Detecting the Dependence of Diversification on Multiple Traits from Phylogenetic Trees and Trait Data

Leonel Herrera-Alsina*, Paul van Els, and Rampal S. Etienne

Groningen Institute for Evolutionary Life Sciences, University of Groningen, PO Box 11103, Groningen 9700 CC, The Netherlands
Correspondence to be sent to: University of Groningen, PO Box 11103, Groningen 9700 CC, The Netherlands;
E-mail: leonelhalsina@gmail.com.

Abstract.—Species diversification may be determined by many different variables, including the traits of the diversifying lineages. The state-dependent speciation and extinction (SSE) framework contains methods to detect the dependence of diversification on these traits. For the analysis of traits with multiple states, MuSSE (multiple-states dependent speciation and extinction) was developed. However, MuSSE and other SSE models have been shown to yield false positives, because they cannot separate differential diversification rates from dependence of diversification on the observed traits. The recently introduced method HiSSE (hidden-state-dependent speciation and extinction) resolves this problem by allowing a hidden state to affect diversification rates. Unfortunately, HiSSE does not allow traits with more than two states, and, perhaps more interestingly, the simultaneous action of multiple traits on diversification. Herein, we introduce an R package (SecSSE: several examined and concealed states-dependent speciation and extinction) that combines the features of HiSSE and MuSSE to simultaneously infer state-dependent diversification across two or more examined (observed) traits or states while accounting for the role of a possible concealed (hidden) trait. Moreover, SecSSE also has improved functionality when compared with its two "parents." First, it allows for an observed trait being in two or more states simultaneously, which is useful for example when a taxon is a generalist or when the exact state is not precisely known. Second, it provides the correct likelihood when conditioned on nonextinction, which has been incorrectly implemented in HiSSE and other SSE models. To illustrate our method, we apply SecSSE to seven previous studies that used MuSSE, and find that in five out of seven cases, the conclusions drawn based on MuSSE were premature. We test with simulations whether SecSSE sacrifices statistical power to avoid the high Type I error problem of MuSSE, but we find that this is not the case: for the majority of simulations where the observed traits affect diversification, SecSSE detects this. [Evolving traits; macroevolution; phylogenetic tools; speciation rates.]

Differences in rates of speciation and extinction might be associated with differences in the diversifying species' traits (strict-sense species selection, Jablonski 2008). A trait could drive phylogenetic branching patterns if certain character states foster (or hinder) species diversification. When this is the case, lineages having such trait states will have an increased (decreased) probability of speciation, and therefore, we expect that the extant species show an over- or underrepresentation of that particular trait state. However, traits evolve over time which influences the prevalence of a character state in extant species. This calls for approaches that simultaneously model diversification and trait dynamics and their effect on diversification. Various approaches have been developed to detect traits that affect speciation and/or extinction rates from phylogenies, together with the trait distribution on the tips. The state-dependent speciation and extinction (SSE) family of models (combined in the R package "diversitree") assumes a model of trait evolution along the branches of a phylogeny and an effect of the state of a trait on speciation and extinction. For instance, for a trait that can take more than two categorical states diversitree offers MuSSE (multiple states-dependent speciation and extinction, Fitzjohn 2012). The procedure starts by assigning probabilities of each trait state based on the information on extant species' traits, and then goes rootward, along the branches and nodes using the parameters (speciation, extinction, and trait evolution rates) to compute the probability of the phylogeny, and

the tip trait data given the state at the crown of the tree. This probability can be used as the likelihood of the model given the phylogeny and tip trait data. The optimization of this computed likelihood allows us to find which parameter combination most likely explains the data.

These SSE methods generally look at variation in diversification rates across trait states and when applied to empirical data sets a significant influence of trait states on speciation has often been found (e.g., Arbuckle and Speed 2015; Marki et al. 2015). However, Maddison and Fitzjohn (2014) noted that the SSE inference framework may be flawed because of pseudoreplication and codistribution of characters: phylogenetically related species are expected to be more similar in their traits, precisely because of their evolutionary relatedness, and not necessarily because their trait value causes a radiation. Rabosky and Goldberg (2015) confirmed that indeed the traditional SSE framework has a high Type I error rate. If there is a shift in diversification rates, an SSE analysis will likely attribute it to the variation in trait states, even if the trait in question has little to do with this shift.

To resolve these problems, Rabosky and Huang (2016) developed a permutation approach where first speciation rates for each lineage are estimated using only the reconstructed phylogenetic tree, and then sets of species with similar rates are shuffled and the correlation with the trait values is calculated. Repeating this procedure several times produces a

distribution of null correlations to which the empirical correlation is compared. In this approach, trait evolution is ignored. Recently, Rabosky and Goldberg (2017) proposed a nonparametric method which considers the difference between the averages of per-tip speciation rates associated with each state (this is their test statistic). Independently, the evolution of a neutral trait is simulated (given a model) on the empirical phylogeny. This assigns trait states to the species at the tips of the phylogeny. For each simulated data set the test statistic is calculated, and the frequency distribution of these values is used for comparison with the observed test statistic in a statistical hypothesis-testing fashion. They found that their method reduces the Type I error rate, but its power to detect trait-dependent diversification is lower than that of Binary-state speciation and extinction (BiSSE). Although such nonparametric approaches are valid to look for association between a given trait in extant species and rates of diversification, one might want to jointly assess how a trait evolves over time and whether and to what extent this leads to the rise and fall of lineages. In line with the conceptual foundation of the SSE framework, Beaulieu and O'Meara (2016) proposed HiSSE (hidden-state-dependent speciation and extinction), a new method to avoid false positives. This method explicitly allows for the possibility that the process of diversification is not related to the focal trait, but rather to some unknown, hidden trait. In other words, under a traditional SSE analysis, a clade with nonhomogeneous diversification rates will be linked to the observed trait variation. In contrast, in HiSSE a hidden-state model may be selected when diversification is heterogeneous but this variation cannot be attributed to the focal trait. This does not mean that there is a single trait responsible for the observed branching patterns, but just that differences in diversification cannot solely be explained by variation in the observed trait. Beaulieu and O'Meara (2016) found that HiSSE is able to distinguish between models with and without an effect of the observed trait on diversification in most of their simulations.

Although HiSSE represents a significant improvement in the toolbox for analysis of state-dependent diversification, it only allows for a single binary observed trait, not for traits with multiple states, or multiple traits. Furthermore, the calculation of the likelihood conditional on survival used in HiSSE (and other SSE models) is not completely correct. Herein, we introduce SecSSE (several examined and concealed states-dependent speciation and extinction) that combines the features of HiSSE and MuSSE to simultaneously infer state-dependent diversification across two or more traits or states while accounting for the role of a possible concealed trait, and correctly conditioning on survival. We provide an R package of the same name. It correctly conditions on survival but the incorrect conditioning is also available in the package for comparison. Furthermore, it can handle data with missing or partial trait information for some species. For instance, consider a species for which we

know that it is in State 1 or 2, but not in State 3. This type of partial information can be specified in SecSSE. This feature is not only useful for missing or partial data but it is also suitable for assigning taxa to more than one state simultaneously (e.g., polymorphic species). We revisit seven studies where particular trait states were found to be linked to differential speciation rates using MuSSE. These studies were published before the problems with the SSE had been identified, and their findings have not yet been screened for false positives. Lastly, we rigorously assess the Types I and II error rates of SecSSE by simulations.

MATERIALS AND METHODS

SecSSE

SecSSE uses the same framework as HiSSE, but allows for multiple examined (observed) and multiple concealed (hidden) states. Coupled differential equations describe the dynamics of the probability of the tipward part of the tree (D_{ij}) and the probability of lineages going extinct (E_{ij}) at time t given that the system is in state ij at time t for all possible examined states i and concealed states j. This means that the user has to specify speciation and extinction rates for all possible states, and a transition rate matrix for all possible transitions between these states.

SecSSE differs from the diversitree models and HiSSE in the way conditioning on nonextinction of the crown lineages is computed. Diversitree/HiSSE calculates the D_i and E_i for every trait state i at the root. Then the unconditional likelihood L is computed as a weighted average over the D_i values:

$$L = \sum_i D_i w_i, \qquad (1)$$

where weights w_i are typically the (normalized) D_i themselves, that is,

$$w_i = \frac{D_i}{\sum_i D_i}. \qquad (2)$$

The likelihood conditioned on nonextinction of the crown lineages according to diversitree/HiSSE is obtained by dividing the unconditional likelihood, Eq. (1), by the weighted sum of nonextinction probabilities multiplied by the speciation rate at the root (because we are looking at a tree, Nee et al. 1994):

$$L = \frac{\sum_i D_i w_i}{\sum_i w_i \lambda_i (1 - E_i)^2}. \qquad (3)$$

In other words, the likelihood is first summed, with weights, over the various traits and then conditioned. However, we argue that the conditioning should occur first, and then the weighted sum over the various traits should be taken:

$$L = \sum_i \frac{D_i w_i}{\lambda_i (1 - E_i)^2}. \qquad (4)$$

We consider that the diversitree/HiSSE implementation of conditioning is incorrect, because it effectively conditions each probability for state i on nonextinction of the process weighted across all root states, whereas it should condition each probability for state i on nonextinction given state i at the root. In other words, we first want to compute the full conditional likelihood given that the trait at the root is in state i, and then apply the weights we want to attach to each root state. These weights can be freely chosen by the user, depending on how much confidence one has about the root state. In Bayesian terms, the weights can be regarded as the prior distribution over the trait states. As stated above, diversitree/HiSSE uses the normalized D_i themselves as weights (Eq. (2)) and hence solely bases the weights on the data. This is similar to how posterior Bayes factors are defined rather than integrating the likelihood over the prior (which is how conventional Bayes factors are defined), posterior Bayes factors integrate over the posterior (Aitkin 1991). In the same spirit, one could argue that the weights should be

$$w_i = \frac{\frac{D_i}{\lambda_i (1 - E_i)^2}}{\sum_i \frac{D_i}{\lambda_i (1 - E_i)^2}}. \qquad (5)$$

In our applications we used the likelihood of Eq. (4) with weights as in Eq. (2). The R package SecSSE also offers the option to use the likelihood of Eqs. (1) and (3), or the weights of Eq. (5).

In SecSSE, uncertainty about the exact state of the species is handled at the start of the likelihood calculation when the initial trait state probabilities D_{ij} are assigned (i.e., at the tree tips). Consider a trait that can take three possible examined states (1, 2, and 3). When we are certain that the examined state of a given species is 2, we assign the initial probabilities of 0, 1, 0 to the three states, respectively. For a species in examined State 2 or 3, the probabilities will be 0, 1, 1, respectively. Finally, when no information is available, the initial probabilities of the states will be 1, 1, 1. These configurations of initial probabilities at the tips are fully consistent with the likelihood calculation. Note that they do not have to add up to 1, because they are not probabilities for the same event; $D_{ij}(t)$ is the probability of the tipward part of the tree including the trait values at the tips given state ij at time t, so if the trait value at the tips is either 2 or 3, then $D_2 = 1$ and $D_3 = 1$. If we have multiple concealed states, then each of them receives the same probability, for example, in the latter case where the examined trait value is either 2 or 3 and there are three concealed states, we have $D_{21} = D_{22} = D_{23} = D_{31} = D_{32} = D_{33} = 1$. The D_{ij} and E_{ij} equations are numerically integrated backwards (from tips to the root), analogous to implementations of other SSE models.

Revisiting MuSSE Analyses with SecSSE

We searched for studies where trait-dependent diversification was analyzed with MuSSE, with potentially elevated Type I errors, using Web of Science and Google Scholar search terms "MuSSE" and "diversification." We used phylogenetic and trait information collected either from the supplementary materials of those papers or provided by the authors upon request. We discarded publications for which the data were incomplete or analyses were not clearly replicable. For instance, some studies removed species from the phylogenetic reconstruction in a rather arbitrary way (e.g., species absent in their geographic scope). We obtained seven complete data sets, all with three examined states (Table 1). These seven studies cover very different taxonomic groups (resulting in highly variable crown ages, 34–269 Myr.), including ants (Insecta: Formicidae, Burchill and Moreau 2016, number of tips $n = 116$), pupfish (Actinoperygii: Cyprinodontidae, Helmstetter et al. 2016, $n = 114$), corvoid birds (Aves: Corvoidea, Marki et al. 2015, $n = 763$), ruminants (Mammalia: Ruminantia, Cantalapiedra et al. 2013, $n = 197$), epiphytic ferns (Plantae: Polypodiaceae, Sundue et al. 2015, $n = 417$), grunters (Teleostei: Terapontidae, Davis et al. 2016, $n = 38$), and amphibians (Amphibia, Arbuckle and Speed 2015, $n = 2683$). State-dependent diversification rates were reported in all seven studies. In most studies, differential extinction and transition rates were not examined or were estimated to be very low (in the case of extinction), so we assumed a single value for the transition rates and a single, trait-independent, value for the extinction rates. We adopted the model settings of

TABLE 1.   Revisited studies where trait-dependent diversification was found

Study	Group	Clade age (my)	Trait	Best supported model
Burchill and Moreau (2016)	Ants	146	Colony size	CTD
Helmstetter et al. (2016)	Cyprinodontiform fishes	89	Life-history strategy	ETD
Marki et al. (2015)	Corvoids	38	Breeding system	CTD
Cantalapiedra et al. (2013)	Ruminants	55	Feeding mode	CTD
Sundue et al. (2015)	Epiphytes	42	Leaf shape	CTD
Davis et al. (2016)	Terapontid fishes	34	Feeding system	CR
Arbuckle and Speed (2015)	Amphibians	269	Coloration	CTD

Note: Phylogenetic and trait information were reanalyzed with SecSSE to compare three different models of diversification: CR, CTD, and ETD. The last column shows the best supported model (according to AIC).
CR = constant-rate; CTD = concealed trait-dependent; ETD = examined trait-dependent.

the original MuSSE analysis as much as possible, which involved the use of the sampling fraction reported in four studies (Arbuckle and Speed 2015; Burchill and Moreau 2016; Davis et al. 2016; Helmstetter et al. 2016). In the case of Helmstetter et al. (2016), we included all taxa present in the tree but missing from the trait data file and assigned them the value NA (equal chance of inclusion in all states), rather than removing them from the tree as done in the original study. In the case of Arbuckle and Speed (2015), we chose to assign conspicuously colored taxa with unknown function to both the conspicuously colored and the camouflaged trait states (but not the polymorphic state) rather than creating a separate fourth state for them as in the original study. In both cases, we report results based on the original methodology in Table S1, Supplementary Material available on Dryad at https://doi.org/10.5061/dryad.qf3g0.

We used the convention introduced in HiSSE to denote observed states with numbers and hidden/concealed states with letters. For example λ_{1A} denotes the speciation rate when the lineage is in a state with observed trait value 1 and concealed trait value A. We defined three models with different settings for the speciation rates. In the constant-rates model (CR, hereafter), all species have the same speciation rate λ regardless of their trait state ($\lambda_{1A} = \lambda_{2A} = \lambda_{3A} = \lambda_{1B} = \lambda_{2B} = \lambda_{3B} = \lambda_{1C} = \lambda_{2C} = \lambda_{3C}$), equivalent to a regular CR birth–death model. In the concealed trait-dependent (CTD) speciation model, speciation rates are allowed to vary only between concealed states ($\lambda_{1A} = \lambda_{2A} = \lambda_{3A} \neq \lambda_{1B} = \lambda_{2B} = \lambda_{3B} \neq \lambda_{1C} = \lambda_{2C} = \lambda_{3C}$). In the examined trait-dependent (ETD) speciation model, speciation rates are allowed to vary only between the examined states ($\lambda_{1A} = \lambda_{1B} = \lambda_{1C} \neq \lambda_{2A} = \lambda_{2B} = \lambda_{2C} \neq \lambda_{3A} = \lambda_{3B} = \lambda_{3C}$). Additionally, we used a model with two and four concealed states, where we allowed speciation rates to vary between concealed states but not the examined states that is, a CTD model but with two concealed states ($\lambda_{1A} = \lambda_{2A} = \lambda_{3A} \neq \lambda_{1B} = \lambda_{2B} = \lambda_{3B}$) or four concealed states ($\lambda_{1A} = \lambda_{2A} = \lambda_{3A} \neq \lambda_{1B} = \lambda_{2B} = \lambda_{3B} \neq \lambda_{1C} = \lambda_{2C} = \lambda_{3C} \neq \lambda_{1D} = \lambda_{2D} = \lambda_{3D}$). We distinguished between the CTD models by labelling them as CTD2, CTD3, and CTD4 where the numeral (2, 3, or 4) refers to the number of concealed states. Note that the CTD3 model has the same number of parameters as ETD, which is analogous to the CID-2 model in HiSSE (Beaulieu and O'Meara 2016).

The Maximum Likelihood search was performed with the simplex algorithm offered by the DDD package (SecSSE also allows for the subplex algorithm from the subplex package). An example of the likelihood landscape can be found in Fig. S2, Supplementary Material available on Dryad. To avoid only finding a local optimum, we used five sets of initial parameters: one starting point came from the standard birth–death model, one in which we halved these values, one in which we doubled these values, one in which we only changed μ to 0.02, and one in which we used the mean of the estimated speciation rates reported in the original studies. Finally, we performed model

selection based on Akaike weights (Wagenmakers and Farrell 2004), to select the best-performing model for each study.

Performance of SecSSE Analysis

To assess the robustness of our conclusions on model selection, we carried out a three-step procedure for each empirical data set. 1) Using the parameters that maximize the likelihood for each of the three models (CR, CTD, and ETD) we simulated 100 data sets (phylogenies and traits; details below) that are structurally similar to the empirical data (Supplementary Material available on Dryad). 2) We ran SecSSE analyses under the three models for each simulated data set. 3) Finally, we compared models and counted the instances where the generating model was selected as the best model according to AIC.

For the simulations, we considered a trait with three examined states and three concealed states, as used for the parameter inference on empirical data (i.e., 1, 2, 3 and A, B, C, respectively), leading to a nine-state system (i.e., 1A, 2A, 3A, 1B, 2B, 3B, 1C, 2C, 3C). The simulation starts with two species with the same trait state; this state can shift to the other states at a transition rate q. Our simulations were conditioned on survival of these two crown species; if one of the species becomes extinct, the simulation was discarded. Speciation and extinction rates depend on trait states, and similar to BiSSE, HiSSE, and MuSSE, the switch from one state another does not cause immediate speciation. Moreover, after a speciation event both daughter species inherit the trait state from the parental species. We performed the simulation for a period of time equal to the crown age of the empirical data. After the simulation, we reconstructed the phylogeny from the record of speciation and extinction events. Then, we relabeled trait states to merge them into a trait with only three states. The traits 1A, 1B, and 1C were coded as 1; 2A, 2B, and 2C were coded as 2, whereas 3A, 3B, and 3C were coded as 3. This is analogous to Beaulieu and O'Meara (2016). We performed SecSSE inference on the simulated data using the same models as for the empirical data. We excluded the data set of Helmstetter et al. (2016) from our performance analysis because their phylogeny was reconstructed at the genus level and because their sampling fraction was very low (10%). We also had to exclude the Arbuckle and Speed (2015) study, because the high estimates of the speciation rate for some trait states (Table 1) combined with a very old crown age (Amphibia ~269 Myr.) caused our simulations to sometimes produce trees that were so large that they were computationally unmanageable.

We also explored the performance of SecSSE with several other parameter combinations (λ from 0.05 to 0.6; μ from 0 to 0.1; $q = 0.05$ and 0.1; see Table 3). These settings were chosen because they produce a realistic number of species and trait distributions across the tree.

We ran five sets of simulations where speciation or extinction rates varied across the states of a trait with either three or five states. Regarding transition rates, we simulated data sets using four parameter sets: both concealed and examined traits had 1) low or 2) high transition rate between states, 3) transition rates between the states of the concealed trait were higher than the rates in the examined trait, and 4) the examined trait had higher transition rates than the concealed trait. We simulated 100 data sets for each parameter set. Finally, we ran a SecSSE analysis on all these data sets, again with the three models we described above, and compared them using AIC in order to evaluate how often the generating model is preferred over the other two.

To explore how well SecSSE performs in model selection when the data is not generated by any of the models implemented in SecSSE, we simulated a scenario where the branching pattern of the clade is completely independent of a trait value (hereafter, heterogeneous rate across lineages, HRL). We simply randomly drew a lineage and increased or decreased its speciation rate at random by a factor of 3 (i.e., shift rate event). Its descendants still kept the old speciation rate. This potentially leads to overrepresentation or underrepresentation of a trait value, but has nothing to do with a (n inherited) trait affecting diversification. We ran two sets of simulations of HRL with either 6 or 12 shift rates events. We applied SecSSE to the simulated tree and trait distributions, and compared which model (CR, CTD, and ETD) fitted better.

For the weights of the root states, we used the observed probability approach (Eq. (2), Maddison et al. 2007) as our default. However, because the choice of root state can influence the likelihood calculation (Goldberg and Igiić 2008), we repeated the analyses on the empirical data sets using different methods to assigns root weights: 1) equal weights to all states, 2) probability 1 to examined State 1 (i.e., assuming we know that the root had state 1), 3) probability 1 to examined State 2, and 4) probability 1 to examined State 3 (Fig. S4, Supplementary Material available on Dryad). Additionally, we repeated our maximum likelihood analyses using Maddison et al. (2007) conditioning on survival for simulated data sets 1 to 6 and compared the parameter estimates to the estimates under our conditioning.

We also analyzed scenarios with partial information on trait states. We considered a three-state trait (States 1, 2, and 3) in which two states (1 and 3) are difficult to tell apart. We used simulated data from Set 5 (see Table 3) to assess the performance in model comparison under three scenarios. In the no information scenario, we deleted trait information of 80% of the species in either State 1 or 3 (i.e., the trait state was replaced by NA). In the partial information scenario, we set the tip-state of these species to "State 1 or 3" (which means that we effectively add information that it was not in State 2). Finally, in the full information scenario the full trait information was used for the analysis.

We explored the behavior of SecSSE with models of different complexity by fitting CTD and ETD models with 2, 3, and 4 concealed states for 100 simulated data sets (using Set 5, see Table 3).

RESULTS

Empirical Data Sets

We detected a significant effect of a concealed trait rather than the examined trait in five out of seven empirical data sets (Tables 1 and 2). Speciation rates in models without concealed state were identical or similar to rates reported in the original studies, except for the two cases where we had to adjust model settings making diversification rates incomparable (Arbuckle and Speed 2015; Helmstetter et al. 2016; see Methods Section). Extinction rates were in most cases close to zero. We will now discuss the results for each data set in more detail.

Burchill and Moreau (2016) focused mainly on transition rates between three different colony sizes (low, medium, and high) across 400 ant species in 118 genera. They also provided estimates of colony size-dependent speciation rates and found that varied approximately three-fold across the three size categories ($\lambda = 0.01 - 0.03$). Their results indicate that small colonies have the lowest rates, and medium colonies the highest. Our reanalysis suggests that there is no significant evidence for colony size being responsible for speciation rate variation.

Helmstetter et al. (2016) found support for the idea that diversification rates in viviparous cyprinodontiform fishes are about three-fold higher than in their oviparous annual or nonannual counterparts. These diversification rates are mainly driven by differences in speciation rates, as extinction rates do not differ between the three groups. Our reanalysis indicates that the model with variation between examined states was most likely, supporting these findings.

In the case of corvoid birds, Marki et al. (2015) results point toward increased net diversification rates (again, mainly as a result of elevated speciation rates, because extinction rates are similar across states) in pair-breeding birds, relative to co-operatively breeding or uniparentally breeding birds. Although their most likely model is one including free transition rates (and fixed extinction rates), they also ran a model including one fixed transition rate (as we did) and still found that speciation rates are higher in pair-breeders. Speciation rates in the model including free transition rates are only slightly higher (0.128) than in the model with fixed transitions (0.125). Our reanalysis finds no support for an influence of breeding system on speciation rates, because a model with variation across concealed states is most likely.

Cantalapiedra et al. (2013) used three different models to explore the relationship between speciation rates and ruminant feeding mode (browser, grazer, and a mix between both), allowing variability in transition rates. They found that ruminants with mixed feeding mode had the highest speciation rate, followed by grazers

TABLE 2. Estimates of speciation rate (λ), extinction rate (μ), and the rate of transition between trait states (q) obtained by applying the ML framework of SecSSE for four different models (CR; CTD with two and three concealed states; ETD diversification) to existing multistate data sets

Ref.	Models	k	ML	AICw	λ1A	λ2A	λ3A	λ1B	λ2B	λ3B	λ1C	λ2C	λ3C	λ1D	λ2D	λ3D	μ	q
1	CR	3	−678.437	0.00001	0.0213	0.0213	0.0213	0.0213	0.0213	0.0213	0.0213	0.0213	0.0213	0.0213	0.0213	0.0213	0.00001	0.00681
	CTD2	4	−669.953	0.01962	0.0269	0.0269	0.0269	0.0001	0.0001	0.0001	—	—	—	—	—	—	0.00003	0.00901
	CTD3	**5**	**−665.135**	**0.89243**	**0.0001**	**0.0001**	**0.0001**	**0.0001**	**0.0001**	**0.0001**	**0.0317**	**0.0317**	**0.0317**	—	—	—	**0.00020**	**0.00869**
	CTD4	6	−666.462	0.08711	0.0244	0.0244	0.0244	0.0343	0.0343	0.0343	0.0001	0.0001	0.0001	0.0000	0.0000	0.0000	0.00001	0.00753
	ETD	5	−672.113	0.00083	0.0070	0.0019	0.0019	0.0070	0.0019	0.0019	0.0070	0.0019	0.0019	—	—	—	0.00001	0.01357
2	CR	3	−503.181	0.00001	0.2454	0.2454	0.2454	0.2454	0.2454	0.2454	0.2454	0.2454	0.2454	0.2454	0.2454	0.2454	0.17618	0.00161
	CTD2	4	−492.183	0.00007	0.0868	0.0868	0.0868	0.2051	0.2051	0.2051	—	—	—	—	—	—	0.00681	0.00162
	CTD3	5	−487.080	0.00398	0.2096	0.2096	0.2096	0.0386	0.0386	0.0386	0.0969	0.0969	0.0969	—	—	—	0.00001	0.00161
	CTD4	6	−489.629	0.00011	0.2294	0.2294	0.2294	0.0346	0.0346	0.0346	0.0939	0.0939	0.0939	0.0847	0.0847	0.0847	0.00001	0.00176
	ETD	**5**	**−481.558**	**0.99584**	**0.0772**	**0.1375**	**0.2084**	**0.0772**	**0.1375**	**0.2084**	**0.0772**	**0.1375**	**0.2084**	—	—	—	**0.00001**	**0.00105**
3	CR	3	−2611.514	0.00001	0.1185	0.1185	0.1185	0.1185	0.1185	0.1185	0.1185	0.1185	0.1185	0.1185	0.1185	0.1185	0.00001	0.00452
	CTD2	4	−2579.453	0.00001	0.2249	0.2249	0.2249	0.0928	0.0928	0.0928	—	—	—	—	—	—	0.00001	0.00446
	CTD3	5	−2569.565	0.00055	0.2283	0.2283	0.2283	0.0994	0.0994	0.0994	0.0001	0.0001	0.0001	—	—	—	0.00001	0.00513
	CTD4	**6**	**−2561.069**	**0.99944**	**0.0051**	**0.0051**	**0.0051**	**0.0036**	**0.0036**	**0.0036**	**0.2338**	**0.2338**	**0.2338**	**0.1057**	**0.1057**	**0.1057**	**0.00001**	**0.00563**
	ETD	5	−2605.539	0.00001	0.1477	0.1244	0.0867	0.1477	0.1244	0.0867	0.1477	0.1244	0.0867	—	—	—	0.00001	0.00451
4	CR	3	−738.191	0.00019	0.1117	0.1117	0.1117	0.1117	0.1117	0.1117	0.1117	0.1117	0.1117	0.1117	0.1117	0.1117	0.00001	0.00979
	CTD2	**4**	**−729.622**	**0.36000**	**0.1324**	**0.1324**	**0.1324**	**0.0274**	**0.0274**	**0.0274**	—	—	—	—	—	—	**0.00001**	**0.00991**
	CTD3	5	−728.798	0.30181	0.1387	0.1387	0.1387	0.0274	0.0274	0.0274	0.0274	0.0274	0.0274	—	—	—	0.00001	0.01002
	CTD4	6	−727.687	0.33708	0.0243	0.0243	0.0243	0.0242	0.0242	0.0242	0.0244	0.0244	0.0244	0.1446	0.1446	0.1446	0.00001	0.00983
	ETD	5	−734.599	0.00091	0.0846	0.1103	0.1103	0.0846	0.1103	0.1103	0.0846	0.1103	0.1103	—	—	—	0.00001	0.00969
5	CR	3	−1430.800	0.00022	0.1372	0.1372	0.1372	0.1372	0.1372	0.1372	0.1372	0.1372	0.1372	0.1372	0.1372	0.1372	0.00001	0.00838
	CTD2	4	−1422.814	0.23683	0.1162	0.1162	0.1162	0.2361	0.2361	0.2361	—	—	—	—	—	—	0.00001	0.00774
	CTD3	**5**	**−1420.864**	**0.61238**	**0.2444**	**0.2444**	**0.2444**	**0.1251**	**0.1251**	**0.1251**	**0.2767**	**0.2767**	**0.2767**	—	—	—	**0.00001**	**0.00785**
	CTD4	6	−1421.272	0.14984	0.0617	0.0617	0.0617	0.1302	0.1302	0.1302	0.1419	0.1403	0.0526	0.1293	0.1293	0.1293	0.00001	0.00768
	ETD	5	−1427.596	0.00073	0.1419	0.1318	0.1372	0.1419	0.1318	0.1372	0.1419	0.1318	0.1372	—	—	—	0.00001	0.00854
6	**CR**	**3**	**−163.708**	**0.34561**	**0.1222**	**0.1222**	**0.1222**	**0.1222**	**0.1222**	**0.1222**	**0.1222**	**0.1222**	**0.1222**	**0.1222**	**0.1222**	**0.1222**	**0.04376**	**0.01907**
	CTD2	4	−163.211	0.20904	0.0001	0.0001	0.0001	0.1150	0.1150	0.1150	—	—	—	—	—	—	0.00319	0.01926
	CTD3	5	−162.889	0.10617	0.1204	0.1204	0.1204	0.1199	0.1199	0.1199	0.0001	0.0001	0.0001	—	—	—	0.01770	0.01884
	CTD4	6	−163.446	0.02237	0.1232	0.1232	0.1232	0.1226	0.1226	0.1226	0.0000	0.0000	0.0000	0.1230	0.1230	0.1230	0.02205	0.01842
	ETD	5	−161.795	0.31681	0.0716	0.1301	0.1599	0.0716	0.1301	0.1599	0.0716	0.1301	0.1599	—	—	—	0.00001	0.01658
7	CR	3	−12,400.770	0.00001	0.0676	0.0676	0.0676	0.0676	0.0676	0.0676	0.0676	0.0676	0.0676	0.0676	0.0676	0.0676	0.01453	0.00110
	CTD2	4	−12,203.330	0.00001	0.0814	0.0814	0.0814	0.0401	0.0401	0.0401	—	—	—	—	—	—	0.00001	0.0161
	CTD3	5	−12,167.480	0.00003	0.0832	0.0832	0.0832	0.0169	0.0169	0.0169	0.0455	0.0455	0.0455	—	—	—	0.00001	0.00123
	CTD4	**6**	**−12,156.120**	**0.99997**	**0.0165**	**0.0165**	**0.0165**	**0.0838**	**0.0838**	**0.0838**	**0.0453**	**0.0453**	**0.0453**	**0.0452**	**0.0452**	**0.0452**	**0.00001**	**0.00124**
	ETD	5	−12,378.910	0.00001	0.0067	0.0276	0.0799	0.0067	0.0276	0.0799	0.0067	0.0276	0.0799	—	—	—	0.00878	0.00137

Note: Models are compared using AICw. The references to the data sets are as follows: 1) Burchill and Moreau (2016), 2) Helmstetter et al. (2016), 3) Marki et al. (2015), 4) Cantalapiedra et al. (2013), 5) Sundue et al. (2015), 6) Davis et al. (2016), and 7) Arbuckle and Speed (2015). Those bold values show the model selected as the best one; it does not have associated any statistical significance value.

CR = constant-rate; CTD = concealed trait-dependent; ETD = examined trait-dependent; ML = maximum likelihood; AICw = Akaike weights.

and browsers. We did not find feeding mode-dependent speciation rates. Allowing variation in transition rates might accommodate a shift in speciation rates. However, we expect that including free transition rates in our analyses will not influence our inferred speciation rates, because our estimated speciation rates in the CTD model are nearly identical to those of Cantalapiedra et al. (2013) who did allow variability in transition rates.

Sundue et al. (2015) ran a number of analyses on an epiphyte data set, to study whether diversification rates differ between plants with simple, 1-pinnate or 2-3-pinnate leaf patterns. We replicated their three-state analysis with SecSSE. The authors found that 1-pinnate plants are associated with higher net diversification rates than simple (next highest rates), and 2-3-pinnate plants. Although Sundue et al. (2015) tested for susceptibility to Type I error rates in their data set using a neutral, simulated trait as suggested by Rabosky and Goldberg (2015), they did not detect spurious effects of leaf shape on diversification patterns. However, we find that a model with variable concealed states is most likely.

Davis et al. (2016) concluded that in terapontid fish, herbivores have considerably higher net diversification rates than either omnivores or carnivores. Our reanalysis with SecSSE points toward a model with one speciation rate across all states; Davis et al. (2016) acknowledged that MCMC plots of their rates overlap considerably due to power issues, which is in line with our findings.

Finally, Arbuckle and Speed (2015) investigated the influence of coloration on amphibian diversification. They found higher net diversification rates for conspicuously colored taxa than for cryptically colored or polymorphic taxa (both have similar rates). Extinction is close to zero across states. We found no significant difference in rates across examined states with SecSSE. While Arbuckle and Speed (2015) assigned a fourth state to taxa with uncertain coloration strategies (multiple strategies used by one species), we kept the three-state set-up and used our new feature (trait state uncertainty) to incorporate this partial information. Nevertheless, we believe that this does not influence the outcome of the SecSSE analysis, because taxa with uncertain coloration strategies represented only 2% of total taxa in the tree ($n = 2683$).

Overall, our confidence in the new conclusions for the seven data sets is high: the results of our simulation analysis show that when SecSSE suggests that trait state is not related to diversification dynamics and a CR model is not appropriate either (i.e., when the CTD model is selected as the best one), this is indeed the case for 96% of the simulations for Study 1, 98.9% for Study 3, 98.8% for Study 4, and 96.6 for Study 5 (Fig. 1). In general, we found that our corrected conditioning on survival yields higher likelihood values than the conditioning used in previous SSE models, but we note that these likelihoods should not be compared to select between the two conditionings. Furthermore, we found that the various conditioning methods produce similar (but not identical) parameter estimates (Fig. S1, Supplementary Material available on Dryad).

Robustness of SecSSE Analysis

In general, our simulation results show that SecSSE is reasonably robust to Type I error. In the robustness analysis for the empirical studies, we found that SecSSE incorrectly suggested that speciation depends on the examined trait between 8 and 15% of all cases (Fig. 1). In other words, for 8–15 out of 100 simulated data sets, SecSSE picked the ETD model as the best performing when ETD was not in fact the generating model. This relatively low Type I error is not accompanied by low power to detect trait-dependent diversification: when the generating model has dependence on the examined trait, it is correctly chosen in on average 84.6% of the cases (Fig. 1).

We now report the analyses on simulations with various parameter settings (Table 3). When speciation is variable between states, the least favorable results are obtained when extinction is elevated ($\mu = 0.1$) or the difference between speciation rates is rather small (parameter Sets 6 and 5, respectively). However, when extinction is zero and the difference in speciation rates is higher (Set 1), accuracy increases: the generating CTD model is correctly selected in the 72% of the cases. In this same set of simulations, SecSSE correctly selected trait-dependent diversification in 93% of all the cases when the ETD was the generating model. Similar to the analysis for empirical data, we found that SecSSE erroneously selects examined trait-dependent diversification as the most likely scenario for 14% of the simulated data sets. In Set 7, where the trait states had the same speciation rate but differed in the extinction rate (i.e., trait-dependent extinction) SecSSE selected ETD as the most likely in 36% of simulations where ETD was the generating model. Out of the 57 cases where ETD was chosen, there were 21 cases where ETD was not the generating model (63% accuracy). For the simulations with low transition rate ($q = 0.05$), SecSSE had a similar performance as when the transition was 0.1 (Table 3). In the simulated data sets, where the concealed trait had higher transition rates than the examined trait, SecSSE performed slightly better, in terms of both Types I and II errors, than when the examined trait had the higher transition rate. We also report that the accuracy of SecSSE to recover the simulated parameters is high (Fig. 2).

For our model of HRL, we found that between 18 and 26% of the times SecSSE incorrectly inferred an ETD model as the best fit. We found worse performance when more shift-rate events took place. Moreover, the rate at which state shift (q) occurred does not affect the results (Table 4).

We found that by incorporating partial information on a certain trait value, the accuracy of parameter estimates and performance in model selection is better than when there is no information at all on that trait value. In the no information scenario, SecSSE picked

TABLE 3.   Performance of SecSSE in model selection using different rates of speciation, extinction, and transition

Generating models	Parameters	Model used for inference		
		CR	CTD	ETD
Set 1				
CR	$\lambda=0.25;\mu=0;q=0.1$	**84**	2	14
CTD	$\lambda=0.05,0.2,0.5;\mu=0;q=0.1$	17	**72**	11
ETD	$\lambda=0.05,0.2,0.5;\mu=0;q=0.1$	7	0	**93**
Set 2				
CR	$\lambda=0.25;\mu=0;q=0.05$	**89**	0	11
CTD	$\lambda=0.05,0.2,0.5;\mu=0;q=0.05$	16	**72**	12
ETD	$\lambda=0.05,0.2,0.5;\mu=0;q=0.05$	3	0	**97**
Set 3				
CR	$\lambda=0.25;\mu=0;q_{conc}=0.1,q_{exam}=0.05$	**90**	1	9
CTD	$\lambda=0.05,0.2,0.5;\mu=0;q_{conc}=0.1,q_{exam}=0.05$	25	**56**	19
ETD	$\lambda=0.05,0.2,0.5;\mu=0;q_{conc}=0.1,q_{exam}=0.05$	9	1	**90**
Set 4				
CR	$\lambda=0.25;\mu=0;q_{conc}=0.05,q_{exam}=0.1$	**89**	1	10
CTD	$\lambda=0.05,0.2,0.5;\mu=0;q_{conc}=0.05,q_{exam}=0.1$	18	**68**	14
ETD	$\lambda=0.05,0.2,0.5;\mu=0;q_{conc}=0.05,q_{exam}=0.1$	13	0	**87**
Set 5				
CR	$\lambda=0.25;\mu=0$	**80**	2	18
CTD	$\lambda=0.25,0.1,0.4;\mu=0$	62	**19**	19
ETD	$\lambda=0.25,0.1,0.4;\mu=0$	23	0	**77**
Set 6				
CR	$\lambda=0.35;\mu=0.1$	**81**	0	19
CTD	$\lambda=0.45,0.2,0.5;\mu=0.1$	65	**17**	18
ETD	$\lambda=0.45,0.2,0.5;\mu=0.1$	10	0	**90**
Set 7				
CR	$\mu=0.1;\lambda=0.4$	**88**	0	12
CTD	$\mu=0.05,0.1,0.2;\lambda=0.4$	91	**0**	9
ETD	$\mu=0.05,0.1,0.2;\lambda=0.4$	64	0	**36**
Set 8				
CR	$\lambda=0.25;\mu=0$	**91**	0	9
CTD	$\lambda=0.1,0.3,0.4,0.15,0.6;\mu=0$	22	**65**	13
ETD	$\lambda=0.1,0.3,0.4,0.15,0.6;\mu=0$	9	0	**91**

Note: One hundred data sets were simulated under CR, CTD, and ETD diversification. SecSSE analyses were carried out under the same models for each simulated data set. Rows show the number of cases in which each model was chosen by SecSSE as the best performing model (lowest AIC). Sets 1–6 show variation in speciation rate between the states of a three-state simulated trait. Simulations in Sets 1 and 2 had a single transition rate (q) which was the same for all the trait states. In Set 3, transition rates among the states of a concealed trait were higher than among states of the examined state, whereas in Set 4 it was the opposite. Set 5 differs from Set 1 in the dissimilarity in speciation rates (λ) among the states. Set 6 includes extinction ($\mu > 0$). In Set 7, trait states have the same speciation rate but different extinction rate (trait-dependent extinction). Finally in Set 8 a trait with five states was simulated. $q=0.1$ in all cases if not specified otherwise. The instances where SecSSE correctly chose the generating model are shown in bold.

ETD model as the best performing in 66% of cases when ETD was not the generating model. In the partial information scenario, the percentage decreased to 41%. With full information SecSSE erroneously selects ETD in 32% of the cases (Table 3, Set 5). With regard to parameter estimation, the confidence intervals became narrower and medians were closer to the real parameter values when partial information was included (Fig. S3, Supplementary Material available on Dryad).

We found that there is no clear tendency to prefer models with higher or lower number of concealed states (Fig. S5, Supplementary Material available on Dryad; Table 2), because models with as many concealed states as examined ones performed similarly to more complex models.

DISCUSSION

We have introduced a generalization of the concealed-state framework that can be applied to traits with multiple states or multiple traits. We used it to re-evaluate trait-dependent diversification for seven studies where a MuSSE-type analysis, that is, ignoring concealed traits, had been done. Our analysis revealed that the conclusions of these studies are not fully supported. We have also shown that SecSSE, like HiSSE, avoids high rates of false positives.

Simulations showed that when extinction is high or differences in diversification are due to variation in extinction rates across states, the reliability of SecSSE decreases. This is not a surprising finding: high-extinction rate generally makes parameter estimations difficult with a diversity-independent model (Etienne et al. 2012), because extinction erases the signature of macroevolutionary processes on reconstructed phylogenies (Rabosky 2010). We observed that for high extinction rates SecSSE tends to erroneously choose a CR model more often than when extinction is zero, that is, the power of SecSSE is compromised. However, Type I error (inferring dependence of speciation on examined traits when it does not exist) remains
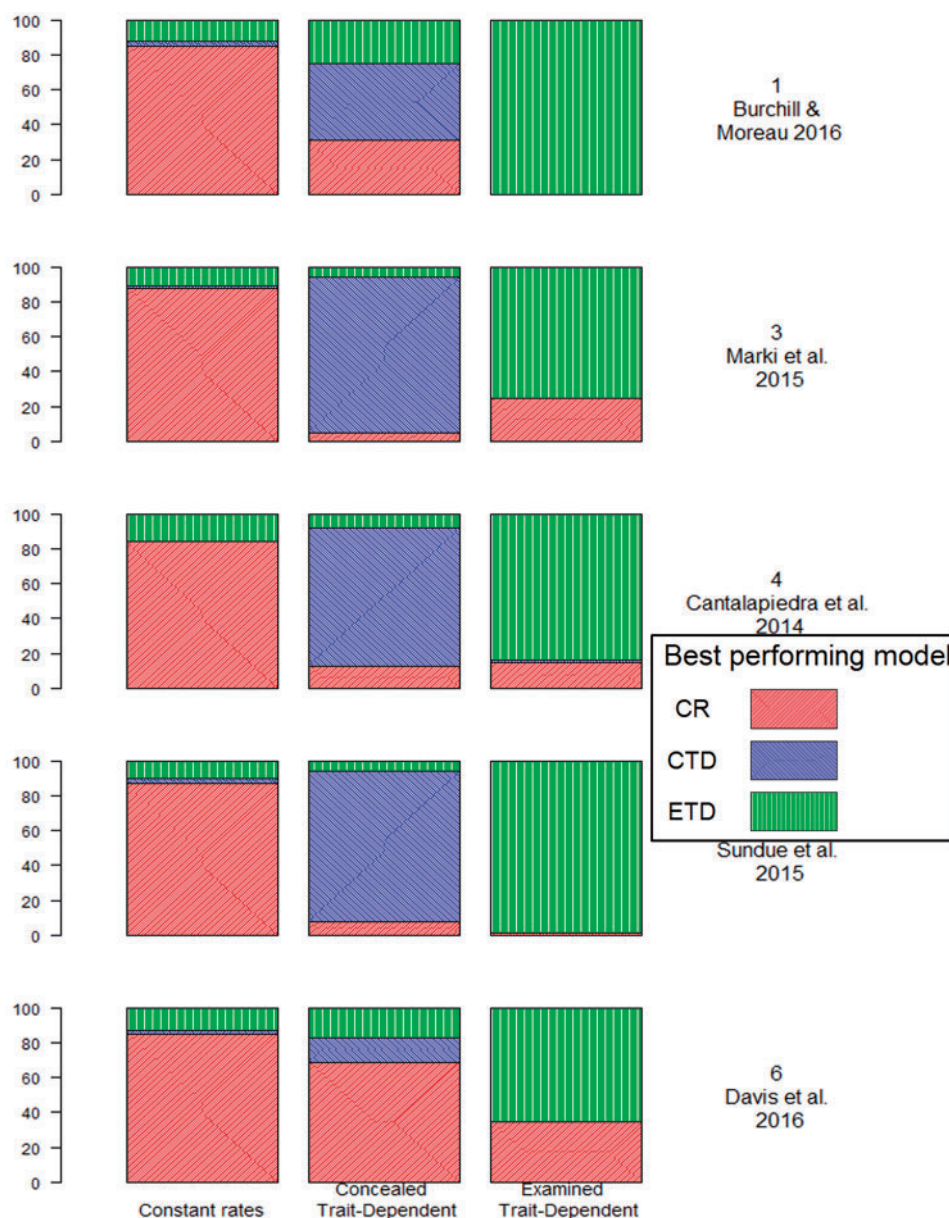
FIGURE 1. Performance of SecSSE in model selection using simulated data sets that are structurally similar to the empirical data sets. One hundred data sets were simulated under CR, CTD, and ETD diversification. SecSSE analyses were carried out under the same models for each simulated data set. Each plot shows the proportion of cases in which each model was chosen by SecSSE as the best performing model (lowest AIC) under different generating models (columns). CR = constant-rate; CTD = concealed trait-dependent; ETD = examined trait-dependent.

relatively low. It is often the case that SecSSE points at a CR model as the best performing model when in fact CTD is the generating model. This indicates that for certain parameter combinations the method might be insensitive to detect shifts in diversification rates. However, we note that SecSSE's purpose is to test dependence of diversification on traits. If the purpose is to find deviations from a CR model without reference to traits, there are other, more appropriate, approaches (e.g., BAMM, Rabosky 2014; DDD, Etienne and Haegeman 2012; MEDUSA, Alfaro et al. 2009).

Phylogenies and trait data contain information on trait-dependent diversification, but if Type I errors are to be avoided, the macroevolutionary dynamics must often be strongly and solely influenced by the trait to be detected. When we simulated trait dependence, all diversification rates were linked to species traits, every single shift in trait state influenced the rate of speciation and for those simulations SecSSE correctly chose an examined-trait-dependent diversification model most of the times. We argue that for empirical data sets SecSSE can find evidence of trait-dependent diversification only when the diversification events are strongly driven by
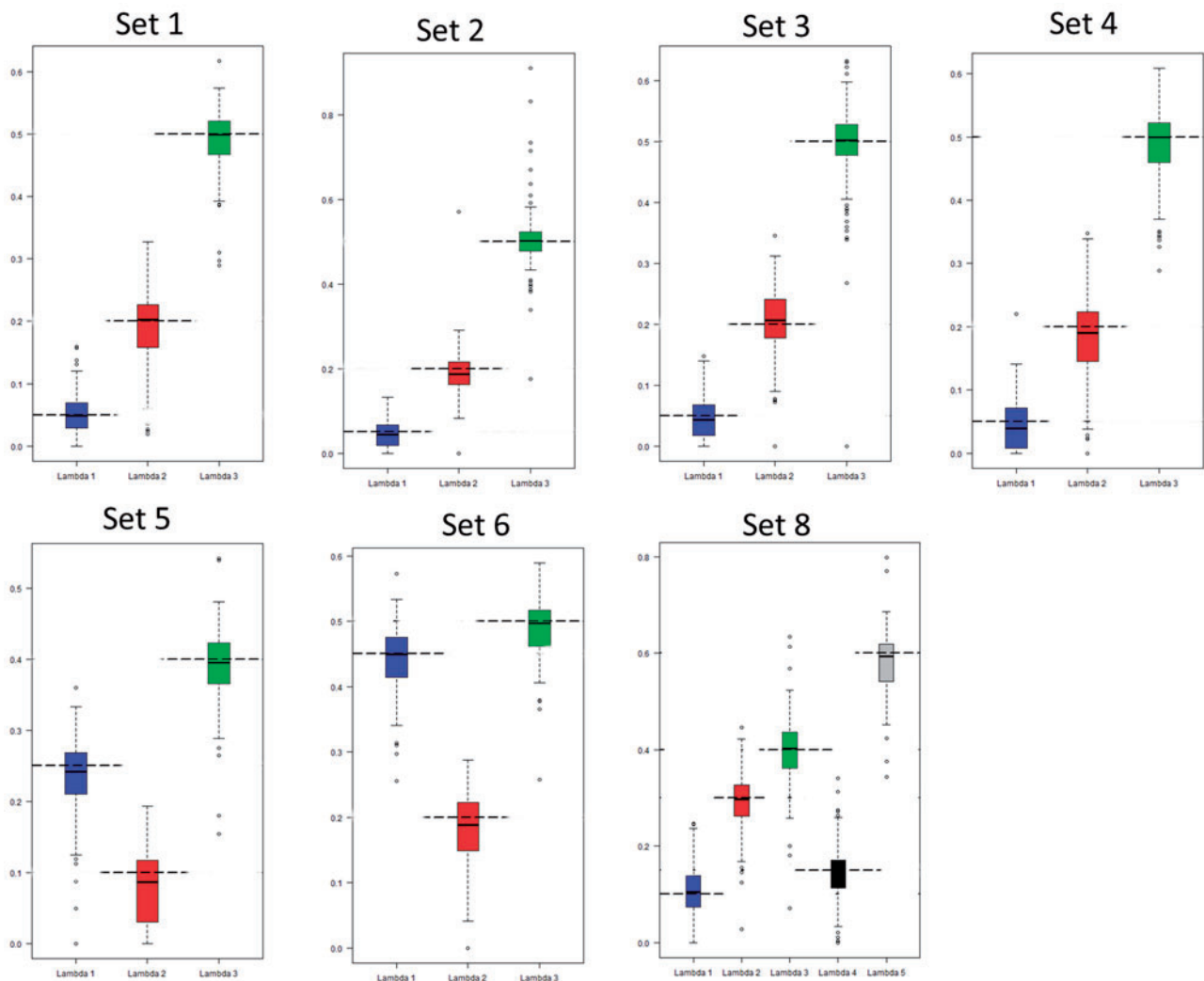
FIGURE 2. Estimation of speciation rate (λ) in different sets of simulations. Phylogenetic trees and trait states were simulated under a model of speciation rates dependent on the state of an evolving character (examined trait-dependent model of diversification). For those instances where SecSSE correctly inferred trait dependence, the parameter estimates are shown in the boxplot, whereas the generating parameters are plotted as horizontal lines. Sets 1–6 show variation in speciation rate between the states of a three-state trait. Simulations in Sets 1 and 2 had a single transition rate (q) which was the same for all the trait states. In Set 3, transition rates among the states of a concealed trait were higher than among states of the examined state whereas in Set 4 it was the opposite. Set 5 differs from Set 1 in the dissimilarity in speciation rates (λ) among the states. Set 6 includes extinction ($\mu > 0$). Finally in Set 8 a trait with five states was simulated.

TABLE 4. Performance of SecSSE in model selection when a model of heterogeneous rate across lineages is used as the generating model

			Model used for inference		
	Generating model		CR	CTD	ETD
Set 9	$q=0.1$	12 shift-rate events	45	29	26
		6 shift-rate events	67	15	18
Set 10	$q=0.05$	12 shift-rate events	33	46	21
		6 shift-rate events	57	21	22

Note: In this model, the evolving trait is completely independent of the dynamics of speciation and extinction. Shift-rate events make lineages switch to a different speciation rate. For each of the 100 simulated data sets SecSSE analyses were carried out under CR, CTD, and ETD diversification. Rows show the number of cases in which each model was chosen by SecSSE as the best performing model (lowest AIC). Sets 9 and 10 differ in the rate of trait evolution (switching rate q).

the trait state. If the trait in question is only partially responsible for branching patterns in the phylogeny, SecSSE might not find support for a link between trait state and macroevolutionary dynamics. In this sense, the concealed-state framework is conservative. When a SecSSE analysis reveals a link between a trait and diversification, we can be quite certain that this is indeed the case. Thus, we foresee that future studies

on macroevolution may fail to find a signal of trait dependence *unless* trait evolution and diversification dynamics are strongly coupled.

The suitability of SSE models for studying trait-dependent diversification depends on the nature of the trait. These models assume that a shift in character states does not cause a lineage split. If, however, the trait is linked to ecological opportunity, a change in the trait state might lead to ecological diversification, violating a core assumption of these models. If this were the case in the revisited studies, neither MuSSE nor SecSSE would be the right tool to use and this might explain why evidence for trait-dependent diversification was not found. A next step in SSE research is to develop a concealed-trait version of the ClaSSE model proposed by Goldberg and Igiić (2012) where cladogenetic changes in trait states are allowed. With such a model, traits potentially associated with ecological diversification can be analyzed. Comparing the performance of SecSSE and a concealed-trait version of ClaSSE would reveal, in a statistically more robust way, whether state shifts occur during cladogenesis or more gradually along branches, but in either case causing differential speciation and/or extinction rates.

The parameter estimates we obtained under the correct and the diversitree/HiSSE conditioning on nonextinction were similar. This can be understood intuitively as follows. Consider, for simplicity, a case of only two states (i.e., BiSSE). Large differences between the two ways of conditioning can arise only if the extinction probabilities for the two root states are very different. But when they are very different, the state with high extinction probability will often get a very low weight (as it is not very likely to have resulted in the observed phylogeny), and hence only the other state matters, and both likelihood formulas will be similar apart from a constant. We note, however, that when we use a different weighting scheme than Eq. (2), this argument no longer holds; for instance, with uniform weights on the root states, the different conditionings will have different results, if the extinction rate of one of the states is high.

The likelihood-free methods to detect trait-dependent diversification (Rabosky and Huang 2016; Rabosky and Goldberg 2017) might appear as a more attractive option over SSE models because of their relatively simplicity and short computation time. However, there are four differences between these two approaches. First, in the nonparametric approaches speciation rate does not vary along a branch according to shifts in trait states (unlike in SSE models), so the relationship between a trait and diversification rates can at most be correlational. In other words, when using such models we could argue for an association between a trait and diversification rate, but we cannot claim causation. Second, there might be instances when diversification rates are not the main focus of research but the evolution of the trait itself. For instance, one could be interested in transition rates between the states of a trait which is linked to speciation. In this case, trait evolution cannot be analyzed separately from branching patterns, and SecSSE is the right tool to use. Third, in SecSSE one can incorporate additional biological information or design models with different constraints, for example some state transitions can be set as forbidden, which allows for testing a number of contrasting hypotheses. Finally, SecSSE is a likelihood-based framework, which is a sound and coherent statistical framework. Because both methods address very similar questions, they can be regarded as complementary to some extent (Zenil-Ferguson and Pennell 2017).

For future applications of SecSSE we advocate to fit models with the same number of concealed states as examined states, and with the same structure in transition rates between the states, for example, when certain transitions are not allowed between examined states, we recommend that transitions between concealed states are also forbidden. This is simply to reduce the otherwise very large number of parameters, and to avoid that conclusions are based on the presence or absence of this structure rather than on the influence of the examined or concealed states on diversification rates. We further suggest that not only various models are fitted to the data, but also that simulations are used to explore Types I and II errors more thoroughly, as we did for the revisited cases, in order to assess the robustness of the model comparison.

We have shown that when trait evolution is coupled to diversification rates, this does leave a signature on phylogenetic trees and extant species traits and that can be detected with SecSSE with relatively high power yet still relatively low Type I error. We see a promising development of models where different modes of speciation and trait evolution are explored to unravel the drivers of diversification.

SUPPLEMENTARY MATERIAL

Data available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.qf3g0.

FUNDING

ACKNOWLEDGMENTS

Technology of the University of Groningen for their support and for providing access to the Peregrine high performance computing cluster.

REFERENCES

Aitkin M. 1991. Posterior Bayes factors. J. R. Stat. Soc. Ser. B. 53:111–142.

Alfaro M.E., Santini F., Brock C., Alamillo H., Rabosky D.L., Carnevale G., Harmon L.J., Hillis D.M., Alfaro M.E., Santini F., Brockb C., Alamillob H., Dornburg A., Raboskyde D.L., Carnevalef G., Harmon L.J. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. Proc. Natl. Acad. Sci. U. S. A. 106:13410–13414.

Arbuckle K., Speed M.P. 2015. Antipredator defenses predict diversification rates. Proc. Natl. Acad. Sci. U. S. A. 112:13597–13602.

Beaulieu J.M., O'Meara B.C. 2016. Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. Syst. Biol. 65:583–601.

Burchill A.T., Moreau C.S. 2016. Colony size evolution in ants: macroevolutionary trends. Insectes Soc. 63:291–298.

Cantalapiedra J.L., FitzJohn R.G., Kuhn T.S., Fernandez M.H., DeMiguel D., Azanza B., Morales J., Mooers A.O. 2013. Dietary innovations spurred the diversification of ruminants during the Caenozoic. Proc. R. Soc. B Biol. Sci. 281:20132746–20132746.

Davis A.M., Unmack P.J., Vari R.P., Betancur R. R. 2016. Herbivory promotes dental disparification and macroevolutionary dynamics in grunters (teleostei: terapontidae), a freshwater adaptive radiation. Am. Nat. 187:320–333.

Etienne R.S., Haegeman B. 2012. A conceptual and statistical framework for adaptive radiations with a key role for diversity dependence. Am. Nat. 180:E75–E89.

Etienne R.S., Haegeman B., Stadler T., Aze T., Pearson P.N., Purvis A., Phillimore A.B. 2012. Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. Proc. Biol. Sci. 279:1300–9.

Fitzjohn R.G. 2012. Diversitree: Comparative phylogenetic analyses of diversification in R. Methods Ecol. Evol. 3:1084–1092.

Goldberg E.E., Igić B. 2008. On phylogenetic test of irreversible evolution. Evolution (N. Y.). 62:2727–2741.

Goldberg E.E., Igiić B. 2012. Tempo and mode in plant breeding system evolution. Evolution (N. Y.). 66:3701–3709.

Helmstetter A.J., Papadopulos A.S.T., Igea J., Van Dooren T.J.M., Leroi A.M., Savolainen V. 2016. Viviparity stimulates diversification in an order of fish. Nat. Commun. 7:11271.

Jablonski D. 2008. Species selection: theory and data. Annu. Rev. Ecol. Evol. Syst. 39:501–524.

Maddison W.P.M., Fitzjohn R.G. 2014. The unsolved challenge to phylogenetic correlation tests for categorical characters. Syst. Biol. 64:127–136.

Maddison W.P., Midford P.E., Otto S.P. 2007. Estimating a binary character's effect on speciation and extinction. Syst. Biol. 56:701–710.

Marki P.Z., Fabre P.H., Jønsson K.A., Rahbek C., Fjeldså J., Kennedy J.D. 2015. Breeding system evolution influenced the geographic expansion and diversification of the core Corvoidea (Aves: Passeriformes). Evolution (N. Y.). 69:1874–1924.

Nee S., Holmes E.C., May R.M., Harvey P.H. 1994. Extinction rates can be estimated from molecular phylogenies. Philos. Trans. R. Soc. B Biol. Sci. 344:77–82.

Rabosky D.L. 2010. Extinction rates should not be estimated from molecular phylogenies. Evolution (N. Y.). 64:1816–1824.

Rabosky D.L. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. PLoS One. 9(2): e89543.

Rabosky D.L., Goldberg E.E. 2015. Model inadequacy and mistaken inferences of trait-dependent speciation. Syst. Biol. 64:340–355.

Rabosky D.L., Goldberg E.E. 2017. FiSSE: a simple nonparametric test for the effects of a binary character on lineage diversification rates. Evolution (N. Y.). 71:1432–1442.

Rabosky D.L., Huang H. 2016. A robust semi-parametric test for detecting trait-dependent diversification. Syst. Biol. 65:181–193.

Sundue M.A., Testo W.L., Ranker T.A. 2015. Morphological innovation, ecological opportunity, and the radiation of a major vascular epiphyte lineage. Evolution (N. Y.). 69:2482–2495.

Wagenmakers E.-J., Farrell S. 2004. AIC model selection using Akaike weights. Psychon. Bull. Rev. 11:192–196.

Zenil-Ferguson R., Pennell M.W. 2017. Digest: trait-dependent diversification and its alternatives. Evolution (N. Y.). 71:1732–1734.