during speciation (Fitzjohn, 2010). Furthermore, transitions are regarded as intraspecific changes, which occur instantaneously along lineages, thus assuming that the period of time in which two states coexist in a polymorphic species is negligible. Transition and speciation events are not allowed to occur simultaneously in this model (Maddison et al., 2007). The equations for calculating the rates are given in the Supplementary Methods.

In both SIS and SDS simulations, we considered diversity-dependent (DD) and diversity-independent (DI) models. In DI models, all the rates are diversity-independent. In DD models, colonization and cladogenetic speciation rates are diversity-dependent, while the other rates are diversity-independent. We implemented a clade-specific DD model, assuming that diversity-dependence only operates between species in the same island clade, i.e., those that descend from the same mainland ancestor (Etienne et al., 2022). We assume that differences in resource utilization due to phylogenetic distance in species belonging to different clades are sufficient to prevent competition between clades. In addition, because there is no strong evidence of an association between trait-dependence and diversity-dependence, in most of our simulations we assumed, for the sake of simplicity, that diversity-dependence is not trait-dependent, i.e., the diversity carrying capacity is the same regardless of the trait state of the species undergoing speciation or colonization. In addition, we also implemented simulations with different diversity carrying capacities for each state. The equations for calculating CES rates for this case are shown in Supplementary Table S1. We implemented the SDS model in the R package DAISIE (Etienne et al., 2020).

The SDS results record the evolutionary history of island species including their colonization and branching times, as well as the richness dynamics of each trait state. We assumed that the inference accuracy of the DAISIE model will be poorer for larger inequality between the numbers of species in each of the two states. To test this assumption, we used the tip ratio (Davis et al., 2013), which we denoted by r, as the number of species in state 2 divided by the number of species in state 1:

$$r = \frac{N_2}{N_1}$$

We used seven island diversity metrics to evaluate the simulated phylogenies from the SDS and the SIS models. Four metrics were used to measure diversity at the end of the simulation: total number of species (N_{Spec}), number of lineages present on the island (N_{Col}) (resulting from independent colonization events), standard deviation of clade size (σ_{CS}), and standard deviation of colonization time among clades (σ_{CT}). The other three metrics measured richness changes through time: total species richness through time (SRTT) (Supplementary Figure S1); endemic species richness through time (ESRTT); and non-endemic species richness through time (NESRTT) (Neves et al., 2022). We also studied the parameter estimates to evaluate how similar the inferred parameters are to the ones used to generate the data.

Simulation scenarios

The mainland pool of 1000 species is assumed to be evenly distributed with 500 species in each state. This is not a strong limitation of our simulations because it is the difference in total colonization rate (the product of mainland pool size and per capita colonization rate) that matters (Valente et al., 2015). In addition, we set a limit of 20 ($K' = 20$) species for each clade for the diversity-dependent model. To measure the effect of transition rates independent of CES rates, we used a symmetric scenario as a control, where all the CES rates are symmetric between binary states (Table 1A). We chose two values for each CES rate, a low one and a high one, in such a way that the total number of species remains between realistic values of 50–150. We set four types of transitions between binary states (Table 1B). For high and low transition rates we used 0.2 and 0.02, respectively. The symmetric scenario consists of 128 combinations of CES rates, transition rates, and carrying capacities (Table 1A).

To investigate the effect of trait dependence, we ran the SDS simulation under a series of scenarios with varying degrees of asymmetry in CES rates. In these scenarios, the mean values of CES rates between binary states were kept the same as in the symmetric scenario, as well as two gradients of mean rates for each parameter (low and high, Table 1A). For the analyses with asymmetry in rates, only one CES rate was set to be the asymmetric per scenario, and all the others were kept symmetric (Table 1C).

We defined the relative rate differential (RRD) as the difference in rates for the two trait states by their mean rate value, to measure the asymmetry level between states (Gamisch, 2016).

$$RRD = \frac{|rate_2 - rate_1|}{\frac{1}{2}(rate_2 + rate_1)}$$

RRD is 0 when the rates are symmetric for the two states, and a larger RRD means a larger rate difference between states. For each CES rate, we ran analyses with three different levels of RRD: 0.5, 1, and 1.5 (Table 1C). Therefore, in total, we ran 13 scenarios: 1 control plus 3 asymmetric scenarios for each of the 4 parameters (rates of cladogenesis, extinction, colonization, and anagenesis) (Table 1C). We did not consider combinations of asymmetric rates as the number of parameter combinations would have been computationally prohibitive. We set CES rate values of state 2 to always be higher or equal to the values of state 1 for all the asymmetric cases to avoid redundancy. In asymmetric scenarios, transition rates were chosen in the same way as for the symmetric scenario, with four transition types (Table 1B). We show the exact parameter values that we used in this paper in Table 1D.

Robustness analysis

We aimed to test whether ignoring trait dependence in inference affects the ability of the SII model to reconstruct diversity dynamics on islands generated by a simulation model that assumes an effect of trait states on CES rates. We used a computational pipeline (Figure 1) to measure the error of the SII model when real insular diversity dynamics involve trait dependence in rates (SDS), adapting the approach of Neves et al. (2022). The pipeline allows investigating the robustness when the inference model does not match the generating (simulation) model. First, we simulated 1000 replicates ("islands") for each parameter set (Table 1), under the SDS model. We converted the SDS results to the SIS output format (i.e., we removed trait-related information), and estimated CES rates and carrying capacity with the SII model. We then used the estimated parameters to simulate with the SIS model, again estimated the parameters for these simulations with the SII model, and used the obtained parameters to simulate a second set of SIS results. Because the SII model can infer diversity

Table 1. Parameter space for the state-dependent simulations. (A) The symmetric (control) scenario with all the CES rates symmetric between binary states, consisting of each combination of the parameters. Numbers separated by comma are the low and high values, respectively. (B) Four transition scenarios with different combinations of q_{12} and q_{21}. (C) The 13 scenarios of asymmetric CES rates (λ^c, λ^a, γ, and μ). The control scenario contains the parameter sets where all the CES rates are symmetric (RRD = 0). For each asymmetric scenario, only one of the four rates is considered to be asymmetric (RRD ≠ 0). The degree of asymmetry is represented by the relative rate differential (RRD). Larger RRD value indicates larger rate difference between states. (D) The exact values of RRD that are used in SDS simulations. As a guide for reading the tables, consider scenario 2 in (C) part of table, which assumes asymmetry only in cladogenesis ($\mathrm{RRD}\lambda^c = 0.5$, $\mathrm{RRD}\mu = 0$, $\mathrm{RRD}\gamma = 0$, and $\mathrm{RRD}\lambda^a = 0$), if the mean rates are $\lambda^c mean = 0.2$, $\lambda^a mean = 0.1$, $\gamma mean = 0.008$, and $\mu mean = 0.2$ (A, low values), we can find the exact parameter values (in bold) in (D) part of table, which are $\lambda^c 1 = 0.15$, $\lambda^c 2 = 0.25$; $\lambda^a 1 = \lambda^a 2 = 0.1$; $\gamma 1 = \gamma 2 = 0.008$; and $\mu 1 = \mu 2 = 0.2$. We consider all the four transition types, which are shown in (B) part of table, for this parameter set.

(A)	Parameters	Parameter values								
	Time	5								
	M1	500								
	M2	500								
	$\lambda^c_1 = \lambda^c_2$	0.2, 0.4								
	$\mu_1 = \mu_2$	0.1, 0.2								
	$\gamma_1 = \gamma_2$	0.008, 0.012								
	$\lambda^a_1 = \lambda^a_2$	0.2, 0.4								
	K'	20, ∞								
(B)	**Transition types**	q_{12}	q_{21}							
	low q_{12} low q_{21}	0.02	0.02							
	high q_{12} low q_{21}	0.2	0.02							
	low q_{12} high q_{21}	0.02	0.2							
	high q_{12} high q_{21}	0.2	0.2							
(C)	**Symmetry of rates**	**Scenario**	**RRD of cladogenesis (λc)**		**RRD of extinction (μ)**		**RRD of colonization (γ)**		**RRD of anagenesis (λa)**	
	Symmetry (control)	1	0		0		0		0	
	Asymmetry in cladogenesis	2	0.5		0		0		0	
		3	1		0		0		0	
		4	1.5		0		0		0	
	Asymmetry in extinction	5	0		0.5		0		0	
		6	0		1		0		0	
		7	0		1.5		0		0	
	Asymmetry in colonization	8	0		0		0.5		0	
		9	0		0		1		0	
		10	0		0		1.5		0	
	Asymmetry in anagenesis	11	0		0		0		0.5	
		12	0		0		0		1	
		13	0		0		0		1.5	
(D)	CES rate	Mean value	RRD = 0		RRD = 0.5		RRD = 1		RRD = 1.5	
			State1	State2	State1	State2	State1	State2	State1	State2
	λ^c	0.2	0.2	0.2	**0.15**	**0.25**	0.1	0.3	0.05	0.35
		0.4	0.4	0.4	0.3	0.5	0.2	0.6	0.1	0.7

Table 1. Continued

(A)	Parameters	Parameter values								
	μ	0.1	**0.1**	**0.1**	0.075	0.125	0.05	0.15	0.025	0.175
		0.2	0.2	0.2	0.15	0.25	0.1	0.3	0.05	0.35
	γ	0.008	**0.008**	**0.008**	0.006	0.010	0.004	0.012	0.002	0.014
		0.012	0.012	0.012	0.009	0.015	0.006	0.018	0.003	0.021
	λ^a	0.2	**0.2**	**0.2**	0.15	0.25	0.1	0.3	0.05	0.35
		0.4	0.4	0.4	0.3	0.5	0.2	0.6	0.1	0.7

dependence by estimating carrying capacity, we used the DD model for both inference processes in the pipeline. We studied several metrics of all simulation results (see below). The difference between the metrics of the two SIS results gives the baseline error (E_0) in the inference that occurs even if the inference model is identical to the simulation model (SII = SIS). The difference between the metrics of the SDS results and the first SIS results gives the error E that occurs when the simulation and inference model differ (SII ≠ SDS) (Figure 1).

We calculated the errors E and E_0 of seven metrics for each replicate, which resulted in two error distributions for each parameter set. E and E_0 for N_{Spec}, N_{Col}, σ_{CS}, and σ_{CT} were calculated as the absolute difference between simulations, while for the three diversity-through-time metrics (SRTT, ESRTT, and NESRTT), errors were calculated using the ΔnLTT (normalized lineage-through-time) statistic (Janzen et al., 2015), by integrating the absolute distance between two diversity-through-time curves (Supplementary Figure S1). ΔnLTT is equal to zero only when the two simulated nLTT curves are identical. To compare the divergence between the distribution of the two errors, E and E_0, we used a metric, ED_{95} (Neves et al., 2022). ED_{95} is the percentage of the distribution of E that exceeds the 95% percentile of the distribution of E_0 (Figure 1B). We use ED_{95} to tease apart inference errors that are due to limited information in the data and even occur when the inference model is identical to the generating model. Higher ED_{95} values indicate larger differences between E and E_0, in other words, the error caused by trait effects is larger than the baseline error inherent to DAISIE. The ED_{95} of all the seven island diversity metrics were calculated for each parameter set. We also calculated Hellinger distance, a metric used to quantify the similarity between two probability distributions, to compare the difference between E and E_0 for each statistical metric (see Supplementary Material) (Csiszár & Shields, 2004).

To evaluate the inference accuracy, we used the same method to calculate the ED_{95} of the estimation errors. The baseline error E_0 for each estimated rate was calculated as the absolute difference between the two SII estimations. The mean rates of the two states are used as reference, to compare with the estimated values from the first SII results, to obtain the estimation error E, which measures the bias caused when the generating model is different from the inference model. We ran a total of 1664 parameter sets (128 for each of the 13 scenarios), with 1000 replicates for each parameter set. The calculation of metrics and the error analysis were implemented in the R package DAISIErobustness (Lambert et al., 2022).

Results

We find that DAISIE is quite robust to trait dependence in rates of colonization and cladogenesis. The inference errors are negligible for all metrics except ΔNESRTT (non-endemic richness through time) and σ_{CS} (clade size standard deviation), and they are only affected when cladogenetic speciation rates are asymmetric. In addition, diversity-dependence and state-dependent transition rates have a negligible effect on the inference errors except under asymmetric cladogenesis. Surprisingly, the errors are not strongly related to the tip ratio, i.e., the ratio of the diversities between two trait states, but they are positively correlated with the variation in clade sizes. We now present these results in more detail. We note that the results using the metrics ED_{95} and

Hellinger Distance are highly similar, so in the main text we only discuss the results of ED_{95}, and the comparisons with Hellinger Distance are shown in Supplementary Figures S8 and S9.

Effect of trait-dependent CES rates on inference

In general, using the SII model to estimate parameters and subsequently the SIS model to simulate with the obtained parameters causes minimal error in reconstructing the SDS diversity patterns for most of the parameter sets (Figure 2). Most metrics measuring the diversity at present (N_{Spec}, N_{Col}, and σ_{CT}) show the minimal difference between E and E_0 for all scenarios (Figure 2). However, among the three ΔnLTT statistical metrics, the error in ΔNESRTT is large when there is a high asymmetry in the cladogenesis rate (RRD = 1.5) (Figure 2). σ_{CS} shows the largest error difference among all the metrics and is positively correlated to the RRD of the cladogenesis rate. Even in the symmetric case, most of the ED_{95} in σ_{CS} is higher than 0.05. In addition, we test the error in σ_{CS} using the SIS model in the whole pipeline to confirm if the symmetric case performs the same as the null model (SIS model). For all the seven metrics, including σ_{CS}, the results of using the symmetric SDS model and using the SIS model are highly similar (Supplementary Figure S11).

Because only asymmetry in the cladogenesis rate has a substantial effect, we zoom in on the comparison of the results of the symmetric scenario with the three scenarios that have different asymmetry levels in the cladogenesis rate. Large inference error in ΔNESRTT only occurs when the mean cladogenesis rate is greater than the anagenesis rate (Supplementary Figure S2), which leads to fewer non-endemic species and more endemic species at present on the island.

Effects of diversity-dependence and state-dependent transitions on inference

Diversity-dependence and state-dependent transitions have a negligible effect on the inference errors in ΔNESRTT and σ_{CS}, except when cladogenesis rates are largely asymmetric between states (Supplementary Figure S3). Among the four transition types, with low or high, symmetric or asymmetric transitions, parameter sets with higher transition rate from high-rate state to low-rate state ("low q_{12} high q_{21}") cause a larger error in σ_{CS} than in the reverse direction ("high q_{12} low q_{21}") (Supplementary Figure S3). Considering trait dependence and diversity dependence simultaneously, i.e., a different capacity for each trait state, has negligible effects on the performance of the model (Supplementary Figure S10).

Parameter estimation

In this study, we consider only cases where only one of the CES rates is asymmetric between binary states, while the other parameters are equal between states. Therefore, the estimation error in the CES rates that are symmetric can be easily calculated from the generating parameters and the estimated values. For the asymmetric parameter in each scenario, we use the mean rates between states in the SDS model as a reference. Note that the reference value for the asymmetric parameter is not the true value, and the bias based on the mean value is not an error.

When comparing the mean parameter values for simulations using the SDS model and the parameters inferred from the SDS results, the colonization and extinction rates are well estimated for most of the parameter sets (Supplementary Figures S6 and S7). However, nearly all scenarios show a systematic bias, with underestimated anagenesis and overestimated cladogenesis rates when asymmetry in cladogenesis

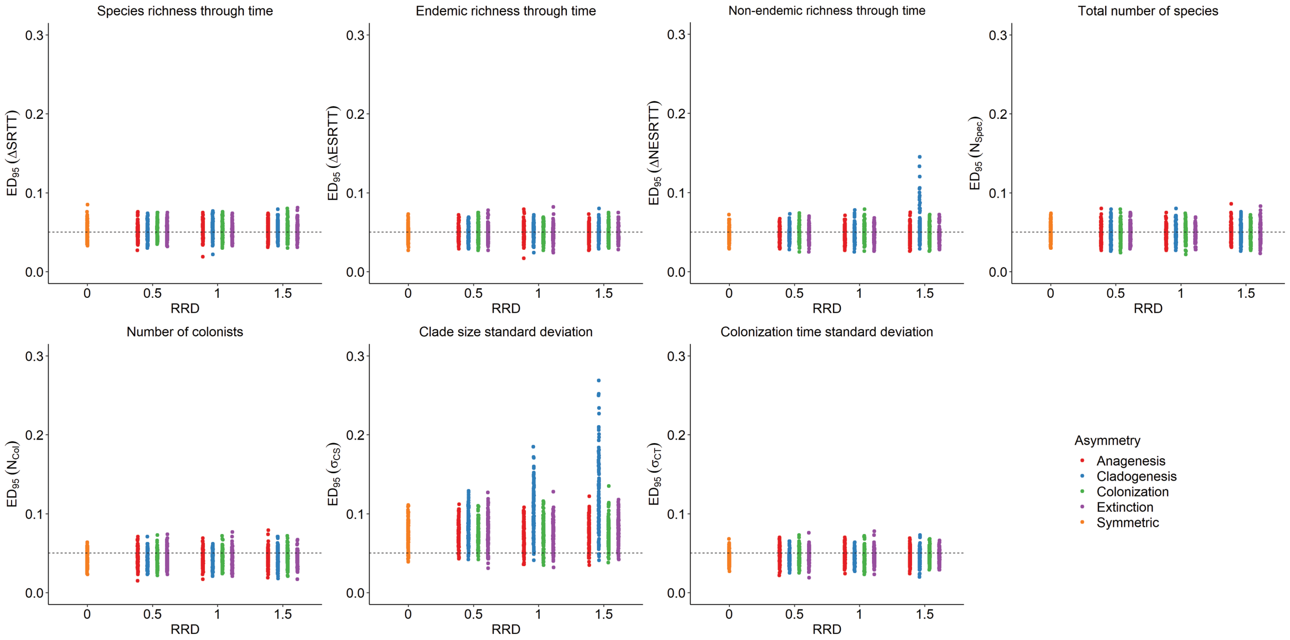

Figure 2. Distribution of ED_{95} for seven island diversity metrics. Each point corresponds to a parameter set. The orange points correspond to the parameter sets in the control scenario with all CES rates symmetric, and the other colors correspond to the 12 scenarios with asymmetric rates in different degrees. The *x*-axis shows the degree of asymmetry (RRD) of each CES rate. The dashed line at 0.05, indicates the expected ED_{95} for the null model.

rate. The inference errors are positively correlated with the bias in cladogenesis and anagenesis when cladogenesis rates are largely asymmetric between states (Supplementary Figures S6 and S7). However, the baseline error in the anagenesis rate for most of the parameter sets is high, resulting in low ED_{95} for the anagenesis rate.

In general, asymmetry in CES rates has a negligible effect on the estimation accuracy of the parameters that are symmetric between states (Figure 3). For example, in scenarios where only the cladogenesis rate is asymmetric, the estimation accuracy is high for the other three CES rates which are symmetric (Figure 3A). When cladogenesis rates are very asymmetric between states, this causes higher estimated cladogenesis rates than the mean of the two generating values, especially when the transition rate from low-rate state to high-rate state is high (Figure 3A). This is because the asymmetry in cladogenesis and transition rates results in most of the existing species being in a high-rate state, and hence the effective rate of cladogenesis is actually higher than simply the mean between the two generating values.

Richness variation among clades and among states

We calculated the richness variation among clades (clade size variation) and the richness among states (tip ratio), using the median value of 1000 replicates to represent the variation for each parameter set. As expected, when all the CES rates and transition rates are symmetric between states, there is nearly no species richness variation between states, with relatively low clade size variation (yellow points in Supplementary Figures S12 and S13). Large cladogenesis rate variation between states in SDS models can result in species richness variation between states, but it can also lead to clade size variation (Supplementary Figure S13). State-dependent transition rates may reinforce or reduce these variations to some extent. To understand what kind of empirical data may cause large errors using the DAISIE model, we checked the relationships between the inference errors with the tip ratio (*r*) and the clade size variation in SDS results. We find that larger inference error does not always occur when the species richness is highly different between states (Supplementary Figure S4). In other words, the tip ratio does not decisively control the robustness of the DAISIE model, because the phylogenies can be accurately reconstructed even with large richness differences between states (Supplementary Figure S4). The clade size variation of SDS results barely affects ΔNESRTT, but substantially affects the inference error in σ_{CS} (Supplementary Figure S5).

In addition, even though asymmetry in colonization rate can also result in large richness variation between states (Supplementary Figure S12), it does not lead to high inference error, because the symmetric cladogenesis rates in these scenarios do not affect the clade size variation. This also suggests that variation among clades, rather than variation among states, is a more likely factor that may cause inference error.

Discussion

Species traits are hypothesized to affect biological assemblages by altering diversification rates (Fitzjohn et al., 2009; Mitter et al., 1988). Our results indicate that not incorporating the

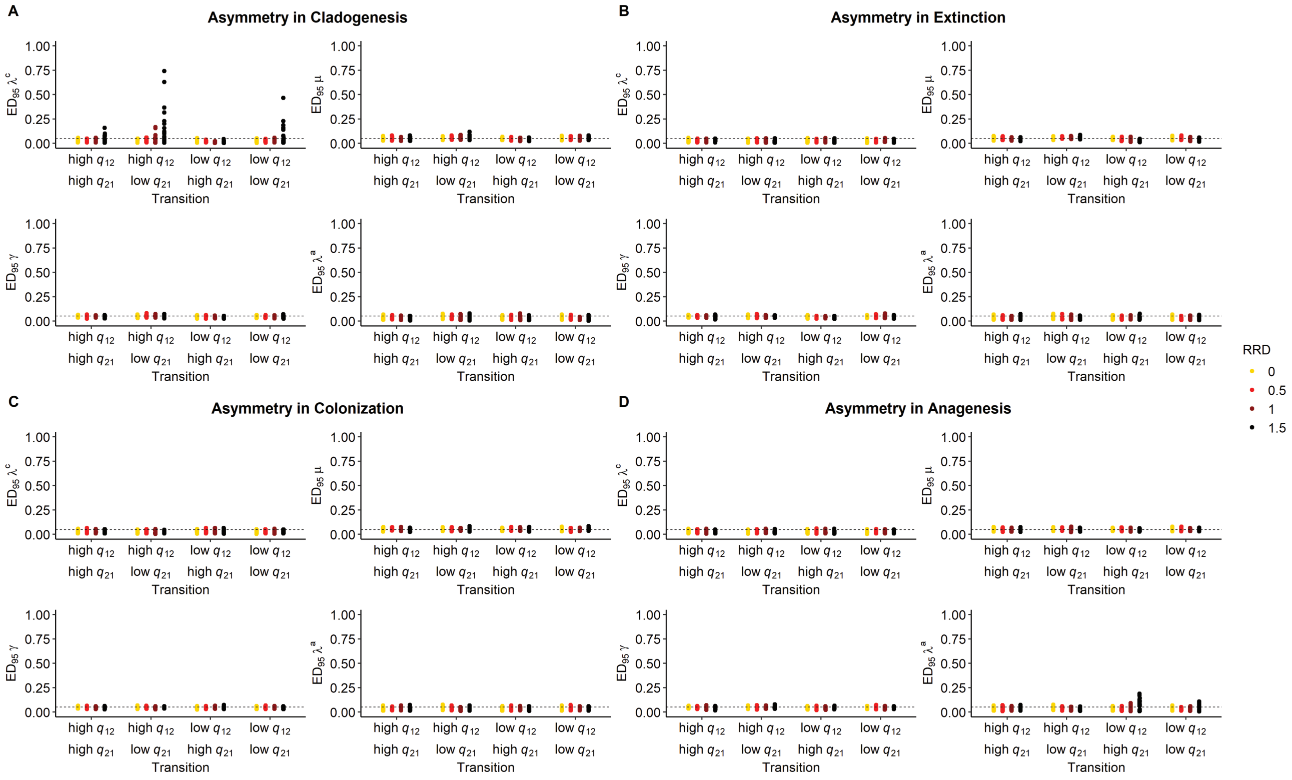

Figure 3. Parameter estimation performance. ED_{95} of the four CES rates across all the parameter combinations. The parameter sets are grouped by RRD and transition types. The colors represent the asymmetry level (RRD) of CES rates between states. The facet plots indicate the combination of the scenarios with different asymmetry levels in (A) cladogenesis rate (scenarios 1 and 2–4 in Table 1C); (B) extinction rate (scenarios 1 and 5–7 in Table 1C); (C) colonization rate (scenarios 1 and 8–10 in Table 1C); (D) anagenesis rate (scenarios 1 and 11–13 in Table 1C).

effect of trait dependence on species diversification in the inference model and subsequently simulating the model with the obtained parameters allows surprisingly accurate estimation of diversification and colonization rates, as well as accurate reconstruction of the evolutionary history of species on an island under a wide range of scenarios. Hence, we conclude that the model is robust to leaving out the details of trait-dependent colonization and diversification. Only in exceptional cases we see large differences between the simulations of a model with trait dependence and a model without. This is specifically the case for two metrics: non-endemic richness through time (ΔNESRTT) and clade size standard deviation (σ_{CS}).

Large differences between endemic species richness and non-endemic species richness may lead to large inference errors in ΔNESRTT. Within the parameter space investigated in this study, a large error in ΔNESRTT occurs only when the mean cladogenesis rate between states is much higher than the mean anagenesis rate (Supplementary Figure S2). In this case, species with the higher cladogenesis rate state can rapidly speciate into a large clade, which leads to the endemic species richness being 5–10 times the non-endemic species richness. Without accounting for trait dependence, the estimated cladogenesis rate is closer to the higher cladogenesis rate than to the mean value of the two states in the SDS model. In contrast, the anagenesis rate is underestimated (Supplementary Figures S6 and S7). This leads to fewer non-endemic species, and more endemic species in the subsequent SIS simulations than in the SDS simulations, resulting in a large error in ΔNESRTT.

The other metric whose estimation is affected by trait dependence is σ_{CS}. When trait states are conserved, and clades with a certain trait state have higher rates of diversification, clades with that trait state will likely become much more species-rich than clades with the other state. In inference, DAISIE assumes that all lineages diversify with the same rates, which generates balanced clades and leads to less clade size variation in the SIS results than in the SDS outputs. In addition, when the cladogenetic speciation rate is diversity-dependent, competition between species in the same clade restricts the increase in species number, preventing clades from growing above a certain diversity level. Therefore, clade size cannot become extremely large, leading to lower error in σ_{CS} in diversity-dependent models than in diversity-independent models. However, we emphasize that even though DAISIE cannot accurately model the fine-scale variation between clades for some exceptional parameter combinations, it can still accurately reconstruct the dynamics of the whole community with multiple independent clades. We attribute the fact that the average level of the error in σ_{CS} is higher than 0.05 even in the scenario where all the CES rates are symmetric to the bias in the parameter estimation, inherent in maximum likelihood, which means in some cases DAISIE may not perform well in reconstructing clade sizes for the phylogenies with multiple clades. However, it was not our aim here to test the inference errors in DAISIE. We showed that even using trait-independent simulations, the error in σ_{CS} is still higher than expected (Supplementary Figure S11).

We attempted to determine the features of the data simulated under trait dependence that led to large inference errors in the two metrics where the non-negligible error was found (ΔNESRTT, σ_{CS}). The results indicate that clade size variation, which is the difference in species richness between clades, has a larger impact on the model accuracy than tip ratio, which is the species richness difference between states (Supplementary Figures S4, S5, S12, and S13). This means DAISIE may cause error when fitting substantially unbalanced phylogenetic data, no matter if the variation between clades is caused by state-dependent diversification. In a study that used the DAISIE model to fit terrestrial birds of the Galapagos (Valente et al., 2015), the species richness of the clade of Darwin's finches is reported to be much higher than the other clades of birds on the islands. The model that best fits the dataset assumes that Darwin's finches have different cladogenesis and extinction rates than non-Darwin's finches. Spectacular adaptive radiations such as Darwin's finches are well-known on oceanic islands, and obviously lead to large clades. However, except for a handful of classic examples of adaptive radiations (Grant & Grant, 2008; Losos & Ricklefs, 2009; Robichaux et al., 1990; Seehausen, 2006), most island lineages do not diversify to form large clades (Patiño et al., 2017). Therefore, for most islands, clade size variation will rarely be extremely large when the whole assemblage of species of a given taxon is considered, suggesting the performance of models ignoring trait dependence may not be affected for typical islands.

Most studies on island traits and diversity are statistical studies based on empirical data, while phylogenetic-based theoretical methods on island traits are very limited, especially compared with the studies on continents. One reason for this is that insular communities are typically less diverse than continents, and thus their phylogenies are comparatively small and information-poor. Even though the process of adaptive radiation can create some large clades on oceanic islands, radiations tend to be limited to certain clades, and most lineages only diversify to a limited extent (Patiño et al., 2014). Therefore, most island clades and radiations are not amenable to fitting SSE models, which generally require relatively large phylogenetic trees (Davis et al., 2013). While SSE models were not designed for this purpose, it is relatively straightforward to extend them to apply to multiple trees. However, insular communities assemble via colonization and potentially subsequent diversification, and thus focusing only on trees of clades that have radiated (and not on colonization times or singleton lineages that have not diversified) ignores an important part of the processes that form insular communities. Currently, the DAISIE framework is the only tool available to test the trait effects on evolutionary rates in the whole community with multiple phylogenies on islands. We believe that the conclusion in our study does not depend on the details of the DAISIE model, but analyses with other still-to-be-developed models with trait dependence have to be performed to confirm this. Our approach will also be applicable to these models.

Because numerous empirical datasets demonstrate the importance of traits for understanding biodiversity, it is crucial to test the effect of trait absence on the accuracy of the existing island models. The power and accuracy of state-dependent biodiversity models have previously been evaluated for ancestral state reconstruction (Holland et al., 2020) and parameter estimation (Davis et al., 2013). However, in these models the parameter inference model is the same as the simulation model, i.e., the model used to generate phylogenies for estimation. Inference may go awry when this is no longer the case, but this is a more general problem of whether the model accurately describes reality. A statistical approach that does not rely on a formal model for the coupling between states and diversification is available to detect the correlation between trait states and diversification with low type I error (Rabosky & Huang, 2016). However, this method cannot be used to reconstruct the ancestral state or to estimate parameters. In our approach, we apply a robustness analysis that uses the data from the complex model (SDS model) to evaluate the inference power of the simple model (SIS model). In general, the comparison between a complex model and a simple model is to use both models to fit the same datasets, and then compare the estimations using the Akaike information criterion to test which model can describe the data better. However, sometimes complex models are vulnerable to overfitting of the data, and this leads to difficulty in accurately estimating parameters (Kelchner & Thomas, 2006). The pipeline used in this study identifies whether the simple model can accurately reconstruct diversity and phylogenies on islands without considering complex factors. This statistical method not only tests if an additional consideration (trait effects) of interest can affect the power of a simple model but can also reveal under what conditions the additional factor needs to be taken into account. In this way, it constitutes a tool to determine whether it is useful to attempt to find a likelihood for the complex model or develop some other method to estimate parameters for the complex model. While we find that the trait-independent DAISIE inference model seems to be robust to trait dependence, it may still be meaningful to develop trait-dependent inference methods if one is interested in detecting the association between trait states and diversification, or in comparing diversification between mainland and island species with different traits (Patiño et al., 2017).

Supplementary material

Supplementary material is available online at *Evolution* (https://academic.oup.com/evolut/qpad006)

Data availability

Data generated and results obtained in this article are archived on an open data repository hosted by Dryad at https://doi.org/10.5061/dryad.m37pvmd65. The R package "DAISIErobustness" is available on GitHub at https://github.com/Neves-P/DAISIErobustness.

Author contributions

S.X., L.V., and R.S.E. conceived the idea for this article. R.S.E. and L.V. implemented the inference model and original simulation model in the package, and S.X. implemented the additional simulation model. S.X. analyzed the data and wrote the first draft of the manuscript. L.V. and R.S.E. contributed to the revisions.

Conflict of interest: The authors declare no conflict of interest.

Acknowledgments

We would like to thank Pedro S. Neves and Joshua W. Lambert for their valuable comments on the manuscript and we thank Raphael Scherrer and Theo Pannetier for the discussions. We thank the editors and reviewers for their helpful comments on the manuscript. We thank the Center for Information Technology of the University of Groningen for providing access to the Peregrine high-performance computing cluster. S.X. was funded by the China Scholarship Council (CSC). R.S.E. and L.V. were funded by Netherlands Organization for Scientific Research (NWO) through VICI and VIDI grants, respectively.

References

Aleixandre, P., Hernández Montoya, J., & Milá, B. (2013). Speciation on Oceanic Islands: Rapid adaptive divergence vs. cryptic speciation in a Guadalupe Island songbird (Aves: Junco). *PLoS One*, 8, e63242. https://doi.org/10.1371/journal.pone.0063242

Algar, A. C., & Losos, J. B. (2011). Evolutionary assembly of island faunas reverses the classic island-mainland richness difference in Anolis lizards. *Journal of Biogeography*, 38(6), 1125–1137. https://doi.org/10.1111/j.1365-2699.2010.02466.x

Auffret, A. G., Aggemyr, E., Plue, J., & Cousins, S. A. O. (2017). Spatial scale and specialization affect how biogeography and functional traits predict long-term patterns of community turnover. *Functional Ecology*, 31, 436–443.

Beaulieu, J. M., & O'Meara, B. C. (2016). Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Systematic Biology*, 65(4), 583–601. https://doi.org/10.1093/sysbio/syw022

Biddick, M., Hutton, I., & Burns, K. C. (2019). Independent evolution of allometric traits: A test of the allometric constraint hypothesis in island vines. *Biological Journal of the Linnean Society*, 126, 203–211.

Burin, G., Kissling, W. D., Guimarães, P. R., Şekercioglu, C. H., & Quental, T. B. (2016). Omnivory in birds is a macroevolutionary sink. *Nature Communications*, 7, 1–10.

Burns, K. C. (2016). Size changes in island plants: independent trait evolution in *Alyxia ruscifolia* (Apocynaceae) on Lord Howe Island. *Biological Journal of the Linnean Society*, 119(4), 847–855. https://doi.org/10.1111/bij.12851.

Cássia-Silva, C., Freitas, C. G., Lemes, L. P., Paterno, G. B., Dias, P. A., Bacon, C. D., & Collevatti, R. G. (2020). Higher evolutionary rates in life-history traits in insular than in mainland palms. *Scientific Reports*, 10(1), 1–10.

Chevin, L. M. (2016). Species selection and random drift in macroevolution. *Evolution*, 70, 513–525.

Cowie, R. H., & Holland, B. S. 2008. Molecular biogeography and diversification of the endemic terrestrial fauna of the Hawaiian Islands. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1508), 3363–76.

Cracraft, J. (1985). Biological diversification and its causes. *Annals of the Missouri Botanical Garden* 72, 794–822.

Csiszár, I., & Shields, P. C. (2004). Information theory and statistics: A tutorial. *Foundations and Trends® in Communications and Information Theory*, 1, 417–528.

Davis, M. P., Midford, P. E., & Maddison, W. (2013). Exploring power and parameter estimation of the BiSSE method for analyzing species diversification. *BMC Evolutionary Biology*, 13, 38. https://doi.org/10.1186/1471-2148-13-38

Etienne, R. S., Haegeman, B., Dugo-Cota, A., Viì, C., Gonzalez-Voyer, A., & Valente, L. (2022). The phylogenetic limits to diversity-dependent diversification. *Systematic Biology*, https://doi.org/10.1093/SYSBIO/SYAC074

Etienne, R., Valente, L. L., Phillimore, A. B., Haegeman, B., Lambert, J., Neves, P., Xie, S., Bilderbeek, R., Hauffe, T., Laudanno, G., Kristensen, N. & Scherrer, R. (2020). DAISIE: Dynamical Assembly of Islands by Speciation, Immigration and Extinction. https://doi.org/10.5281/ZENODO.4054059.

Fitzjohn, R. G. (2010). Quantitative traits and diversification. *Systematic Biology*, 59(6), 619–633. https://doi.org/10.1093/sysbio/syq053

Fitzjohn, R. G., Maddison, W. P., & Otto, S. P. (2009). Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Systematic Biology*, 58(6), 595–611. https://doi.org/10.1093/sysbio/syp067

Gamisch, A. (2016). Notes on the statistical power of the binary state speciation and extinction (BiSSE) model. *Evolutionary Bioinformatics*, 12, 165–174.

García-Verdugo, C., Baldwin, B. G., Fay, M. F., & Caujapé-Castells, J. (2014). Life history traits and patterns of diversification in oceanic archipelagos: A meta-analysis. *Botanical Journal of the Linnean Society*, 174, 334–348.

García-Verdugo, C., Monroy, P., Pugnaire, F. I., Jura-Morawiec, J., Moreira, X., & Flexas, J. (2020). Leaf functional traits and insular colonization: Subtropical islands as a melting pot of trait diversity in a widespread plant lineage. *Journal of Biogeography*, 47, 2362–2376.

Goldberg, E. E., & Igić, B. (2012). Tempo and mode in plant breeding system evolution. *Evolution*, 66, 3701–3709.

Goldberg, E. E., Lancaster, L. T., & Ree, R. H. (2011). Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Systematic Biology*, 60(4), 451–465. https://doi.org/10.1093/sysbio/syr046

Grant, P. R., & Grant, B. R. (2008). *How and why species multiply: The radiation of Darwin's finches* (p. 218). Princeton University Press.

Herrera-Alsina, L., Van Els, P., & Etienne, R. S. (2019). Detecting the dependence of diversification on multiple traits from phylogenetic trees and trait data. *Systematic Biology*, 68(2), 317–328. https://doi.org/10.1093/sysbio/syy057.

Hetherington-Rauth, M. C., & Johnson, M. T. J. (2020). Floral trait evolution of angiosperms on pacific islands. *American Naturalist*, 196(1), 87–100. https://doi.org/10.1086/709018

Holland, B. R., Ketelaar-Jones, S., O'Mara, A. R., Woodhams, M. D., & Jordan, G. J. (2020). Accuracy of ancestral state reconstruction for non-neutral traits. *Scientific Reports*, 10, 7644.

Jablonski, D. (2008). Species selection: Theory and data. *Annual Reviews of Ecology, Evolution, and Systematics*, 39, 501–524.

Janzen, T., Höhna, S., & Etienne, R. S. (2015). Approximate Bayesian Computation of diversification rates from molecular phylogenies: Introducing a new efficient summary statistic, the nLTT. *Methods in Ecology and Evolution*, 6(5), 566–575. https://doi.org/10.1111/2041-210x.12350.

Kelchner, S. A., & Thomas, M. A. (2006). Model use in phylogenetics: Nine key questions. *Trends in Ecology and Evolution*, 22(2), 87–94.

Kisel, Y., Mcinnes, L., Toomey, N. H., & Orme, C. D. L. (2011). How diversification rates and diversity limits combine to create large-scale species-area relationships. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366, 2514–2525.

Kozak, K. H., & Wiens, J. J. (2016). Testing the relationships between diversification, species richness, and trait evolution. *Systematic Biology*, 65(6), 975–988. https://doi.org/10.1093/sysbio/syw029.

Lambert, J. W., Neves, P., & Xie, S. (2022). *DAISIErobustness: Test the robustness of DAISIE to geodynamics and traits*. https://doi.org/10.5281/ZENODO.6811591.

Leidinger, L., & Cabral, J. S. (2017). Biodiversity dynamics on Islands: Explicitly accounting for causality in mechanistic models. *Diversity*, 9(3), 30.

Li, P., & Wiens, J. J. (2022). What drives diversification? Range expansion tops climate, life history, habitat and size in lizards and snakes. *Journal of Biogeography*, 49(2), 237–247. https://doi.org/10.1111/jbi.14304.

Losos, J. B., & Ricklefs, R. E. (2009). Adaptation and diversification on islands. *Nature*, 457(7231), 830–836. https://doi.org/10.1038/nature07893

Losos, J. B., Warheit, K. I., & Schoener, T. W. (1997). Adaptive differentiation following experimental island colonization in Anolis lizards. *Nature*, 387, 70–73.

Maddison, W. P., Midford, P. E., & Otto, S. P. (2007). Estimating a binary character's effect on speciation and extinction. *Systematic Biology*, 56(5), 701–710. https://doi.org/10.1080/10635150701607033

McPeek, M. A., & Brown, J. M. (2007). Clade age and not diversification rate explains species richness among animal taxa. *American Naturalist*, 169, 97–106.

Millien, V. (2006). Morphological evolution is accelerated among island mammals. *PLoS Biology*, 4, e3841863–e3841868. https://doi.org/10.1371/journal.pbio.0040384. Public Library of Science

Mitter, C., Farrell, B., & Wiegmann, B. (1988). The phylogenetic study of adaptive zones: Has phytophagy promoted insect diversification? *American Naturalist*, 132(1), 107–128. https://doi.org/10.1086/284840

Onstein, R. E., Baker, W. J., Couvreur, T. L. P., Faurby, S., Svenning, J. C., & Kissling, W. D. (2017). Frugivory-related traits promote speciation of tropical palms. *Nature Ecology and Evolution*, 1(12), 1903–1911. https://doi.org/10.1038/s41559-017-0348-7

Parent, C. E., Caccone, A. & Petren, K. (2008). Colonization and diversification of Galápagos terrestrial fauna: A phylogenetic and biogeographical synthesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1508), 3347–3361.

Patiño, J., Carine, M., Fernández-Palacios, J. M., Otto, R., Schaefer, H., & Vanderpoorten, A. (2014). The anagenetic world of spore-producing land plants. *New Phytologist*, 201(1), 305–311. https://doi.org/10.1111/nph.12480

Patiño, J., Whittaker, R. J., Borges, P. A. V., Fernández-Palacios, J. M.J. M., Ah-Peng, C., Araújo, M. B.M. B., ÁvilaS. P.S. P., Cardoso, P., Cornuault, J., de Boer, E. J., Nascimento, L. de, Gil, A., González-Castro, A.A., Gruner, D. S., Heleno, R., Hortal, J., Illera, J. C., Kaiser-Bunbury, C. N., Matthews, T. J., … Emerson, B. C. (2017). A roadmap for island biology: 50 fundamental questions after 50 years of The Theory of Island Biogeography. *Journal of Biogeography*, 44, 963–983.

Rabosky, D. L., & Goldberg, E. E. (2015). Model inadequacy and mistaken inferences of trait-dependent speciation. *Systematic Biology*, 64(2), 340–355. https://doi.org/10.1093/sysbio/syu131

Rabosky, D. L., & Huang, H. (2016). A robust semi-parametric test for detecting trait-dependent diversification. *Systematic Biology*, 65(2), 181–193. https://doi.org/10.1093/sysbio/syv066

Rabosky, D. L., & McCune, A. R. (2010). Reinventing species selection with molecular phylogenies. *Trends in Ecology and Evolution*, 25(2), 68–74. https://doi.org/10.1016/j.tree.2009.07.002

Robichaux, R. H., Carr, G. D., Liebman, M., & Pearcy, R. W. (1990). Adaptive radiation of the Hawaiian Silversword Alliance (Compositae-Madiinae): Ecological, morphological, and physiological diversity. *Annals of the Missouri Botanical Garden*, 77, 64.

Santos Neves, P., Lambert, J. W., Valente, L., & Etienne, R. S. (2022). The robustness of a simple dynamic model of island biodiversity to geological and sea-level change. *Journal of Biogeography*, 49(11), 2091–2104. https://doi.org/10.1111/jbi.14519

Scholl, J. P., & Wiens, J. J. (2016). Diversification rates and species richness across the Tree of Life. *Proceedings of the Royal Society B: Biological Sciences*, 283(1838), 20161334.

Seehausen, O. (2006). African cichlid fish: A model system in adaptive radiation research. *Proceedings of the Royal Society B: Biological Sciences*, 273, 1987–1998.

Simpson, C. (2013). Species selection and the macroevolution of coral coloniality and photosymbiosis. *Evolution*, 67, 1607–1621.

Stanley, S. M. (1975). A theory of evolution above the species level. *Proceedings of the National Academy of Sciences*, 72(2), 646–650. https://doi.org/10.1073/pnas.72.2.646

Supsup, C. E., Asis, A. A., Carestia, U. V., Diesmos, A. C., Mallari, N. A. D., & Brown, R. M. (2020). Variation in species richness, composition and herpetological community structure across a tropical habitat gradient of Palawan Island, Philippines. *Herpetozoa*, 33, 95–111.

Tietje, M., Antonelli, A., Baker, W. J., Govaerts, R., Smith, S. A., & Eiserhardt, W. L. (2022). Global variation in diversification rate and species richness are unlinked in plants. *Proceedings of the National Academy of Sciences of the United States of America*, 119, e2120662119.

Valente, L. M., Phillimore, A. B., & Etienne, R. S. (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the

Galápagos Islands. *Ecology Letters*, 18(8), 844–852. https://doi.org/10.1111/ele.12461

Valente, L., Phillimore, A. B., & Etienne, R. S. (2018). Using molecular phylogenies in island biogeography: It's about time. *Ecography*, 41, 1684–1686.

Valente, L., Phillimore, A. B., Melo, M., Warren, B. H., Clegg, S. M., Havenstein, K., Tiedemann, R., Illera, J. C., Thébaud, C., Aschenbach, T., & Etienne, R. S. (2020). A simple dynamic model explains the diversity of island birds worldwide. *Nature*, 579(7797), 92–96. https://doi.org/10.1038/s41586-020-2022-5

Zheng, J. M., & Zheng. 2022. Exploring the species and phylogenetic diversity, phylogenetic structure of mixed communities along the coastal gradient. A case study in a subtropical Island, China. *Applied Ecology and Environmental Research*, 20(4), 3129–3141 https://doi.org/10.15666/aeer/2004_31293141.

Zhu, X. -M., Du, Y., Qu, Y. -F., Li, H., Gao, J. -F., Lin, C. -X., Ji, X., & Lin, L. -H. (2020). The geographical diversification in varanid lizards: The role of mainland versus island in driving species evolution. *Current Zoology*, 66(2), 165–171. https://doi.org/10.1093/cz/zoaa002