RESEARCH ARTICLE

Methods in Ecology and Evolution | BRITISH ECOLOGICAL SOCIETY

Continuous traits and speciation rates: Alternatives to state-dependent diversification models

Michael G. Harvey [ID] | Daniel L. Rabosky

Department of Ecology and Evolutionary Biology and Museum of Zoology, University of Michigan, Ann Arbor, MI, USA

Correspondence
Michael G. Harvey
Email: mgh272@gmail.com

Abstract

1. Many quantitative traits, for example body size, have been hypothesized to influence the diversification dynamics of lineages over macroevolutionary time-scales. The Quantitative State Speciation-Extinction (QuaSSE) model and related methods provide an elegant framework for jointly modelling the relationship between change in continuous traits and diversification. However, model misspecification and phylogenetic pseudoreplication can result in elevated false discovery rates in this and other state-dependent speciation-extinction models.

2. Here, we evaluate alternative trait-dependent diversification methods that do not formally model the relationship between traits and diversification, but instead test for correlations between summary statistics of phylogenetic branching patterns and trait variation at the tips of a phylogenetic tree (hereafter tip-rate correlations or TRCs). We compare alternative branching pattern statistics and significance tests, and we evaluate their performance relative to QuaSSE under a range of evolutionary scenarios.

3. We found that a simple statistic derived from branch lengths (inverse equal splits) can detect trait-associated rate variation, and that a simulation-based method performs better than phylogenetic generalized least squares for testing the significance of trait-rate correlations. This test (*ES-sim*) had better power to detect trait-dependent diversification than other TRCs. By testing the approach across a diverse set of simulation scenarios, we found that *ES-sim* is similar to QuaSSE in statistical power. However, the approach rarely led to false inferences of trait-dependent diversification, even under conditions that are problematic for formal state-dependent models. We illustrate the application of *ES-sim* to real data by re-assessing the relationship between dispersal ability and diversification in Furnariid birds.

4. We conclude that simple, semi-parametric tests like *ES-sim* represent a promising approach for trait-dependent diversification analyses in groups with heterogeneous diversification histories and provide a useful alternative or complement to formal state-dependent speciation-extinction models.

KEYWORDS
comparative methods, inverse equal splits statistic, phylogenetic generalized least squares, state-dependent speciation and extinction models, trait-dependent diversification

　　　　1

1 | INTRODUCTION

Traits of organisms can impact their propensity for evolutionary diversification through time (Stanley, 1975; Jablonski, 2008). Many traits thought to be responsible for trait-dependent diversification are quantitative or continuous, rather than discrete. Body size may be associated with diversification, for example, if the higher metabolic rates or faster generation times typical of smaller-bodied species lead to higher evolutionary rates (Gittleman & Purvis, 1998; Glazier, 1987; Marzluff & Dial, 1991). Other examples of continuous traits with hypothesized links to diversification rates include dispersal ability (Claramunt, Derryberry, Remsen, & Brumfield, 2012; Phillimore, Freckleton, Orme, & Owens, 2006), ecological specialization (Futuyma & Moreno, 1988), strength of sexual selection (Panhuis, Butlin, Zuk, & Tregenza, 2001; West-Eberhard, 1983), range size (Rosenzweig, 1995) and latitudinal range (Cardillo, 1999).

Early investigations of trait-dependent diversification involved comparing the diversities of sister clades that differed in some trait of interest (Barraclough, Nee, & Harvey, 1998; Farrell, Dussourd, & Mitter, 1991; Mitter, Farrell, & Wiegmann, 1988). In recent years, the study of trait-dependent diversification has focused on jointly modelling diversification dynamics and trait evolution across a phylogeny (e.g. Bokma, 2008; Paradis, 2005). The most recent such method for continuous traits, quantitative state speciation and extinction or QuaSSE (FitzJohn, 2010), allows speciation and extinction rates to vary as arbitrary (user-defined) functions of trait values. The degree to which the phylogeny and trait data are explained by models with and without trait-dependent diversification can then be compared in a likelihood framework. QuaSSE and related state-dependent speciation-extinction (SSE) models for binary (BiSSE) and multistate (MuSSE) characters are powerful tests for detecting trait-dependent diversification (Beaulieu & O'Meara, 2016; FitzJohn, 2012; FitzJohn, Maddison, & Otto, 2009; Maddison, Midford, & Otto, 2007). However, various authors have found high incidences of false inference of trait-dependent relationships using SSE methods (Maddison & FitzJohn, 2015; Rabosky & Goldberg, 2015; Rabosky & Huang, 2016), including QuaSSE (FitzJohn, 2010; Machac, 2014).

Recently, Beaulieu and O'Meara (2016) noted that many false inferences of state-dependent diversification ultimately follow from an incorrectly formulated hypothesis-testing framework. Specifically, formal tests for trait-dependent diversification have typically involved comparing a model with trait-dependent diversification (e.g. BiSSE) to a model with no diversification rate variation (e.g. constant-rate birth–death process). This procedure is problematic, because state-dependent models frequently provide a good fit whenever diversification rate variation is present in the data, even if it is unlinked to the character state of interest. As noted by Beaulieu and O'Meara (2016), this outcome is not a "false positive" in the statistical sense, because it reflects correct rejection of an overly simplistic null hypothesis rather than incorrect rejection of a true null hypothesis. Nonetheless, we continue to refer to "false positives" and "false discovery rates" in the remainder of the text, partly for brevity and partly because the biological interpretation of the result is that observed diversification

dynamics are associated with trait variation even though in actuality they are not.

As an alternative to overly simplistic null models, Beaulieu and O'Meara developed several models (CID-2, CID-4) that allow diversification rates to vary across the phylogeny as a function of unobserved character states. Use of these hidden-state models in conjunction with BiSSE can dramatically reduce false inferences of trait-dependent diversification (Beaulieu & O'Meara, 2016; Rabosky & Goldberg, 2017). However, an equivalent hidden-state model has yet to be developed for quantitative characters, and modelling continuous variation in diversification rates across a phylogeny as a function of an unobserved latent variable poses a challenging problem in numerical analysis.

An alternative class of methods for trait-dependent diversification analyses involves assessing the correlation between variation in a trait of interest across the tips of a phylogeny and tip-specific estimates of speciation rates. These tip-rate correlation (hereafter TRC) methods bypass the need for a fully parameterized model of diversification and trait evolution. Speciation rate metrics used in TRC tests are generally simple indices based on the waiting times between speciation events and ignore extinction; as such, they provide a more reliable index of speciation than net diversification in many scenarios (Belmaker & Jetz, 2015). Freckleton, Phillimore, and Pagel (2008) introduced a TRC method for continuous traits, measuring speciation rate as the mean internode distance (branch lengths) between the root and a given tip. Jetz, Thomas, Joy, Hartmann, and Mooers (2012) used a related measure (the "DR statistic") that assigns more weight to recent branch lengths than to branches early in the clade's history. Bromham, Hua, and Cardillo (2016) and Hua and Bromham (2016) present a suite of alternative summary statistics describing phylogenetic branching patterns.

Tip-rate correlation methods involve, in addition to choice of speciation rate metrics, a strategy for assessing the significance of correlations between traits and diversification. Most TRC tests have used phylogenetic generalized least squares (PGLS) to assess the significance of correlations while accounting for shared evolutionary history among relatives (Freckleton et al., 2008; Harvey et al., 2017; Jetz et al., 2012). PGLS accounts for shared history using the expected covariance of residuals based on the phylogenetic distance between species and assuming some model of evolutionary change (e.g. random Brownian motion). Although this strategy may be appropriate for modelling covariance among species in many traits, it is unclear whether Brownian motion and similar models appropriately account for covariance in comparisons involving speciation rate metrics based on phylogenetic branching patterns, which change in concert between sister lineages at each node rather than randomly along branches.

The significance of trait-speciation correlations can also be assessed by testing whether the observed correlation between trait values and speciation rate metrics lies outside a distribution constructed by simulation under a null evolutionary model (e.g. Garland, Dickerman, Janis, & Jones, 1993). Rabosky and Huang (2016) developed a test (STRAPP) that builds a null distribution of associations between speciation metrics and trait variation by permuting trait

values among diversification rate regimes inferred using BAMM (Rabosky, 2014) or potentially other multiprocess diversification models, but the power of this approach is limited by the number of distinct rate regimes present in a given phylogeny. Bromham et al. (2016) and Hua and Bromham (2016) developed tests that construct null distributions of trait-speciation associations by backward simulation of phylogenetic trees with or without trait dependence. The FiSSE approach (Rabosky & Goldberg, 2017) constructs a null distribution by simulating change in a binary trait across the empirical phylogeny under a simple Markovian model. Rabosky and Goldberg (2017) demonstrated that this strategy performed well across a diverse range of testing scenarios, although FiSSE was limited to analysis of discrete characters.

Here, we explore the performance of TRC tests for trait-dependent speciation in quantitative characters. We use simulations to evaluate the performance of alternative tip-specific speciation rate metrics. We also compare strategies for significance testing including PGLS and simulation-based approaches to generating a null distribution of speciation-trait correlations. We then evaluate the performance of our best-performing TRC method relative to QuaSSE using simulated and empirical data. Our simulation scenarios encompass a range of possible model violations that might lead to spurious inference of relationships between traits and diversification.

2 | MATERIALS AND METHODS

2.1 | Tip-rate correlation tests

We evaluated three tip-specific metrics of speciation rate for use in TRC tests. The node density (ND) is the simplest measure of speciation rate and is simply the ratio of the number of speciation events (nodes) along a particular root-to-tip path divided by the age of the clade, or

$$ND_i = \frac{N_i}{T}$$

where ND_i is the speciation rate for tip i, N_i is the number of nodes between tip i and the root of the tree, and T is the total evolutionary time between the tips and the root. Alternatively, we can estimate the speciation rate for a particular tips as the inverse of the corresponding equal splits (ES) measure, which was originally designed to capture the amount of unique evolutionary history that could be apportioned among each tip in a phylogenetic tree (Redding & Mooers, 2006):

$$ES_i = \sum_{j=1}^{N_i} l_j \frac{1}{2^{j-1}}$$

Here, ES_i is the speciation rate for tip i, N_i is the number of edges between tip i and the root of the tree, and l_j is the length of each edge j beginning with the terminal edge ($j = 1$) and terminating with the root edge ($j = N_i$). Effectively, ES represents the sum of the lengths of the edges subtending a tip, with each edge root-ward down-weighted by 1/2. The log-transformed ES is the diversification rate statistic ("DR statistic") employed in trait-dependent diversification tests by Jetz

et al. (2012). Finally, the inverse of the terminal branch lengths (TB) can be used as a measure of the time since the last speciation event, with lineages exhibiting higher speciation rates expected to have shorter terminal branches. This statistic has been used recently for trait-dependent diversification analyses (e.g. Bromham et al., 2016; Gomes, Sorenson, & Cardoso, 2016). In summary, ND captures splitting dynamics over the entire history of the lineage leading to a tip, TB captures only the dynamics at the tips, and ES uses information from the full root-to-tip path but is weighted towards branching patterns nearer the tips.

We evaluated two methods of determining the significance of associations between trait variation and speciation rate metrics: phylogenetic generalized least squares (PGLS) and a simulation test involving comparison of the observed correlation with a null set of associations between the speciation metrics and trait values. We used caper (Orme et al., 2013) to fit PGLS models assuming a Brownian motion model for the error structure, following prior studies (Freckleton et al., 2008; Gomes et al., 2016; Jetz et al., 2012). For the simulation test, we simulated Brownian trait evolution 1,000 times across the empirical tree using root state and diffusion rate (σ^2) parameters from the maximum-likelihood fit of a Brownian motion model to the original data. Note that PGLS and the simulation approach need not yield identical results: PGLS assumes that the residuals of the relationship between traits and speciation rates can be modelled as a Brownian motion on the phylogeny (Revell, 2010); the simulation approach assumes Brownian motion in the trait only. Two-tailed p-values were computed by comparing the Pearson's correlation between the speciation rate metric and trait values in the original data to the correlation between the speciation rate metric and the simulated trait values. We note that test statistics aside from Pearson's correlation could certainly be used, including statistics that accommodate nonlinear associations between traits and diversification (see Section 4).

2.2 | General overview of performance tests

We used simulated datasets to evaluate the performance of TRC methods. First, we compared the power of the three speciation rate metrics (ND, ES and TB) to detect associations between speciation and traits changing at different rates. Second, we evaluated the two strategies for significance testing (PGLS and simulations), based on both power and false discovery rates, in datasets of different sizes. Third, we evaluated whether power was reduced when the assumption of Brownian motion used in our simulation-based significance test was violated. Fourth, we compared the power of our best-performing TRC test to that of QuaSSE. Finally, we compared false discovery rates of the TRC test to those of QuaSSE across a wide range of evolutionary scenarios.

2.3 | Speciation rate metrics

We evaluated the ability of the three speciation rate metrics (ND, ES and TB) to infer true relationships between continuous traits and speciation rates by assessing their performance on trees simulated with a QuaSSE process (FitzJohn, 2010). Using diversitree

(FitzJohn, 2012), we performed forward-in-time pure-birth simulations in which speciation rate was related to trait values according to a linear function (slope = 0.004). Traits evolved along the tree under a Brownian motion process. Different speciation rate metrics may perform better depending on the rate of trait evolution and associated rate of change in diversification rates in a dataset. For example, in rapidly evolving traits we might expect trait variation at the tips to be associated with length variation only in the most recent branches. For such traits, TB may be the best diversification metric. For slowly evolving traits, ND may be preferred because it captures variation in diversification back to the root of the phylogeny. Therefore, we simulated trait-dependent diversification under a series of diffusion rates of trait change (σ^2) encompassing a range of values (0.00006, 0.0006, 0.006, 0.06, 0.6, 6, 60) similar to the spectrum of body size evolution rates observed in empirical studies (Harmon et al., 2010). At each rate of trait change, we simulated 100 datasets with 250 species each and assessed the power of all three speciation rate metrics to recover the signal of trait-dependent diversification. We evaluated power by calculating the proportion of simulated datasets for which trait-dependent diversification was correctly inferred using both of the significance testing approaches described below.

2.4 | Significance tests

We compared PGLS and simulation-based significance tests using the 250-tip datasets simulated at an intermediate rate of trait change ($\sigma^2 = 0.06$) from the previous section, but added sets of datasets ($n = 100$) containing 50 tips and 1,250 tips to assess the effect of dataset size on test performance. We also simulated datasets in which there was no relationship between speciation rate and trait values (simulated using simple Brownian motion) to measure the false discovery rate of each test. For clarity, a full list of the trait-dependent diversification tests examined in the study is presented in Table 1.

2.5 | Evaluating power with violations of Brownian trait evolution

Our simulation-based significance test relies on a simple Brownian motion process to generate the null distribution of trait values. Trait model misspecification can, however, lead to spurious results in comparative analyses (Diaz-Uriarte & Garland, 1996; Pennell, FitzJohn, Cornwell, & Harmon, 2015). To investigate the sensitivity of our method to misspecification of the model of trait evolution, we simulated datasets under an Ornstein-Uhlenbeck (OU) model and compared the performance of the Brownian motion simulation test to an alternative test in which the correct (OU) model was used to generate the null distribution. We simulated trees and OU trait evolution using diversitree with "pull" towards the optimum determined by the linear function $\alpha(\hat{x} - x)$ as suggested by FitzJohn (2010). We examined α values of 0.002, 0.02 and 0.2. These absolute values mean little because α is scaled to tree depth (Cooper, Thomas, Venditti, Meade, & Freckleton, 2016), but this range included the parameter space across

TABLE 1 Trait-dependent diversification tests examined in this study

	Test	References
Joint model of trait evolution and diversification		
1	QuaSSE	FitzJohn (2010)
Tip-rate correlation (TRC) tests		
PGLS tests		
2	*ES-pgls*	Jetz et al. (2012)
3	*ND-pgls*	Freckleton et al. (2008)
4	*TB-pgls*	Gomes et al. (2016)
Simulation tests		
5	*ES-sim*	This study
6	*ND-sim*	This study
7	*TB-sim*	This study

which all methods lost power to detect trait dependence. At each α value, we simulated 100 datasets with 250 tips with trait dependence and examined the power of simulation-based tests using Brownian and OU models. OU models were fit using the R package geiger (Harmon, Weir, Brock, Glor, & Challenger, 2008) and OU simulations used phytools (Revell, 2012).

2.6 | Power comparison with QuaSSE

We compared the best-performing TRC test of trait-dependent diversification based on the above analyses to QuaSSE (FitzJohn, 2010). We used the same sets of datasets with different numbers of tips (50, 250 and 1,250 species) and with and without trait dependence that were examined in "Significance tests" above to evaluate the power and false discovery rates of both tests. We used QuaSSE to fit a model in which the trait exhibited a linear relationship with speciation vs. one in which speciation was constant with respect to trait variation. We used likelihood ratio tests for model comparison and to determine whether trait dependence was supported in each case.

2.7 | False discovery rate comparison with QuaSSE

A major goal of this study is to evaluate methods that may overcome the erroneous inferences of trait-dependent diversification ("false discovery" for brevity) often observed in analyses with formal state-dependent speciation-extinction tests (Machac, 2014; Rabosky & Goldberg, 2015; but see Beaulieu & O'Meara, 2016). We therefore examined false discovery rates of our best-performing TRC test and QuaSSE in datasets simulated under a broad spectrum of scenarios where the focal trait was unlinked to diversification rates, roughly following Rabosky and Goldberg (2017). These scenarios included sets of trees simulated under a constant diversification rate, a diversification rate slowdown, a QuaSSE tree with trait dependence, a BiSSE tree with trait dependence, the coral supertree from Huang and Roy (2015), the carnivore tree from Nyakatura and Bininda-Emonds (2012) and a set of diversity-dependent multiprocess trees

with a single shift between decoupled diversification processes from Rabosky (2014). These were combined with each of the following trait simulation scenarios: Brownian motion, Brownian motion with a single rate shift, Brownian motion with a jump in the mean values in one clade, no phylogenetic signal in the trait (i.e. evolving as if along a star-shaped tree), Brownian motion across most of the tree but white noise (no phylogenetic signal) in a single subclade, Brownian motion but with one clade fixed for a single trait value, shifts between two discrete trait distributions (normally distributed), an OU process with a single optimum and weak "pull" towards the optimum and an OU process with a single optimum and strong "pull" towards that optimum. The resulting scenarios represent 63 unique combinations of diversification and trait evolution settings, but in none of the scenarios is diversification rate linked to trait values (Table S1). For each combination, one iteration of trait evolution was simulated on each of 50 trees from the tree set, except in combinations involving the coral supertree, for which 50 iterations of trait evolution were simulated on the single tree. Thus, 50 simulated datasets were generated for each of the 63 scenarios. We then ran the TRC test and QuaSSE on each iteration of each scenario and tabulated the frequency with which each method incorrectly inferred state-dependent diversification. In some scenarios, the find.mle optimizer from QuaSSE failed under the default settings. In these cases, we used the optim function with the Nelder–Mead algorithm using starting parameters estimated by QuaSSE. If both optimization strategies failed for any particular iteration, we treated the iteration as failed and excluded it from further analysis.

2.8 | Trait-dependent diversification in *Furnariidae*

We evaluated the results of different tests of trait-dependent diversification on an empirical dataset previously found to exhibit trait-dependent diversification dynamics (Claramunt et al., 2012). This dataset includes a time-calibrated phylogenetic tree of birds in the family *Furnariidae* and measurements of the hand-wing index (HWI), a morphological metric that predicts dispersal ability. In continental settings, high dispersal ability is expected to inhibit speciation in birds, because it allows populations to maintain genetic cohesion in the presence of biogeographic barriers. Accordingly, Claramunt et al. (2012) found that species with high HWI had relatively low speciation rates based on a QuaSSE analysis. In fact, their best model (log-Likelihood [ln$L$] = −1531.6) included a sigmoidal relationship in which lineages with high HWI had low speciation rates, those with low-to-moderate HWI had high speciation rates, and those with the smallest HWI again had somewhat lower speciation rates (i.e. an "intermediate dispersal" model). However, a simple linear model in which HWI was negatively correlated with speciation rate was still a better fit (ln$L$ = −1535.6) than a model in which speciation was unrelated to HWI (ln$L$ = −1539.7). Thus, we expect a significant negative linear correlation between HWI and speciation rate in this dataset.

We first examined the Furnariid dataset using our best-performing TRC method assuming Brownian trait evolution as described above. We removed one species (*Asthenes luizae*) lacking HWI information,

resulting in a final set of 282 species. Although Brownian simulations perform reasonably well in TRC tests even when the trait evolved under a different model (see Section 3.1), comparing the fit of alternative trait evolution models may still be advisable in analyses of empirical datasets. We therefore compared the fit of a model of Brownian motion, an OU model, an early burst model, and a white noise model assuming no covariance among species to the Furnariid dataset using AICc scores. We also used parametric bootstrapping to evaluate model adequacy by simulating 1000 trait datasets under the best-fit model and assessing whether the log likelihood of the real data fell outside the 95% confidence interval of log likelihoods from the simulated datasets. We compared the results of *ES-sim* using Brownian motion, *ES-sim* using the best-fit trait evolution model, and QuaSSE.

3 | RESULTS

3.1 | Comparison of performance among TRC tests

The most powerful tip-rate correlation (TRC) test for trait-dependent diversification combined ES (the inverse of the equal splits measure) with a simulation-based significance test (Figure 1). We refer to this test hereafter as *ES-sim*. TB (the inverse of terminal branch lengths) and ND (node depth) both exhibited lower power than ES in tests using the simulation-based significance test. Pearson's correlation performed similar to or better than other test statistics in the simulation test (Table S2). PGLS-based tests had lower power than simulation-based tests in ES and TB. PGLS with ND actually performed better
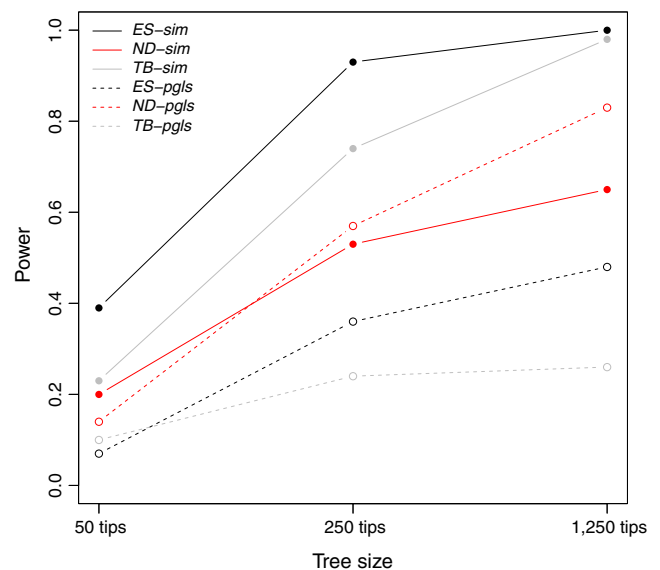

FIGURE 1 A comparison of the power of tip-rate correlation (TRC) tests of trait-dependent diversification differing in the speciation rate metrics examined and in the approach for significance testing. The diversification metrics examined were the inverse of the equal splits metric (ES), node density (ND) and the inverse of the terminal branch length (TB). The significance tests examined were phylogenetic generalized least squares (PGLS) and a simulation test in which the observed correlation was compared to a null distribution of trait-diversification correlations

than the simulation-based test with ND, but was still less powerful than *ES-sim*. All tests performed better on 250-tip trees than on 50-tip trees, with more modest improvements on 1,250-tip trees relative to 250-tip trees. Rates of false positives were low across all tests when they were used to examine datasets simulated without trait-dependent diversification (Table S3).

All TRC tests examined had the greatest power at intermediate rates of trait change given a linear relationship between the trait and speciation rate with a slope of 0.004 (σ^2; Figure 2 and Figure S1). In the simulation tests, all three metrics performed poorly at very slow rates ($\sigma^2 \leq 0.0006$) presumably due to minimal variation in speciation rate at this value, ES had the highest power at intermediate rates, and ES and TB performed similarly at very high rates ($\sigma^2 \geq 6$; Figure 2). ES, therefore, may be the best metric for use in simulation-based tests of trait-dependent diversification across a broad range of rates of trait evolution.

ES-sim in which Brownian motion was used for trait simulations had lower power to detect trait-dependent diversification when the true model of trait evolution was an OU model, particularly as the "pull" towards an optimum increased (Table 2). However, an *ES-sim* test in which the correct, OU model was used for simulations performed similarly to *ES-sim* with the Brownian motion model, suggesting that a mismatched trait evolution model is not the problem but rather that the signal of trait-dependent diversification is obscured by an OU model of trait change. On a related note, we also found that QuaSSE showed similar reductions in power with greater deviation from Brownian motion in the trait evolution model (Table S4).
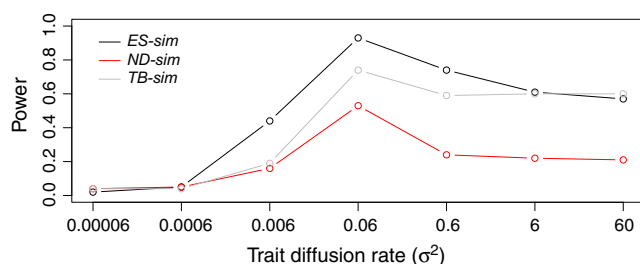

FIGURE 2 A comparison of the power of simulation-based TRC (Tip-rate correlation) tests with alternative speciation rate metrics across different rates of trait evolution and associated rates of change in diversification dynamics equal-splits (ES), node density (ND), terminal branch lengths (TB)

TABLE 2 Performance of *ES-sim* when trait analysed was simulated under OU model

	ES-sim (Brownian)		ES-sim (OU)	
	Power	**FDR**	**Power**	**FDR**
OU with alpha = 0.002	0.89	0.04	0.85	0.05
OU with alpha = 0.02	0.33	0.01	0.36	0.01
OU with alpha = 0.2	0.01	0.00	0.04	0.03

FDR, false discovery rate.
"Brownian" and "OU" in parentheses reflect the trait evolution model used for the simulation-based significance test.

3.2 | Comparison of performance relative to QuaSSE

QuaSSE had slightly more power to detect trait-dependent diversification in datasets of 50 and 250 tips than *ES-sim* (Table 3). In the set of 63 diversification and trait evolution scenarios modelled after that of Rabosky and Goldberg (2017), we found false discovery rates were substantially higher (5% or more) in QuaSSE than in *ES-sim* in 43 of 63 scenarios (Figure 3). False discovery rates were similar (within 5%) in eight scenarios, and were higher in *ES-sim* in nine scenarios. QuaSSE results failed in all iterations in the remaining three scenarios. The *ES-sim* false discovery rate was 10% or lower in all but one scenario (it was 18% in the coral tree with trait simulations in which one clade had trait values with no phylogenetic signal). However, QuaSSE false discovery rates were higher than 18% in 54 scenarios. The scenarios with the highest false discovery rates were those including the empirical carnivore tree and the simulated diversity-dependent multi-process trees with a single shift between decoupled diversification processes, which were (along with the coral supertree) the largest trees examined.

TABLE 3 Power of *ES-sim* relative to QuaSSE

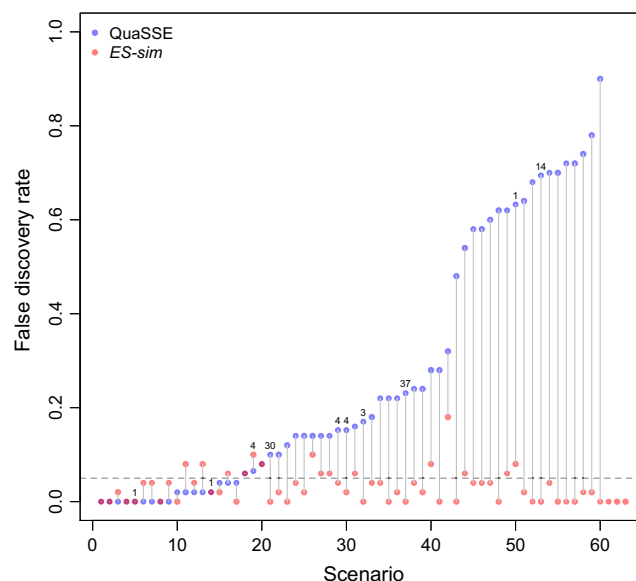|	**50 tips**	**250 tips**	**1,250 tips**
ES-sim	0.38	0.93	1.00
QuaSSE	0.45	0.98	1.00

FIGURE 3 False discovery rates of *ES-sim* compared to QuaSSE across 63 diversification and trait evolutionary scenarios. Scenarios are numbered across the bottom axis, and vertical lines connect the false discovery rates of *ES-sim* and QuaSSE. The numbers above individual points denote the number of iterations for that scenario (of 50) for which no QuaSSE results could be obtained due to numerical failures; no number is given for scenarios where QuaSSE worked for all iterations. In the four scenarios furthest to the right, QuaSSE failed on all replicates and no point is presented for QuaSSE

3.3 | Trait-dependent diversification in *Furnariidae*

Consistent with the results of Claramunt et al. (2012), our QuaSSE results indicated a model containing a linear association between the hand-wing index (HWI) and speciation rate was a better fit than a model in which speciation was constant with respect to HWI in Furnariid birds (likelihood ratio test: χ^2 = 8.054, *p* = .005). The best-fit model of trait evolution for HWI was an Ornstein-Uhlenbeck (OU) model (AICc = 1467.1 vs. AICc = 1481.7 with Brownian motion; Table S5). However, OU models can be incorrectly favoured over Brownian motion in some cases (Cooper et al., 2016). Parametric bootstrapping indicated that the real data were not distinguishable from datasets simulated under either a Brownian (*p* = .094) or OU (*p* = .108) model. We therefore conducted *ES-sim* tests using both Brownian and OU models. We failed, however, to detect significant trait-dependent correlations in the Furnariid dataset using *ES-sim* with either OU (*p* = .33) or Brownian motion (*p* = .40). The Pearson's correlation coefficient [ρ] was -0.16, indicating 2.56% of the variance in speciation rate was explained by variation in HWI. The slope of a linear model fit to the data was -0.02, which equates to model-based speciation rates 0.11 species/My higher in species with the lowest HWI values vs. the highest (speciation rates observed across species in the dataset ranged from 0.04 to 1.37 species per My). Although these effect size measures do not account for covariance among related species, they do provide additional evidence that dispersal ability is a weak predictor of speciation rates in this group. The Furnariid tree appears to show some heterogeneity in diversification dynamics (Figure 4a), which might explain the inference of trait-dependent diversification with QuaSSE. QuaSSE analysis of 100 traits simulated with random Brownian motion on the Furnariid tree revealed a high rate (40%) of false positives. The positive result in QuaSSE may also be partly due to phylogenetic pseudoreplication; many of the points with high values of HWI and low speciation rates are in one clade, the Sclerurinae (Figure 4a,b).

4 | DISCUSSION

We assessed the performance of a series of TRC methods for testing hypotheses about the relationship between continuous-valued traits and lineage diversification rates. We focused on three measurements of tip-specific speciation rate (ND, ES and TB) under two general approaches for significance testing (PGLS and null simulations). Our results highlight differences in performance both among TRC tests and between TRC tests and QuaSSE under a set of simple evolutionary scenarios. Consistent with prior results (FitzJohn, 2010; Machac, 2014), we found that QuaSSE exhibits a high rate of false positives when trees contain diversification rate variation unlinked to the focal trait. QuaSSE false discovery rates were especially high in datasets containing large trees with heterogeneous diversification dynamics, such as the carnivore trees (Nyakatura & Bininda-Emonds, 2012) and the diversity-dependent multiprocess trees from Rabosky (2014). The use of more sophisticated null models is an important way forward in addressing false positives in SSE methods and in phylogenetic comparative methods generally (Beaulieu & O'Meara, 2016; Uyeda, Zenil-Ferguson, & Pennell, 2017). This approach may be possible with QuaSSE, but implementations are lacking and the computational challenges associated with fitting such models in a QuaSSE framework are expected to be non-trivial.

We found that a simulation-based test using ES (*ES-sim*) had nearly as much power as QuaSSE to detect trait dependence across trees of different sizes (Table 3) and was robust to false inferences of
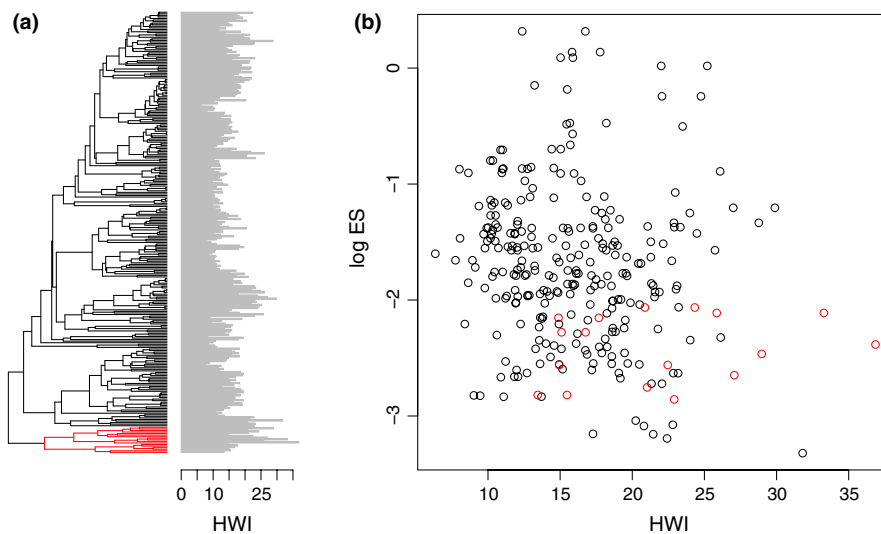

FIGURE 4 Plots of the empirical dataset from Furnariid ovenbirds. (a) The time-calibrated phylogeny of ovenbirds with a bar graph indicating the value of a morphological measure of dispersal ability (hand-wing index; HWI) for each tip. (b) A scatterplot showing the association between ES and the HWI. An association between diversification and HWI was significant based on QuaSSE analysis, but not *ES-sim*. This is likely because the simple null model used in QuaSSE failed to account for the complex diversification dynamics evident across the Furnariid tree. In addition, many of the large values of HWI were confined to one slowly diversifying clade (Sclerurinae), coloured red on the phylogeny and in the scatter plot

trait-dependent diversification across a range of evolutionary scenarios (Figure 3). The null trait-speciation associations used in *ES-sim* are simple to simulate and may be sufficiently realistic to avert false positives in many evolutionary scenarios. *ES-sim* performed better than simulation-based tests using the other speciation rate metrics we considered, ND and TB. TB performed as well or slightly better than ES at very high rates of trait evolution, and may be preferred in analyses of rapidly evolving traits, but ES performed better across a wide range of evolutionary rates. Tests that used PGLS to evaluate significance also were less powerful than simulation-based tests, a result that bears further investigation but may be related to the fact that speciation rate metrics change in non-Brownian fashion. Even when traits were simulated using non-Brownian models, we found that *ES-sim* with Brownian motion simulations had roughly equivalent power to an alternative approach where the true trait evolution model (OU) was used to construct the null distribution (Table 2). This suggests that, like FiSSE for discrete characters (Rabosky & Goldberg, 2017), *ES-sim* may be reasonably robust to model misspecification in terms of statistical power as well as false discovery rates (Figure 3).

ES-sim is a powerful test because it incorporates relatively fine-scale variation in speciation rates across phylogenies. It is therefore useful in small trees with few dramatic diversification rate shifts, in contrast to methods like STRAPP (Rabosky & Huang, 2016). However, the sensitivity of *ES-sim* needs to be taken into account in empirical studies, and researchers should evaluate the effect size as well as significance of their results. Effect size in a test like *ES-sim* could correspond either to the amount of variance in speciation rate explained by trait variation (i.e. the spread of points away from the correlation line), or the magnitude of the difference in speciation rates between lineages with the minimum and maximum trait values (the slope of the correlation line). Although the Pearson's correlation from *ES-sim* does not account for covariance between closely related species, it does provide an index of the amount of variance in speciation rate that might be explained by variation in the trait of interest. The slope of a linear model fit to the data can provide an index of the magnitude of the change in speciation rates across the observed range of trait values. We encourage researchers to report both the variance explained by the trait of interest and the slope of the correlation, as we did for the Furnariid dataset. Plotting the relationship between a trait and tip rates can also provide informal but useful insights into effect size. Moreover, sensitivity tests can provide quantitative information about the robustness of results to stochastic noise, measurement error and the impact of phylogenetic pseudoreplication. Moving forward, it would be useful to develop formal measures of trait-diversification effect size that estimate the change in species richness—or potentially, the among-clade variance in richness—that is attributable to the correlation with traits. Such a metric could compare the magnitude of the observed difference in species richness to that which would be present if the clade evolved in the absence of a relationship between traits and diversification rates.

In our empirical analysis, we found that the relationship between the hand-wing index (HWI), a measure of dispersal ability, and speciation rate in Furnariid ovenbirds identified using QuaSSE (Claramunt et al., 2012) was not supported by *ES-sim*. However, this result does not conclusively reject an association between HWI and speciation in this group. The best-fit model found by Claramunt et al. (2012) included a sigmoidal relationship between HWI and speciation, but we tested only for a linear relationship between speciation and traits using *ES-sim* and may have failed to capture a more complex relationship. QuaSSE has higher power than *ES-sim* based on simulations, and it is possible our non-significant *ES-sim* result simply reflects inadequate power. Researchers should generally be wary of over-interpretation when TRC tests reveal a negative result. Even a strong causal relationship between traits and speciation rates could be difficult to detect with TRC methods if there is insufficient replication across the phylogeny. Nonetheless, there is no clear visual signal of a relationship between HWI and ES in Furnariids (Figure 4b). Independent evidence supports the association between high dispersal ability and limited divergence in birds (Burney & Brumfield, 2009; Salisbury, Seddon, Cooney, & Tobias, 2012; Weeks & Claramunt, 2014), but additional study will surely reveal a more nuanced understanding of their association and interactions with other predictors.

The methods examined in this study are amenable to modification and extension. *ES-sim* can readily accommodate missing trait information. The method can even be used with sparsely sampled trait data across a tree, provided the sample reflects the spectrum of trait variation across the phylogeny as a whole. However, the estimation of tip-specific speciation rates will be biased by incomplete taxon sampling. For phylogenies with substantial and/or non-random missing taxa, we suggest that researchers estimate speciation rates from distributions of phylogenies where the unsampled species have been placed on the tree according to constraints, but integrating over possible placements of the unsampled lineages (e.g. Kuhn, Mooers, & Thomas, 2011; Thomas et al., 2013). The trait values for these unsampled taxa should not be included in the analyses, due to biases in the rate of trait evolution that can emerge when unsampled species are placed randomly on trees with respect to trait values (Rabosky, 2015).

TRC methods could also be devised that allow for nonlinear relationships between traits and diversification and, potentially, multiple predictor variables. In the present article, we assessed the performance of *ES-sim* only under scenarios where speciation rates are a strict linear function of the underlying traits. However, we should be clear that there are many potential functional relationships between speciation rate and phenotypes, including unimodal (hump) functions, logistic/threshold functions, step functions and others. As noted above for the Furnariids, QuaSSE can already accommodate sigmoidal and other potential relationships. *ES-sim* could also be modified to fit nonlinear models to datasets and incorporate different test statistics, for example the absolute difference between the upper and lower limits in a sigmoid function, to assess significance. We expect that *ES-sim* will perform better for some types of relationships than others, and for some functional relationships the method may fail entirely. The interpretation of parameters from *ES-sim* may be difficult if the true evolutionary process deviates substantially from a simple linear relationship, even if the method recovers a significant relationship. These concerns provide another argument for always visualizing the relationships

between tip rates, phenotypes and fitted values; simple visual inspection may help diagnose potential problems with the analyses.

In summary, *ES-sim* provides a powerful test for trait-dependent speciation with relatively low rates of false positives. *ES-sim* is also appealing because the inverse equal splits measure provides an intuitive metric of speciation rate that is closely connected to the underlying data (e.g. the branch lengths) and lends itself to visual inspection of the trait-speciation relationship. It may be an appropriate alternative or supplement to likelihood-based state-dependent speciation-extinction analyses, particularly in datasets with heterogeneous diversification dynamics. Finally, the computational speed of *ES-sim* makes it feasible for use with very large datasets that may be computationally intractable with other methods.

ACKNOWLEDGEMENTS

AUTHORS' CONTRIBUTIONS

M.G.H. and D.L.R. developed the method and developed the simulation scenarios, M.G.H. and D.L.R. implemented the method, M.G.H. ran analyses, M.G.H. and D.L.R. wrote the manuscript. Both authors contributed critically to subsequent drafts and approved the final publication.

DATA ACCESSIBILITY

R scripts and simulated data are available on Github https://github.com/mgharvey/ES-sim, https://doi.org/10.5281/zenodo.1067144.

ORCID

Michael G. Harvey http://orcid.org/0000-0001-8050-6068

REFERENCES

Orme, D., Freckleton, R., Thomas, G., Petzoldt, T., Fritz, S., Isaac, N., & Pearse, N. (2013). *caper: Comparative analyses of phylogenetics and evolution in R*. R package Version 0.5.2 https://CRAN.R-project.org/package=caper

Barraclough, T. G., Nee, S., & Harvey, P. H. (1998). Sister-group analysis in identifying correlates of diversification. *Evolutionary Ecology*, *12*, 751–754.

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

Belmaker, J., & Jetz, W. (2015). Relative roles of ecological and energetic constraints, diversification rates, and region history on global species richness gradients. *Ecology Letters*, *18*, 563–571. https://doi.org/10.1111/ele.12438

Bokma, F. (2008). Detection of "punctuated equilibrium" by Bayesian estimation of speciation and extinction rates, ancestral character states, and rates of anagenetic and cladogenetic evolution on a molecular phylogeny. *Evolution*, *62*, 2718–2726. https://doi.org/10.1111/j.1558-5646.2008.00492.x

Bromham, L., Hua, X., & Cardillo, M. (2016). Detecting macroevolutionary self-destruction from phylogenies. *Systematic Biology*, *65*, 109–127. https://doi.org/10.1093/sysbio/syv062

Burney, C. W., & Brumfield, R. T. (2009). Ecology predicts levels of differentiation in Neotropical birds. *The American Naturalist*, *174*, 358–368. https://doi.org/10.1086/603613

Cardillo, M. (1999). Latitude and rates of diversification in birds and butterflies. *Proceedings of the Royal Society of London B: Biological Sciences*, *266*, 1221–1225. https://doi.org/10.1098/rspb.1999.0766

Claramunt, S., Derryberry, E. P., Remsen, J. V., & Brumfield, R. T. (2012). High dispersal ability inhibits speciation in a continental radiation of passerine birds. *Proceedings of the Royal Society of London B: Biological Sciences*, *279*, 1567–1574. https://doi.org/10.1098/rspb.2011.1922

Cooper, N., Thomas, G. H., Venditti, C., Meade, A., & Freckleton, R. P. (2016). A cautionary note on the use of Ornstein Uhlenbeck models in macroevolutionary studies. *Biological Journal of the Linnean Society*, *118*, 64–77. https://doi.org/10.1111/bij.12701

Diaz-Uriarte, R., & Garland, T. (1996). Testing hypotheses of correlated evolution using phylogenetically independent contrasts: Sensitivity to deviations from Brownian motion. *Systematic Biology*, *45*, 27–41. https://doi.org/10.1093/sysbio/45.1.27

Farrell, B. D., Dussourd, D. E., & Mitter, C. (1991). Escalation of plant defense: Do latex and resin canals spur plant diversification? *The American Naturalist*, *138*, 881–900. https://doi.org/10.1086/285258

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

FitzJohn, R. G. (2012). Diversitree: Comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution*, *3*, 1084–1092. https://doi.org/10.1111/j.2041-210X.2012.00234.x

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

Freckleton, R. P., Phillimore, A. B., & Pagel, M. (2008). Relating traits to diversificiation: A simple test. *The American Naturalist*, *172*, 102–115. https://doi.org/10.1086/588076

Futuyma, D. J., & Moreno, G. (1988). The evolution of ecological specialization. *Annual Review of Ecology and Systematics*, *19*, 207–233. https://doi.org/10.1146/annurev.es.19.110188.001231

Garland, T. G., Dickerman, A. W., Janis, C. M., & Jones, J. A. (1993). Phylogenetic analysis of covariance by computer simulation. *Systematic Biology*, *42*, 265–292. https://doi.org/10.1093/sysbio/42.3.265

Gittleman, J. L., & Purvis, A. (1998). Body size and species-richness in carnivores and primates. *Proceedings of the Royal Society of London B: Biological Sciences*, *265*, 113–119. https://doi.org/10.1098/rspb.1998.0271

Glazier, D. S. (1987). Energetics and taxonomic patterns of species richness. *Systematic Zoology*, *36*, 62–71. https://doi.org/10.2307/2413308

Gomes, A. C. R., Sorenson, M. D., & Cardoso, G. C. (2016). Speciation is associated with changing ornamentation rather than stronger sexual selection. *Evolution*, *70*, 2823–2838. https://doi.org/10.1111/evo.13088

Harmon, L. J., Losos, J. B., Davies, T. J., Gillespie, R. G., Gittleman, J. L., Jennings, W. B., … Mooers, A. O. (2010). Early bursts of body size and shape evolution are rare in comparative data. *Evolution*, *64*, 2385–2396.

Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E., & Challenger, W. (2008). GEIGER: Investigating evolutionary radiations. *Bioinformatics*, *24*, 129–131. https://doi.org/10.1093/bioinformatics/btm538

Harvey, M. G., Seeholzer, G. F., Smith, B. T., Rabosky, D. L., Cuervo, A. M., & Brumfield, R. T. (2017). Positive association between

population genetic differentiation and speciation rates in New World birds. *Proceedings of the National Academy of Sciences of the United States of America*, 114, 6328–6333. https://doi.org/10.1073/pnas.1617397114

Hua, X., & Bromham, L. (2016). Phylometrics: An R package for detecting macroevolutionary patterns, using phylogenetic metrics and backward tree simulation. *Methods in Ecology and Evolution*, 7, 806–810. https://doi.org/10.1111/2041-210X.12531

Huang, D., & Roy, K. (2015). The future of evolutionary diversity in reef corals. *Philosophical Transactions of the Royal Society B*, 370, 20140010. https://doi.org/10.1098/rstb.2014.0010

Jablonski, D. (2008). Species selection: Theory and data. *Annual Review of Ecology, Evolution, and Systematics*, 39, 501–524. https://doi.org/10.1146/annurev.ecolsys.39.110707.173510

Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., & Mooers, A. O. (2012). The global diversity of birds in space and time. *Nature*, 491, 444–448. https://doi.org/10.1038/nature11631

Kuhn, T. S., Mooers, A. O., & Thomas, G. H. (2011). A simple polytomy resolver for dated phylogenies. *Methods in Ecology and Evolution*, 2, 427–436. https://doi.org/10.1111/j.2041-210X.2011.00103.x

Machac, A. (2014). Detecting trait-dependent diversification under diversification slowdowns. *Evolutionary Biology*, 41, 201–211. https://doi.org/10.1007/s11692-013-9258-z

Maddison, W. P., & FitzJohn, R. G. (2015). The unsolved challenge to phylogenetic correlation tests for categorical characters. *Systematic Biology*, 64, 127–136. https://doi.org/10.1093/sysbio/syu070

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

Marzluff, J. M., & Dial, K. P. (1991). Life history correlates of taxonomic diversity. *Ecology*, 72, 428–439. https://doi.org/10.2307/2937185

Mitchell, J. S., & Rabosky, D. L. (2016). Bayesian model selection with BAMM: Effects of the model prior on the inferred number of diversification shifts. *Methods in Ecology and Evolution*, 8, 37–46.

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

Nyakatura, K., & Bininda-Emonds, O. R. P. (2012). Updating the evolutionary history of Carnivora (Mammalia): A new species-level supertree complete with divergence time estimates. *BMC Biology*, 10, 12. https://doi.org/10.1186/1741-7007-10-12

Panhuis, T. M., Butlin, R., Zuk, M., & Tregenza, T. (2001). Sexual selection and speciation. *Trends in Ecology and Evolution*, 16, 364–371. https://doi.org/10.1016/S0169-5347(01)02160-7

Paradis, E. (2005). Statistical analysis of diversification with species traits. *Evolution*, 59, 1–12. https://doi.org/10.1111/j.0014-3820.2005.tb00889.x

Pennell, M. W., FitzJohn, R. G., Cornwell, W. K., & Harmon, L. J. (2015). Model adequacy and the macroevolution of angiosperm functional traits. *The American Naturalist*, 186, E33–E50. https://doi.org/10.1086/682022

Phillimore, A. B., Freckleton, R. P., Orme, C. D. L., & Owens, I. P. F. (2006). Ecology predicts large-scale patterns of phylogenetic diversification in birds. *The American Naturalist*, 168, 220–229.

Rabosky, D. L. (2014). Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS ONE*, 9, e89543. https://doi.org/10.1371/journal.pone.0089543

Rabosky, D. L. (2015). No substitute for real data: A cautionary note on the use of phylogenies from birth-death polytomy resolvers for downstream comparative analyses. *Evolution*, 69, 3207–3216. https://doi.org/10.1111/evo.12817

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

Rabosky, D. L., & Goldberg, E. E. (2017). FiSSE: A simple non-parametric test for the effects of a binary character on lineage diversification rates. *Evolution*, 71, 1432–1442. https://doi.org/10.1111/evo.13227

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

Rabosky, D. L., Mitchell, J. S., & Chang, J. (2017). Is BAMM flawed? Theoretical and practical concerns in the analysis of multi-rate diversification models. *Systematic Biology*, 66, 477–498. https://doi.org/10.1093/sysbio/syx037

Redding, D. W., & Mooers, A. O. (2006). Incorporating evolutionary measures into conservation prioritization. *Conservation Biology*, 20, 1670–1678. https://doi.org/10.1111/j.1523-1739.2006.00555.x

Revell, L. J. (2010). Phylogenetic signal and linear regression on species data. *Methods in Ecology and Evolution*, 1, 319–329. https://doi.org/10.1111/j.2041-210X.2010.00044.x

Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, 3, 217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x

Rosenzweig, M. L. (1995). *Species diversity in space and time*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511623387

Salisbury, C. L., Seddon, N., Cooney, C. R., & Tobias, J. A. (2012). The latitudinal gradient in dispersal constraints: Ecological specialization drives diversification in tropical birds. *Ecology Letters*, 15, 847–855. https://doi.org/10.1111/j.1461-0248.2012.01806.x

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

Thomas, G. H., Hartmann, K., Jetz, W., Joy, J. B., Mimoto, A., & Mooers, A. (2013). PASTIS: An R package to facilitate phylogenetic assembly with soft taxonomic inferences. *Methods in Ecology and Evolution*, 4, 1011–1017. https://doi.org/10.1111/2041-210X.12117

Uyeda, J. C., Zenil-Ferguson, R., & Pennell, M. W. (2017). Rethinking phylogenetic comparative methods. *bioRxiv*, https://doi.org/10.1101/222729

Weeks, B. C., & Claramunt, S. (2014). Dispersal has inhibited avian diversification in Australasian archipelagos. *Proceedings of the Royal Society of London B: Biological Sciences*, 281, 20141257. https://doi.org/10.1098/rspb.2014.1257

West-Eberhard, M. J. (1983). Sexual selection, social competition, and speciation. *Quarterly Review of Biology*, 58, 155–183. https://doi.org/10.1086/413215

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Harvey MG, Rabosky DL. Continuous traits and speciation rates: Alternatives to state-dependent diversification models. *Methods Ecol Evol*. 2017;00:1–10. https://doi.org/10.1111/2041-210X.12949