

Notes on the Statistical Power of the Binary State Speciation and Extinction (BiSSE) Model

Alexander Gamisch

Department of Ecology and Evolution, University of Salzburg, Salzburg, Austria.

ABSTRACT: The Binary State Speciation and Extinction (BiSSE) method is one of the most popular tools for investigating the rates of diversification and character evolution. Yet, based on previous simulation studies, it is commonly held that the BiSSE method requires phylogenetic trees of fairly large sample sizes (>300 taxa) in order to distinguish between the different models of speciation, extinction, or transition rate asymmetry. Here, the power of the BiSSE method is reevaluated by simulating trees of both small and large sample sizes (30, 60, 90, and 300 taxa) under various asymmetry models and root state assumptions. Results show that the power of the BiSSE method can be much higher, also in trees of small sample size, for detecting differences in speciation rate asymmetry than anticipated earlier. This, however, is not a consequence of any conceptual or mathematical flaw in the method per se but rather of assumptions about the character state at the root of the simulated trees and thus the underlying macroevolutionary model, which led to biased results and conclusions in earlier power assessments. As such, these earlier simulation studies used to determine the power of BiSSE were not incorrect but biased, leading to an overestimation of type-II statistical error for detecting differences in speciation rate but not for extinction and transition rates.

KEYWORDS: BiSSE, simulation, low sample size, type-II statistical error, key innovation model

CITATION: Gamisch. Notes on the Statistical Power of the Binary State Speciation and Extinction (BiSSE) Model. *Evolutionary Bioinformatics* 2016;12:165–174. doi: 10.4137/EBO.S39732.

TYPE: Original Research

RECEIVED: March 22, 2016. **RESUBMITTED:** June 09, 2016. **ACCEPTED FOR PUBLICATION:** June 10, 2016.

ACADEMIC EDITOR: Liuyang Wang, Associate Editor

PEER REVIEW: Six peer reviewers contributed to the peer review report. Reviewers' reports totaled 2293 words, excluding any confidential comments to the academic editor.

FUNDING: The present study was funded by the FWF (Austrian Science Fund) grant P17124-B0 to Hans Peter Comes (University of Salzburg) and by the Stiftungs- und Förderungsgesellschaft der Paris-Lodron-Universität Salzburg. The author confirms that the funder had no influence over the study design, content of the article, or selection of this journal.

COMPETING INTERESTS: Author discloses no potential conflicts of interest.

CORRESPONDENCE: alexandergamisch@gmx.at

COPYRIGHT: © the authors, publisher and licensee Libertas Academica Limited. This is an open-access article distributed under the terms of the Creative Commons CC-BY-NC 3.0 License.

Paper subject to independent expert blind peer review. All editorial decisions made by independent academic editor. Upon submission manuscript was subject to anti-plagiarism scanning. Prior to publication all authors have given signed confirmation of agreement to article publication and compliance with all applicable ethical and legal requirements, including the accuracy of author and contributor information, disclosure of competing interests and funding sources, compliance with ethical requirements relating to human and animal study participants, and compliance with any copyright requirements of third parties. This journal is a member of the Committee on Publication Ethics (COPE).

Published by Libertas Academica. Learn more about this journal.

Introduction

How species traits influence rates of diversification is of central interest in evolutionary biology.^{1–3} The Binary State Speciation and Extinction (BiSSE) model¹ allows the joint estimation of state-dependent but time-invariant rates of binary character change (q_{01} , q_{10}) as well as state-dependent rates of extinction (μ_0 , μ_1) and speciation (λ_0 , λ_1). Together with its derivations (eg, QuaSSE,⁴ GeoSSE,⁵ ClaSSE,⁶ MuSSE,⁷ BiSSE-BMA,⁸ HiSSE⁹), the BiSSE method has become one of the most popular tools for the detection of asymmetries in diversification rates from phylogenies. However, till date, there has been little work on the statistical power of BiSSE.¹⁰ In fact, only three studies have so far explored the potential for type-I (incorrect rejection of a true null hypothesis) and type-II (failure to reject a false null hypothesis) statistical errors associated with this method.^{1,10,11} While the potential for type-I errors can be assessed by using simulated (“neutral”) trait datasets,¹¹ and potentially be counteracted using the HiSSE framework,^{9,11} there has been little guidance in interpreting the sources of type-II errors and thus the ability of BiSSE to distinguish among macroevolutionary hypotheses of interest.^{1,10}

According to previous accounts,^{1,10} BiSSE suffers from two general limitations that complicate the detection of macroevolutionary processes underlying observed heterogeneities of state frequencies in a given phylogeny, namely, (i) the

requirement for fairly large phylogenetic trees (with c. 300–500 extant species) and (ii) a low tip ratio bias (one character state should be not more frequent among the tips than a ratio of 10:1). The reason for this could be that many similar ways could have generated an observed phylogeny and associated character states, which only larger datasets are able to disentangle.^{1,10} In other words, if tree size is too small, or when one of the binary states in question is exceedingly rare, confounding effects may lead to the acceptance of a (false) null model of diversification rate symmetry; thus, the potential for a type-II error is extremely high.^{1,10} Based on these earlier findings, it is now the general perception that the BiSSE method performs poorly with low sample sizes and/or high tip ratio bias,^{1,10} which by itself is often a reason for causing concern (eg, see Igić and Busch¹²; Soltis et al.¹³ for a critical reappraisal of Mayrose et al.¹⁴).

Maddison et al.¹ and Davis et al.¹⁰ have previously assessed the type-II error of BiSSE using simulations. However, it remains unclear whether these earlier power explorations included the full range of macroevolutionary hypotheses. The latter are essentially defined by asymmetries in one of the diversification parameters (speciation or transition or extinction rate) of the derived state compared to the ancestral one (Table 1). For example, if state 1 confers a higher speciation rate than state 0 ($\lambda_0 < \lambda_1$), then a tree with the pattern of a *key innovation* is simulated if the ancestral state is 0; or a tree



with a *lonely state* pattern is simulated, if the ancestral state is 1 (see also Table 1). This connotation has been ignored in the test setups of Maddison et al.¹ and Davis et al.¹⁰, probably because the option to use different root priors for joint tree–trait simulations (eg, to fix state 0 or 1 at the root) has become available (for the DIVERSE package of MESQUITE) only after the initial description of the BiSSE method.^{1,7,24,25} Instead, Maddison et al.¹ and Davis et al.¹⁰ used simulated trees with the *stationary frequency* (sometimes termed *equilibrium frequency*) root state prior, which randomly chooses a state of the diversification model's equilibrium distribution. Under this setting, a state whose diversification parameter increases the frequency of that state among the tips (eg, $> \lambda$ or $< \mu$ or $> q$; all else being equal) is more likely to be selected as the state at the root of that tree.^{1,2,25,26} It is thus unclear whether and how the potentially circular logic of the stationary frequency root state assumption affected the conclusions of Maddison et al.¹ and Davis et al.¹⁰ in their exploration of the statistical power of the BiSSE method. It is known, however, that the root state assumption can have a dramatic effect on phylogenetic tests of irreversibility, eg, by frequently leading to an incorrect rejection of Dollo's law.²⁵

The present simulation study investigates the power of BiSSE to detect rate asymmetries in diversification parameters in trees simulated under a variety of macroevolutionary models. For this purpose, trees were simulated under the three alternative root state assumptions and variable tree sizes (300, 90, 60, and 30 extant species). The present findings suggest that the root state assumption has a dramatic effect on the trees simulated. This will be discussed in the context of what is known

about the power of the BiSSE method, especially with regard to small sample size trees and the effect of tip ratio biases.

Methods

BiSSE performance evaluation using simulated data.

In general, BiSSE computes the probability of a phylogenetic tree and the observed state distribution of a binary trait (0, 1) among extant taxa, given a particular model of character evolution, speciation, and extinction.^{1,2} BiSSE assumes that the phylogenetic tree is fully resolved and complete (but see Fitzjohn et al.² for methodology dealing with incomplete or unresolved phylogenies) and that the diversification parameters are constant and operate independently throughout the tree.¹

The power of BiSSE hypothesis testing was explored using simulated trees of 30, 60, and 90 terminal taxa, and for direct comparison with the study of Davis et al.¹⁰, 300 extant terminal taxa were used under various simple models of rate asymmetry. The R package DIVERSITREE version 0.9–2⁷ was used for joint tree–trait simulations and subsequent BiSSE likelihood calculations and hypothesis tests. The *tree.bisse* function was used to simultaneously simulate binary characters that affect speciation and/or extinction, and the tree history itself. For each of the states 0 and 1, rates of speciation (λ_0, λ_1), extinction (μ_0, μ_1), and transition (q_{01}, q_{10}) as well as the character state at the root were specified (all R functions used are provided as Supplementary File 1). The state at the root was fixed to either 0 or 1 or set to stationary frequency (NA in DIVERSITREE). In the latter case, the root is weighted according to the expected equilibrium of frequencies of taxa with

Table 1. Simple alternative macroevolutionary hypotheses that can be tested using BiSSE. These macroevolutionary hypotheses are essentially defined by asymmetries in one of the diversification parameters (speciation [$\lambda_0 < \lambda_1$] or transition [$q_{01} > q_{10}$] or extinction [$\mu_0 > \mu_1$] rate) of the derived state compared to the ancestral one (with all other parameters being symmetrical).

ANCESTRAL STATE	RATE ASYMMETRY	PROCESS	MACROEVOLUTIONARY MODEL	EXAMPLE STUDIES
0	$\lambda_0 < \lambda_1$	Higher net diversification rate of the derived state due to higher speciation rate ^{3,8}	<i>Key innovation</i>	Herbivory in mammals; ¹⁵ CAM photosynthesis in Bromeliaceae; ⁸ heterostyly in primroses ¹⁶
1		Reduced net diversification rate of the derived state due to lower speciation rate ¹⁷	<i>Lonely state</i>	Moth pollination in <i>Ruella</i> ; ¹⁸ unpigmented flowers in <i>Ipomoea</i> subgenus <i>Quamoclit</i> ; ¹⁹ omnivory in mammals ¹⁵
0	$\mu_0 > \mu_1$	Higher net diversification rate of the derived state due to lower extinction rate ^{3,8}	<i>Key innovation</i>	Tank habit of Bromeliaceae ⁸
1		Negative net diversification rate of the derived trait (usually) due to higher extinction rate ^{3,12,17}	<i>Dead end</i>	Loss of self-incompatibility in Solanaceae; ²⁰ polyploidy in plants; ¹⁴ selfing in Madagascan <i>Bulbophyllum</i> ²¹
0	$q_{01} > q_{10}$	Higher transition from the ancestral state to the derived state ³	<i>Asymmetrical transitions</i>	Selfing in Madagascan <i>Bulbophyllum</i> ; ²¹ shifts from lowland to highland in Madagascan <i>Bulbophyllum</i> ²²
1		Lower transition from the ancestral state to the derived state ³	<i>Asymmetrical transitions</i>	Sociality in spiders ²³

state 0 and 1, as defined by the diversification parameters (λ_0 , λ_1 , μ_0 , μ_1 , q_{01} , q_{10}).^{1,2,10,25,26,27}

Using the same diversification parameter settings as Maddison et al.¹ and Davis et al.¹⁰, 500 trees and associated binary characters, each with 300, 90, 60, and 30 terminal taxa, were simulated for each root state under two scenarios of rate asymmetry. Under the first scenario, one parameter (in the following marked in bold) has twice the value than the other one with all other parameters being equal (see Maddison et al.¹ and the 2x models of Davis et al.¹⁰), that is, asymmetrical rates of speciation (2x): $\lambda_0 = 0.1$, $\lambda_1 = \mathbf{0.2}$, $\mu_0 = \mu_1 = 0.03$, $q_{01} = q_{10} = 0.01$; asymmetrical rates of extinction (2x): $\lambda_0 = \lambda_1 = 0.1$, $\mu_0 = \mathbf{0.06}$, $\mu_1 = 0.03$, $q_{01} = q_{10} = 0.01$; and asymmetrical rates of transitions (2x): $\lambda_0 = \lambda_1 = 0.1$, $\mu_0 = \mu_1 = 0.03$, $q_{01} = \mathbf{0.01}$, $q_{10} = 0.005$. Under the second scenario of rate asymmetry, one parameter (in the following marked in bold) has ten times the value (compare to the 10x models of Davis et al.¹⁰) with all other parameters being equal, that is, asymmetrical rates of speciation (10x): $\lambda_0 = 0.1$, $\lambda_1 = \mathbf{1.0}$, $\mu_0 = \mu_1 = 0.03$, $q_{01} = q_{10} = 0.01$; asymmetrical rates of extinction (10x): $\lambda_0 = \lambda_1 = 0.1$, $\mu_0 = \mathbf{0.03}$, $\mu_1 = 0.003$, $q_{01} = q_{10} = 0.01$; asymmetrical rates of transitions (10x): $\lambda_0 = \lambda_1 = 0.1$, $\mu_0 = \mu_1 = 0.03$, $q_{01} = \mathbf{0.01}$, $q_{10} = 0.001$.

The influence of the root state assumption on the simulations was assessed by fixing the root to either state 0 or 1 or stationary frequency (see above). For each simulation, the average percentage of (i) the root state with state 0 and (ii) the terminal taxa with state 0 was recorded from 500 simulated trees, respectively. The observed root state can differ from the fixed one if the early branching taxa go extinct after a transition to the new state has occurred (as the tree only contains extant taxa, the next descendant node with the new state is then perceived as the root of the tree). For the purpose of this study, the state at the root of the tree is referred to as the ancestral state. Representative tree histories reflecting the average observed frequency of state 0 at the root and tips (see Tables 2–4) are depicted in Figure 1.

All simulated trees were examined to explore the power of the BiSSE method to reject a (false) null model of diversification rate symmetry, that is, either $\lambda_0 = \lambda_1$ or $\mu_0 = \mu_1$ or $q_{01} = q_{10}$ when in fact the trees were simulated under an asymmetric model, that is, $\lambda_0 < \lambda_1$ or $\mu_0 > \mu_1$ or $q_{01} > q_{10}$.¹ To do so, the fit of the full (six-parameter) BiSSE model with unconstrained parameters (ie, λ , μ , and q are allowed to vary) was compared to three models with those parameters constrained to be equal. All models were fitted by maximum likelihood (ML) optimization across all 500 trees for each of the tree size and diversification parameter combinations. For testing hypotheses of speciation, extinction, or transition rate asymmetry likelihood ratio tests (LRTs) were used.¹ However, for speciation rate asymmetry, the power levels of hypotheses significantly differed from those reported in the literature^{1,10} (see Results). Hence, in this case, the power was additionally tested and confirmed using empirically determined critical (5% cutoff) values (see Table 2). The cutoff values were

determined by simulating a set of 500 trees with symmetrical rate parameters ($\lambda_0 = \lambda_1$, $\mu_0 = \mu_1$, $q_{01} = q_{10}$) and by using these trees to test the fit of the full model (more complex) and the constrained (symmetrical; eg, $\lambda_0 = \lambda_1$) H_0 model.^{1,10} The lowest value of the upper 5% of the 2x log-likelihood differences between the full and constrained model was used as the cutoff value.^{1,10} Power was estimated as the percentage of simulations that had a higher likelihood difference than the empirically determined 5% cutoff value.^{1,10}

Likewise, LRTs were calculated as twice the difference in log-likelihoods between the full (more complex) model and the constrained (simpler) model (H_0), which should follow a χ^2 distribution with degrees of freedom equal to the difference in the number of free parameters allowed by the two models.¹ Power was estimated as the percentage of trees, where the full model provided a significant improvement ($P < 0.05$) over the simpler constrained model (H_0).

Results

Influence of the root state assumption on observed root/tip state frequencies and simulated tree histories.

The results of the three types of asymmetrical rate simulations (speciation, extinction, and transition) are shown in Tables 2–4. As a general outcome, the frequency enhancing diversification parameters associated with state 1 (higher speciation, transition, or lower extinction rate compared to state 0) had, depending on the root state assumption (the state from which the simulations started), a dramatic effect on the simulated tree histories and tip frequencies (Tables 2–4). In detail, in the asymmetrical speciation rate simulations, the observed frequencies of state 0 at the root and tips ranged between 85–88.2% (root) and 25.6–7.6% (tips) under the 2x scenario and between 86.4–88.6% (root) and 4.5–28.4% (tips) under the 10x scenario if the root was fixed to state 0 (Table 2). In contrast, the observed frequency of state 0 at the root and tips ranged between 9.8–13% (root) and 10–13.1% (tips) under the 2x scenario and between 0.4–1.8% (root) and 1.1–1.6% (tips) under the 10x scenario if the root was set to stationary frequency (NA) and between 3.4–5% (root) and 8.1–8.6% (tips) under the 2x scenario and between 0.6–1.2% (root) and 1.1–1.2% (tips) under the 10x scenario if the root was fixed to state 1 (Table 2). In the asymmetrical extinction rate simulations (Table 3), with the root fixed to state 0, the observed frequencies of state 0 ranged between 73.8–89.2% (root) and 34.7–59.6% (tips) over both scenarios; when the root was set to stationary frequency or to state 1, the observed frequencies of state 0 ranged between 17.8–23.8% (root) and 23.3–28.7% (tips) and between 6–8.8% (root) and 18.5–23.8% (tips) over both scenarios, respectively. Finally, in the asymmetrical transition rate simulations (Table 4), with the root fixed to state 0, the observed frequencies of state 0 ranged between 88.2–89.6% (root) and 44.2–69.5% (tips) over both scenarios; when the root was set to stationary frequency or state 1, the corresponding values were, respectively, 31.2–34.4% (root)



Table 2. Power of asymmetrical speciation rate simulations. The remaining parameters were symmetrical for each simulation (2x: $\lambda_0 = 0.1$, $\lambda_1 = 0.2$, $\mu_0 = \mu_1 = 0.03$, $q_{01} = q_{10} = 0.01$; 10x: $\lambda_0 = 0.1$, $\lambda_1 = 1.0$, $\mu_0 = \mu_1 = 0.03$, $q_{01} = q_{10} = 0.01$). The trees were simulated with the root set to either state 0 or 1 or stationary frequency (NA) and used to test the null hypothesis (H_0) of rate symmetry ($\lambda_0 = \lambda_1$) in BiSSE. The observed percentage of root and terminal taxa with state 0 is indicated by the mean value from 500 simulations. The corresponding macroevolutionary pattern is indicated according to Table 1. See Figure 1A–C for a representative set of simulated trees (60 taxa) under different root state assumptions.

NUMBER OF TAXA	ROOT STATE ASSUMPTION	ASYMMETRY	% OBSERVED STATE 0 (ROOT)	% OBSERVED STATE 0 (TIPS)	MACROEVOLUTIONARY PATTERN	5% CUTOFF VALUE	POWER% (CUTOFF)*	POWER (LRT% REJECTING H_0)
300	root = 0	2x	88.2	25.6	<i>Key innovation</i>	3.4	75.2	72
90		2x	85	39.5		4.1	39.8	42.6
60		2x	87	45.5		3.3	38.8	33.8
30		2x	88.2	57.6		3.3	18	14.8
300	10x	10x	87.2	4.5		3.4	88.4	86.4
90		10x	86.4	12.6		4.1	82.6	83.2
60		10x	87	18.4		3.3	85.2	82.2
30		10x	88.6	28.4		3.3	77.2	72.2
300	root = NA	2x	13	10	<i>Lonely state</i>	3.5	49.4	44.6
90		2x	10	10.8		3.5	22	18.4
60		2x	12	12.1		3.5	9.4	8.2
30		2x	9.8	13.1		3.2	4.6	7.6
300	10x	10x	0.4	1.1		3.5	32.8	29.6
90		10x	1.8	1.2		3.5	17	13.2
60		10x	1.6	1.2		3.5	8	6
30		10x	1	1.6		3.2	7.8	5
300	root = 1	2x	4	8.6	<i>Lonely state</i>	3.6	46.6	43.8
90		2x	4.6	8.5		3.3	20.6	15.8
60		2x	5	8.1		3.3	14.2	11.2
30		2x	3.4	8.2		3.2	6.6	3.2
300	10x	10x	0.6	1.1		3.6	30.4	28
90		10x	0.6	1.1		3.3	11.6	6.8
60		10x	1.2	1.2		3.3	9.4	6.2
30		10x	0.6	1.1		3.2	4.6	2.8

Note: *Percentages of trees with a greater log likelihood difference than the empirically determined critical (5% cutoff) values leading to the rejection of the (false) H_0 (see also Maddison et al.¹; Davis et al.¹⁰).

and 30.8–34.2% (tips) under the 2x scenario and between 8.6–11.2% (root) and 8–10.1% (tips) under the 10x scenario, and 5–7.6% (root) and 15.5–26.2% (tips) under the 2x scenario and between 1.2–1.6% (root) and 3.8–5.5% (tips) under the 10x scenario. Thus, as a general pattern across all simulations, the observed frequency of state 0 at the root and tips was relatively high when the root was fixed to state 0 but much lower when the root was fixed to state 1 or stationary frequency (Tables 2–4). Furthermore, trees simulated under stationary frequency tended to mimic trees that were simulated with the root fixed to state 1 in terms of dominance of state 1 at the root and tips (see also Fig. 1B, C, E, F, H, I). This is simply because the root state is pushed toward the state with higher speciation and transitions rates (here state 1) and lower extinction rates (here state 1), if the root is set to stationary frequency.

Macroevolutionary models are essentially defined by diversification rate asymmetries of the derived state compared to the ancestral one (see Table 1). Not unexpectedly, therefore, the root state assumption was found to have a determining effect on the macroevolutionary tree histories simulated. In detail, if the root was set to stationary frequency or state 1, trees were simulated following the *lonely state* model (for asymmetrical speciation), or the *dead end* model (for asymmetrical extinction) or a model of *asymmetrical transitions* in terms of higher transitions from the derived trait to the ancestral one. In contrast, if the root was set to state 0, trees were simulated following (i) the *key innovation* model due to higher speciation of the derived trait than the ancestral one (for asymmetrical speciation), or (ii) the *key innovation* model due to lower extinction rate of the derived trait (for asymmetrical extinction) than the ancestral one, or (iii) a model of *asymmetrical*

Table 3. Power of asymmetrical extinction rate simulations. The remaining parameters were symmetrical for each simulation (2x: $\lambda_0 = \lambda_1 = 0.1$, $\mu_0 = 0.06$, $\mu_1 = 0.03$, $q_{01} = q_{10} = 0.01$; 10x: $\lambda_0 = \lambda_1 = 0.1$, $\mu_0 = \mathbf{0.03}$, $\mu_1 = 0.003$, $q_{01} = q_{10} = 0.01$). The trees were simulated with the root set to either state 0 or 1 or stationary frequency (NA) and used to test the null hypothesis (H_0) of rate symmetry ($\mu_0 = \mu_1$) in BiSSE. The observed percentage of root and terminal taxa with state 0 is indicated by the mean value from 500 simulations. The corresponding macroevolutionary pattern is indicated according to Table 1. See Figure 1D–F for a representative set of simulated trees (60 taxa) under different root state assumptions.

NUMBER OF TAXA	ROOT STATE ASSUMPTION	ASYMMETRY	% OBSERVED STATE 0 (ROOT)	% OBSERVED STATE 0 (TIPS)	MACROEVOLUTIONARY PATTERN	POWER (LRT% REJECTING H_0)
300	root = 0	2x	77	34.7	<i>Key innovation</i>	6.4
90		2x	78.2	43.7		2.6
60		2x	73.8	47.8		2
30		2x	77	57.8		1
300		10x	85.8	43		3.8
90		10x	89.2	52.1		1.4
60		10x	84.8	52.1		0.6
30		10x	84.4	59.6		1
300	root = NA	2x	17.8	24	<i>Dead end</i>	8.8
90		2x	19.6	23.3		2
60		2x	17.4	26.2		0.6
30		2x	19.8	26.1		0.8
300		10x	23.2	27.4		4.2
90		10x	22.6	28.7		1.6
60		10x	23.8	28.3		0.2
30		10x	22.4	27.4		0.6
300	root = 1	2x	8.8	22	<i>Dead end</i>	7.6
90		2x	7.2	20.6		2.2
60		2x	7	19.9		1.6
30		2x	6.6	18.5		0.8
300		10x	6.4	23.8		4.4
90		10x	7.8	23.3		1
60		10x	7.8	22.6		0.6
30		10x	6	20.7		0.6

transitions in terms of higher transitions from the ancestral trait to the derived one. These results were robust to different sample sizes (300, 90, 60, and 30 terminal taxa) and to different degrees of rate asymmetry (2x and 10x), indicating that the stationary root state assumption is not the most appropriate root prior for ensuring a balanced set of simulated tree histories.

Power of the BiSSE method: speciation, extinction, and transition rates. The present simulations further indicate that the power of the BiSSE method depends on the tree size, with the lowest power in small trees and highest power in large trees (Tables 2–4). However, the power of the method was found to differ depending on the root state assumption under which the trees were simulated (Tables 2–4). In this respect, the power to detect speciation rate asymmetries was highest in trees simulated with the root set to 0 (ie, trees with the derived state had the higher speciation rate), while the power was relatively low in trees simulated with the root set to

either 1 or stationary frequency (ie, trees in which the ancestral state had the higher speciation rate). Further, the degree of rate asymmetry apparently modulates the signal for BiSSE to detect the speciation rate – and to a lesser extent the transition rate – but not extinction rate asymmetries (Tables 2–4). That said, under the scenarios of high rate asymmetry (10x), the power strongly increased for detecting differences in the speciation rate compared to the 2x scenarios when the root is set to 0 but decreased when the root is set to stationary frequency or state 1 (see Table 2). Also, the power to detect transition rate asymmetries increased compared to the 2x scenario, while the power to detect extinction rate asymmetries remained relatively low under the scenario of high rate asymmetry (10x; see Tables 3 and 4).

In detail, the statistical power for trees simulated with the root set to 0 ranged between 72 and 75.2% for trees with 300 tips, 33.8 and 42.6% for trees with 60 and 90 tips, and between 14.8 and 18% for trees with 30 tips under the 2x scenario and



Table 4. Power of asymmetrical transition rate simulations. The remaining parameters were symmetrical for each simulation (2x: $\lambda_0 = \lambda_1 = 0.1$, $\mu_0 = \mu_1 = 0.03$, $q_{01} = 0.01$, $q_{10} = 0.005$; 10x: $\lambda_0 = \lambda_1 = 0.1$, $\mu_0 = \mu_1 = 0.03$, $q_{01} = 0.01$, $q_{10} = 0.001$). The trees were simulated with the root set to either state 0 or 1 or stationary frequency (NA) and used to test the null hypothesis (H_0) of rate symmetry ($q_{01} = q_{10}$) in BiSSE. The observed percentage of root and terminal taxa with state 0 is indicated by the mean value from 500 simulations. The corresponding macroevolutionary pattern is indicated according to Table 1. See Figure 1G–I for a representative set of simulated trees (60 taxa) under different root state assumptions.

NUMBER OF TAXA	ROOT STATE ASSUMPTION	ASYMMETRY	% OBSERVED STATE 0 (ROOT)	% OBSERVED STATE 0 (TIPS)	MACROEVOLUTIONARY PATTERN	POWER (LRT% REJECTING H_0)
300	root = 0	2x	88.2	57.4	<i>Asymmetrical transitions</i>	17.4
90		2x	87.8	62.9		6.2
60		2x	86.2	65.3		4.2
30		2x	89.6	69.5		2.4
300	root = NA	10x	86.2	44.2	<i>Asymmetrical transitions</i>	63.2
90		10x	89.6	53.3		19.8
60		10x	87.8	54.6		10.4
30		10x	86.8	59.8		3.4
300		2x	31.2	32.5		14.8
90		2x	34.4	34.2		6.8
60		2x	32	30.8		5.6
30	2x	32.2	34	2.2		
300	root = 1	10x	9	8.9	<i>Asymmetrical transitions</i>	22.2
90		10x	10	9.6		7.2
60		10x	8.6	8		6
30		10x	11.2	10.1		3
300		2x	5	22		13.4
90		2x	7.6	17.4		4.2
60		2x	5	26.2		4.6
30		2x	7	15.5		2.8
300		10x	1.2	5.5		17
90		10x	1.6	5.2		8.2
60	10x	1.2	4.2	4		
30	10x	1.6	3.8	2		

between 86.4 and 88.4% for trees with 300 tips, 82.2 and 88.4% for trees with 60 and 90 tips, and between 72.2 and 77.2% for trees with 30 tips under the 10x scenario (Table 2). Overall, these results demonstrate an approximately 1.6 to 4.6 times higher power under the 2x scenario, and an even 2.9 to 25.8 times higher power under the 10x scenario in trees simulated with the root set to 0 than in trees simulated with the root set to either 1 or stationary frequency (Table 2). This effect was most pronounced in trees of small sample size (30 and 60 taxa) with an approximately 2 to 4.6, and even 13.7 to 25.8 times increase in power under the 2x and 10x scenarios, respectively (Table 2). In contrast, when using trees simulated with the root set to 1 or stationary frequency, the power was much lower and ranged between 43.8–49.4% (2x) and 28–32.8% (10x) for trees of 300 taxa, between 15.8–22% (2x) and 6.8–17% (10x) for trees of 90 taxa and between 3.2–9.4% (2x) and 2.8–9.4% (10x) for trees of both 30 and 60 taxa.

On the other hand, the power of BiSSE for testing hypotheses of extinction and transition rates was mostly unaffected by the root state assumption (but see below) and much lower than that observed for speciation rate asymmetry, but also dependent on tree size ($300 > 90 > 60 > 30$; Tables 3, 4). The only exception is found in the *asymmetrical transitions* models under the 10x scenarios in trees of 300 taxa. Here, BiSSE showed a 2.8 to 3.7 times increase in power in trees simulated with the root set to 0 compared to trees simulated with the root set to either 1 or stationary frequency (Table 4). That said, the power to reject the false hypothesis of transition rate symmetry ranged from 2.4 and 17.4% (root set to 0) under the 2x scenario and between 3.4 and 63.2% under the 10x scenario, and between 2.2 and 22.2% (root set to stationary frequency), and between 2 and 17% (root set to 1) over both scenarios (2x and 10x) respectively (Table 4). Whereas, the power to reject the false hypothesis of extinction rate symmetry over both scenarios (2x and 10x)

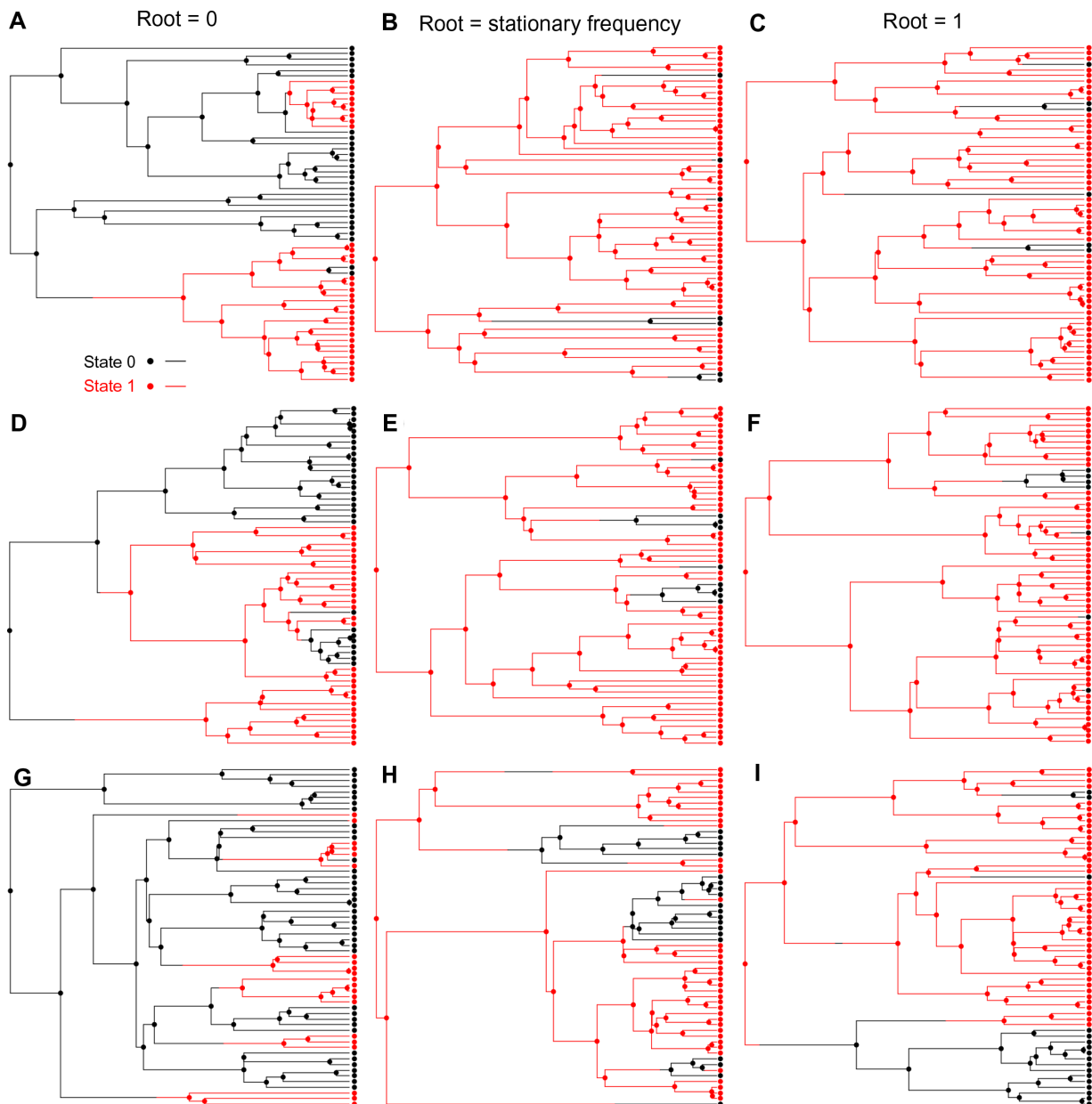


Figure 1. Representative tree histories and character states (for 60 extant taxa) simulated under different root state assumptions (state 0; stationary frequency [the root assumption used by Maddison et al.¹ and Davis et al.¹⁰]; or state 1) and models of diversification rate (2x) asymmetry: **(A–C)** asymmetrical speciation ($\lambda_0 = 0.1$, $\lambda_1 = 0.2$, $\mu_0 = \mu_1 = 0.03$, $q_{01} = q_{10} = 0.01$); **(D–F)** asymmetrical extinction ($\lambda_0 = \lambda_1 = 0.1$, $\mu_0 = 0.06$, $\mu_1 = 0.03$, $q_{01} = q_{10} = 0.01$); **(G–I)** asymmetrical transition ($\lambda_0 = \lambda_1 = 0.1$, $\mu_0 = \mu_1 = 0.03$, $q_{01} = 0.01$, $q_{10} = 0.005$). In each case, nodes and branches of the trees are colored (state 0: black; state 1: red) to indicate the known (because simulated) character state at a particular time.

ranged between 0.6 and 6.4% (root set to 0), 0.2 and 8.8% (root set to stationary frequency), and between 0.6 and 7.6% (root set to 1; Table 3).

Discussion

The circular logic of the stationary frequency root state prior in joint tree–trait simulations. The main purpose of this study was to reassess the power of BiSSE to detect simple rate asymmetries in diversification parameters (in terms of speciation or extinction or transitions) in trees simulated under a variety of simple macroevolutionary models and, especially,

those of small sample sizes. Results show that during the joint tree–trait simulations, the root state assumption has a determining effect on the simulated tree histories in terms of the frequency of state 0 at the root and tips (Tables 2–4). More specifically, trees simulated with the root fixed to state 0 have a higher frequency of this state at both the root and the tips, while state 0 becomes exceedingly rare when the root is set to state 1 or stationary frequency (Tables 2–4). However, the frequency of state 0 at the tips is lower when the simulation (viz. evolution) starts from state 1 rather than state 0 (Tables 2–4). The simple reason for this is that if state 1



has frequency enhancing diversification parameters (higher speciation, transition, or lower extinction rate compared to state 0), and the simulation starts from the ancestral state 0, then at least one shift to the derived state 1 has to occur (ie, evolutionary time must pass) in order to allow taxa with state 1 to diversify with a higher rate (Fig. 1A, D, G). In contrast, if taxa with the ancestral state evolve immediately at higher rate (state 1), then the derived one (state 0) becomes exceedingly rare and isolated (twiggy/tippy) across several clades (Fig. 1B, C, E, F, H, I). Furthermore, under the stationary frequency root state assumption, the state with more frequency enhancing diversification (state 1) is also much more frequently selected as the state at the root of the trees than the alternative one (state 0; eg, Table 2), just as expected under this root prior.^{1,2,25,26}

Evidently, the stationary frequency root state assumption favors the state with frequency-enhancing diversification parameters (state 1) and the root state assumption determines the macroevolutionary model under which the tree evolves (see Table 1). Given a predefined diversification rate asymmetry, this results in circular logic. As such, the joint tree–trait simulations using the stationary frequency root state prior (as previously also conducted by Maddison et al.¹ and Davis et al.¹⁰) favor trees evolving under the same macroevolutionary models as those simulated with the root fixed to state 1, namely, the *lonely state* (for asymmetrical speciation), the *dead end* (for asymmetrical extinction), and the *asymmetrical transition* model (with higher transitions from the derived trait to the ancestral one). These models are all known to produce “tippy” patterns of trait distribution across phylogenies¹⁷ and are likely the cause of a high tip ratio bias in favor of one of the binary states as observed by Davis et al.¹⁰ and also as observed in the present study (Fig. 1B, C, E, F, H, I). Conversely, assigning state 0 to the root for the same simulations results in different macroevolutionary models, namely, the *key innovation* (for asymmetrical speciation), the *key innovation* (for asymmetrical extinction), and the *asymmetrical transition* model (with higher transitions from the ancestral trait to the derived one). In other words, there is a bias in macroevolutionary models selected caused by the joint influence of the time-invariant but state-dependent diversification rates and the stationary frequency root state prior on the tree simulation. Probably because of this bias, the latter three macroevolutionary models escaped the otherwise rigorous statistical power testing of the BiSSE method by Maddison et al.¹ and Davis et al.¹⁰ This latter notion is supported by the fact that under the stationary frequency root state prior, the levels of statistical power for trees of 300 taxa, as observed here, mostly resemble those reported by Davis et al.¹⁰ [compare Tables 2–4 with S1–S3 in the Supplement of Davis et al.¹⁰]. Overall, the present data indicate that the power assessments of Davis et al.¹⁰ and Maddison et al.¹ were confounded with respect to the macroevolutionary models tested, rather than being incorrect.

Consequences for statistical power assessment. As shown here, the power of the BiSSE method for testing

hypotheses of extinction and transition rates is relatively low as already shown previously by Maddison et al.¹ and Davis et al.¹⁰ and mostly unaffected by the root state assumption (Tables 3 and 4). As a consequence, earlier statements by Maddison et al.¹ and Davis et al.¹⁰ regarding the statistical power of this method (eg, the necessity for large datasets) still hold true for testing hypotheses of extinction and transition rates. However, there is one exception: if one examines trees simulated under the (hitherto untested) *key innovation* model (for asymmetrical speciation), BiSSE has a significantly higher statistical power, that is, almost twice as high as inferred previously¹⁰ (see Table 2). In this regard, BiSSE performs at least as good or even better with trees of 300, 60, and 90 and even with trees of 30 taxa than, as previously thought, with trees of 500, 300, and 100 taxa under the 2x scenario, respectively.^{1,10} This effect is even stronger under a scenario of high speciation rate asymmetry (10x), where BiSSE constantly reaches power levels well above 70% and thus about 1.3 to 1.6 times the power in trees of 30, 60, 90, and 300 taxa (see Table 2) than previously thought with trees of 500 taxa.¹⁰ The reason is that Davis et al.¹⁰ and Maddison et al.¹ did not acknowledge that the hypothesis of diversification rate asymmetry can be subdivided into different macroevolutionary models of different tree histories and trait distributions (Table 1), which (in case of speciation rate – and to a lesser extent transition rate asymmetry) are evidently associated with different power levels (Tables 2 and 4). In this sense, trees simulated under the *lonely state* and *key innovation* model (for asymmetrical speciation) represent, respectively, the worst and best case scenarios with regard to the statistical power of BiSSE. As a consequence, it seems feasible to use BiSSE in tests of speciation rate asymmetry, even when using phylogenies of low sample size (Table 2), provided that neither state is (too) scattered or twiggy across the tips of a given phylogeny (as might be the case when a newly evolved character state is associated with an increased rate of speciation).

However, one may conjecture that the increased power of BiSSE to detect speciation rate asymmetries in trees under the *key innovation* model is simply caused by the more balanced tip frequency (a tip ratio of lower than 10:1¹⁰) of the binary states (frequency of state 0) as compared to the *lonely state* model (4.5–57.6% vs. 1.1–13.1%, respectively; see Table 2). This popular view (eg, Soltis et al.¹³), however, is ill founded because Davis et al.¹⁰ showed that the power in trees with asymmetrical rates of speciation and tip ratios of 20:1, 30:1, and 40:1 is much higher than in trees with tip ratios of 3:1, 5:1, and 10:1 but also in those with ratios of 90:1 and 180:1. Moreover, there is apparently no linear relationship between the tip state frequency (or tip state ratio) and statistical power (see Tables 2–4; but also Davis et al.¹⁰). In addition, the present results indicate that a strong signal of key innovation can still be detected by BiSSE with high confidence even though the ancestral state is exceedingly rare among the tips (eg, the frequency of state 0 is 12.6 and 4.5% for trees of 90 and 300 taxa, respectively,

but the power of BiSSE is still well above 80% in either case; see the *key innovation* model under the 10x scenario; Table 2). Overall, there is no firm evidence for assuming that the statistical power of BiSSE decreases for trees with a tip ratio of greater than 10:1 (in favor of either state) as cautioned previously by Davis et al.¹⁰ Hence, the present results indicate that, rather than the tip state frequency, it is the tip state distribution (eg, taxa with derived states either confined to clades under the *key innovation* model or scattered across the tree under the *lonely state* model) that more reliably serves as an a priori indicator of the statistical power of BiSSE for detecting asymmetrical rates of speciation (Table 2; compare Fig. 1A vs. B, C). The question remains as to why, for simulated trees, the statistical power of BiSSE can be apparently high and quite sensitive to the root state assumption for testing the hypothesis of speciation rate asymmetry (Table 2) but relatively low and unresponsive to the root state assumption for testing hypotheses of extinction/transition asymmetries (Tables 3 and 4). Of course, this discrepancy could simply reflect the general difficulty to estimate transition and, especially, extinction rates.^{1,10,28,29} On the other hand, this discrepancy might also reflect the different intensity of signal each diversification parameter asymmetry is able to leave within a phylogenetic tree. For example, in Figure 1A, there are 28 speciation events associated with state 0 and 32 speciation events with state 1 but only two transitions from state 0 to 1, and only one from 1 to 0. Apparently, BiSSE is able to distinguish differential rates of speciation from this sample, but this does not imply statistical certainty of this method when it comes to distinguishing among transitions that have equal rates and those that involve, for example, two forward and one reverse transitions. Here it becomes apparent why more species-rich (simulated) trees are generally associated with higher power (see Davis et al.¹⁰; but also Tables 2–4), simply because larger trees (with, by necessity, more speciation events) offer a broader scene for state-dependent differential diversification parameters (speciation, extinction, transition) to act on, thus leaving a statistically distinguishable footprint. In other words, while the present study demonstrates that asymmetrical speciation can leave a strong unambiguous signal even in small phylogenetic trees (Fig. 1A, Table 2), the findings agree with those of Maddison et al.¹ and Davis et al.¹⁰ in that the accumulated evolutionary information of large trees is necessary for the BiSSE method to be able to distinguish among the comparably weak signatures of transitions and extinctions.

Conclusions

The present study demonstrates that Maddison et al.¹ and Davies et al.¹⁰ underestimated the power of BiSSE for detecting speciation rate asymmetries due to a bias, which is based on the joint influence of the time-invariant but state-dependent diversification rates and the *stationary frequency* root state prior on the tree simulation. In particular, these earlier simulation studies^{1,10} lead to an overestimation of type-II

statistical error because they did not account for the fact that hypotheses of diversification rate asymmetry can be subdivided into different macroevolutionary models with different tree histories and trait distributions, and which – in the case of speciation rate asymmetry – can have different power levels. Thus, as shown here, trees simulated under the *lonely state* vs. *key innovation* models (for asymmetrical speciation) represent, respectively, the worst and best case scenarios with regard to the statistical power of BiSSE. In this context, BiSSE performs at least as good, or even better, with trees of 300, 60, and 90 and even 30 taxa than, as previously thought, with trees of 500, 300, and 100 taxa under the 2x scenario, respectively. However, even in the presence of a high tip ratio bias, a scenario of high speciation rate asymmetry (10x) may already be detectable with confidence in trees with a sample size ≥ 30 taxa, provided there is a strong signal of a key innovation. It seems feasible, therefore, to use BiSSE in tests of speciation rate asymmetry even for trees of relatively low sample size (eg, when the aim of a phylogenetic study is simply to confirm suspected cases of a key innovation). Nonetheless, the present results also reemphasize those of Maddison et al.¹ and Davis et al.¹⁰, indicating that still large sample sizes are required when the full state-dependent diversification history needs to be disentangled (in terms of speciation and extinction and transition rates). As a side note, researchers should be less concerned about tree size and tip ratio biases but rather be focused on the trait distribution within a given tree. Although the results reported here only refer to power tests in BiSSE, they might also bear significance for similar tests in related *State dependent Speciation and Extinction* (SSE) models (eg, QuaSSE, GeoSSE, ClaSSE, MuSSE, BiSSE-BMA, HiSSE, etc.). In any event, future power tests should involve simulated trees with different root state assumptions in order to obtain as far as possible an unbiased set of trees. In sum, the present study indicates that the BiSSE method performs better than its reputation and, almost ten years following its initial description,¹ still offers fascinating insights into its versatility to detect diversification rates asymmetries of interest.

Acknowledgment

Special thanks are expressed to Hans Peter Comes (University of Salzburg) and six anonymous reviewers for very helpful and insightful comments on earlier versions of this manuscript.

Author Contributions

Conceived and designed the experiments: AG. Analyzed the data: AG. Wrote the first draft of the manuscript: AG. Contributed to the writing of the manuscript: AG. Agree with manuscript results and conclusions: AG. Jointly developed the structure and arguments for the paper: AG. Made critical revisions and approved final version: AG. The author reviewed and approved of the final manuscript.



Supplementary File 1

R functions used for the joint tree–trait simulations as well as for recording the root and tip states of the simulated tree histories.

REFERENCES

- Maddison WP, Midford PE, Otto SP. Estimating a binary character's effect on speciation and extinction. *Syst Biol.* 2007;56:701–10.
- FitzJohn RG, Maddison WP, Otto SP. Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Syst Biol.* 2009;58:595–611.
- Ng J, Smith SD. How traits shape trees: new approaches for detecting character state-dependent lineage diversification. *J Evol Biol.* 2014;27:2035–45.
- FitzJohn RG. Quantitative traits and diversification. *Syst Biol.* 2010;59:619–33.
- Goldberg EE, Lancaster LT, Ree RH. Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Syst Biol.* 2011;60:451–65.
- Goldberg EE, Igić B. Tempo and mode in plant breeding system evolution. *Evolution.* 2012;66:3701–9.
- FitzJohn RG. Diversitree: comparative phylogenetic analyses of diversification in R. *Methods Ecol Evol.* 2012;3:1084–92.
- Silvestro D, Zizka G, Schulte K. Disentangling the effects of key innovations on the diversification of Bromelioideae (Bromeliaceae). *Evolution.* 2014;68:163–75.
- Beaulieu JM, O'Meara BC. Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Syst Biol.* 2016;65:583–601.
- Davis M, Midford P, Maddison W. Exploring power and parameter estimation of the BiSSE method for analyzing species diversification. *BMC Evol Biol.* 2013;13:38.
- Rabosky DL, Goldberg EE. Model inadequacy and mistaken inferences of trait dependent speciation. *Syst Biol.* 2015;64:340–55.
- Igić B, Busch JW. Is self-fertilization an evolutionary dead end? *New Phytol.* 2013;198:386–97.
- Soltis DE, Segovia-Salcedo MC, Jordon-Thaden I, et al. Are polyploids really evolutionary dead-ends (again)? A critical reappraisal of Mayrose et al. (2011). *New Phytol.* 2014;202:1105–17.
- Mayrose I, Zhan SH, Rothfels CJ, et al. Recently formed polyploid plants diversify at lower rates. *Science.* 2011;333:1257.
- Price SA, Hopkins SSB, Smith KK, Roth VL. Tempo of trophic evolution and its impact on mammalian diversification. *Proc Natl Acad Sci U S A.* 2012;109:7008–12.
- de Vos JM, Hughes CE, Schneeweiss GM, Moore BR, Conti E. Heterostyly accelerates diversification via reduced extinction in primroses. *Proc Biol Sci.* 2014;281:20140075.
- Bromham L, Hua X, Cardillo M. Detecting macroevolutionary self-destruction from phylogenies. *Syst Biol.* 2016;65:109–27.
- Tripp EA, Manos PS. Is floral specialization an evolutionary dead-end? Pollination system transitions in *Ruellia* (Acanthaceae). *Evolution.* 2008;62:1712–32.
- Smith SD, Miller RE, Otto SP, FitzJohn RG, Rausher MD. The effects of flower color transitions on diversification rates in morning glories (*Ipomoea* subgenus *Quamoclit*, Convolvulaceae). In: Long M, Gu H, Zhou Z, eds. *Darwin's Heritage Today*. Beijing: Higher Education Press; 2010:205–26.
- Goldberg EE, Kohn JR, Lande R, Robertson KA, Smith SA, Igić B. Species selection maintains self-incompatibility. *Science.* 2010;330:493–5.
- Gamisch A, Fischer GA, Comes HP. Multiple independent origins of autopolyploidization in tropical orchids (*Bulbophyllum*) in light of the hypothesis of selfing as an evolutionary dead end. *BMC Evol Biol.* 2015;15:192.
- Gamisch A, Fischer GA, Comes HP. Frequent but asymmetric niche shifts in *Bulbophyllum* orchids support environmental and climatic instability in Madagascar over Quaternary time scales. *BMC Evol Biol.* 2016;16:14.
- Agarsson I, Avilés L, Coddington JA, Maddison WP. Sociality in theridiid spiders: repeated origins of an evolutionary dead end. *Evolution.* 2006;60:2342–51.
- Maddison WP, Maddison DR. Mesquite: a modular system for evolutionary analysis. 2007. Version 2.0b3. Available at: <http://mesquiteproject.org/>.
- Goldberg EE, Igić B. On phylogenetic tests of irreversible evolution. *Evolution.* 2008;62:2727–41.
- Felsenstein J. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol.* 1981;17:368–76.
- Maddison WP. Confounding asymmetries in evolutionary diversification and character change. *Evolution.* 2006;60:1743–6.
- Rabosky DL. Extinction rates should not be estimated from molecular phylogenies. *Evolution.* 2010;64:1816–24.
- Rabosky DL. Challenges in the estimation of extinction from molecular phylogenies: a response to Beaulieu and O'Meara. *Evolution.* 2016;70:218–28.