

APPLICATION

Diversitree: comparative phylogenetic analyses of diversification in R

Richard G. FitzJohn*

Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

Summary

1. The R package ‘diversitree’ contains a number of classical and contemporary comparative phylogenetic methods. Key included methods are BiSSE (binary state speciation and extinction), MuSSE (a multistate extension of BiSSE), and QuaSSE (quantitative state speciation and extinction). Diversitree also includes methods for analysing trait evolution and estimating speciation/extinction rates independently.

2. In this note, I describe the features and demonstrate use of the package, using a new method, MuSSE (multistate speciation and extinction), to examine the joint effects of two traits on speciation.

3. Using simulations, I found that MuSSE could reliably detect that a binary trait that affected speciation rates when simultaneously accounting for additional traits that had no effect on speciation rates.

4. Diversitree is an open source and available on the Comprehensive R Archive Network (CRAN). A tutorial and worked examples can be downloaded from <http://www.zoology.ubc.ca/prog/diversitree>.

Key-words: birth–death process, comparative methods, extinction, macroevolution, speciation

Introduction

The tree of life is remarkably uneven in both taxonomic and trait diversity; describing this unevenness and revealing its underlying causes are major focuses of evolutionary biology. Comparative phylogenetic methods have been widely used to study patterns and rates of both trait evolution (Felsenstein 1985; Pagel 1994) and diversification (Nee, May & Harvey 1994). A recently developed set of models unites both trait evolution and species diversification, avoiding biases that occur when the two are treated separately (Maddison 2006). This includes the ‘BiSSE’ method (binary state speciation and extinction; Maddison, Midford & Otto 2007), as well as similar methods that generalise the approach to nonanagenetic trait evolution and to quantitative traits.

In this note, I describe the ‘diversitree’ package for R (R Development Core Team, 2012). Diversitree implements several recently developed methods for analysing trait evolution, speciation, extinction, and their interactions. Below, I describe the general approach of the package and the methods that it contains. I introduce a generalisation of the BiSSE method to multistate characters or to combinations of binary

traits (MuSSE: multistate speciation and extinction). Finally, I demonstrate the package, and MuSSE, with an example of social trait evolution in primates.

The methods

The diversitree package implements a series of methods for detecting associations between species traits and rates of speciation and/or extinction, given a phylogeny and trait data, including the BiSSE method (Maddison, Midford & Otto 2007). Under BiSSE, speciation and extinction follow a birth–death process, where the rate of speciation and extinction may vary with a binary trait, itself evolving following a continuous-time Markov process. BiSSE has been used to look at the associations between many different traits and speciation or extinction, including migration in warblers (Winger, Lovette & Winker 2012), fruiting body morphology in fungi (Wilson, Binder & Hibbett 2011) and recombination in plants (Johnson *et al.* 2011).

In its original formulation, BiSSE assumes that character change occurs only along branches (anagenetic change), using the same model of character evolution as used in the ‘discrete’ (Pagel 1994) or ‘Mk’ models (Lewis 2001). This may not always be a reasonable assumption, and we might expect some

*Correspondence author. E-mail: fitzjohn@zoology.ubc.ca

characters to show considerable change during speciation (cladogenetic change). One such example is geographic range; while geographic ranges are expected to change anagenetically, allopatric speciation should also alter range sizes. The Geographic SSE (GeoSSE; Goldberg, Lancaster & Ree 2011) method allows speciation rates to vary depending on a species' presence in two different geographic regions, allowing within- and between-region speciation. This has been used to examine diversification in plants endemic to serpentine regions (Anacker *et al.* 2010). More recently, the BiSSE-ness (BiSSE-node enhanced state shift; Magnuson-Ford & Otto 2012) and Cladogenetic SSE (ClaSSE; Goldberg & Igić *in press*) models have been developed to allow both anagenetic and cladogenetic character evolution, such as that expected for traits involved in ecological speciation (Schluter 2009). Importantly, with extinction or incomplete taxonomic sampling, not all speciation events will appear as nodes in a phylogeny; these missing nodes must be modelled to accurately estimate the rate of cladogenetic trait change (Nee, May & Harvey 1994 and Bokma 2008, and note that the placement of these missing nodes is nonlinear in time).

Diversitree also includes methods for nonbinary traits. Quantitative SSE (QuaSSE; FitzJohn 2010) allows speciation and extinction rates to be modelled as any user-supplied function of a continuously varying trait, which itself evolves under Brownian motion. This has been used to test for associations between diversification rates and body size in snakes (Burbrink, Ruane & Pyron 2012) and dispersal ability in birds (Claramunt *et al.* 2012). Finally, MuSSE extends BiSSE to multistate traits or combinations of binary traits.

Diversitree includes variants that relax some of the original assumptions of the included methods. Birth–death-based speciation/extinction models will give biased parameter estimates unless all extant taxa in the focal clade are present in a phylogeny. For cases where not all extant species are included in a phylogeny, diversitree includes methods for where species are included randomly or where all species are represented in 'unresolved clades' (FitzJohn, Maddison & Otto 2009). Rates of speciation, extinction or character change can be set to vary as any user-supplied function of time. Similar approaches have been used elsewhere to model slowdowns in speciation or diversification over time (Rabosky & Glor 2010).

Rates of speciation, extinction and character change may also be allowed to vary in different regions of a tree. This is similar to MEDUSA (modelling evolutionary diversification under stepwise AIC: Alfaro *et al.* 2009) for diversification and AUTEUR (Eastman *et al.* 2011) for continuous character evolution. Such methods can be used to test whether membership of a clade that has undergone a shift in diversification rates is misleading BiSSE or other methods. For example, if particular trait values are concentrated in a highly diverse clade, BiSSE may detect an association when none exists (see applications in Johnson *et al.* 2011 and FitzJohn 2010, the diversitree tutorial for a worked example, and further discussion in Read & Nee 1995).

In the above models, if speciation and extinction do not vary with character state, the models converge on classical models of character evolution (Pagel, 1994) and state-independent speciation and extinction (Nee, May & Harvey 1994). For completeness, these models are also included. However, when comparing models to determine whether traits are associated with speciation or extinction using likelihood ratio tests, comparisons must involve only nested models to be valid. For example, BiSSE and Mk2 are not directly comparable, but BiSSE can be compared with a constrained version of BiSSE that disallows state-dependent diversification. See Table 1 for a summary of included methods.

In addition to the likelihood calculations, tree simulation routines are implemented for birth–death models, BiSSE, MuSSE and QuaSSE. Simulating character evolution on a given tree is possible for discrete (binary or multistate) characters and continuous characters under Brownian motion and Ornstein–Uhlenbeck processes. Ancestral state reconstruction (Schluter *et al.* 1997) and stochastic character mapping (Bollback 2006) are implemented for discrete characters.

The approach

In diversitree, the inference process is decoupled from the likelihood calculations, allowing users to take advantage of the programmatic flexibility of R. Analyses, therefore, require at least two steps. First, the user creates a likelihood function from their tree and data, using a `make.xxx` function (where `xxx` is one of the model types available). For example, to model character evolution under a two-state Markov model (Lewis 2001), the user would enter:

```
lik<-make.mk2(tree, states)
```

Secondly, we can find the maximum likelihood (ML) parameter vector for this function:

```
fit<-find.mle(lik, starting.parameters)
```

or use it in a Bayesian analysis by running an MCMC (Markov chain Monte Carlo) chain (with an appropriate prior):

```
samples<-mcmc(lik, starting.parameters,
nsteps, proposal.widths, prior)
```

or in some other use (for example, integrating the function numerically to compute the 'integrated likelihood' for Bayes factors, for example, Kass & Raftery 1995).

Between these steps, the likelihood function can be constrained arbitrarily. Diversitree's `constrain` function allows several natural constraints, such as setting one parameter equal to another, or to a specific numerical value. For example, to constrain the forward and backward transition rates to be equal (reducing the Mk2 model to the Jukes–Cantor model):

```
lik.jc<-constrain(lik, q01~q10)
```

We could then find the ML parameter by entering

```
fit.jc<-find.mle(lik.jc, starting.parameters)
```

Table 1. Summary of model types available in diversitree (as of version 1.0)

Name	Trait ^a	Missing taxa ^b	Extensions ^c	Description and reference
bd	–	Sk, Un	Sp, Tv	Constant-rate birth–death (Nee, May & Harvey 1994)
mk2, mkn	B,M	–	Sp, Tv	Markov discrete character evolution (Pagel, 1994; Lewis, 2001)
bisse	B	Sk, Un	Sp, Tv	Binary State Speciation and Extinction (Maddison, Midford & Otto 2007; FitzJohn, Maddison & Otto 2009)
bisness	B	Sk, Un	–	BiSSE-ness (Magnuson-Ford & Otto 2012)
geosse	T	Sk	–	Geographic state speciation and extinction (Goldberg, Lancaster & Ree 2011)
musse	M	Sk, Un	Sp, Tv	Multistate speciation and extinction
classe	M	Sk	–	Clade-state speciation and extinction (Goldberg & Igić in press)
bm	Q	–	–	Brownian motion
ou	Q	–	–	Ornstein–Uhlenbeck
quasse	Q	Sk	Sp	Quantitative state speciation and extinction (FitzJohn 2010)

^a Trait type key: B = binary (0/1), T = ternary (three combinations of presence/absence in two regions), M = multistate (1, 2, 3, ...), Q = quantitative (real-valued). ^b Missing taxa support: Sk = ‘skeleton tree’ (random sampling) correction, Un = ‘unresolved clade’. ^c Extensions: Sp = ‘split tree’ (allows MEDUSA-style different rate classes in different areas of the tree), Tv = time-varying rates.

These nested models could then be compared using a likelihood ratio test.

Most of the methods included in diversitree are computationally challenging, but there are a number of options for controlling how the calculations are performed. Amongst these, the user can use different ODE solvers, and the accuracy of the calculations can be traded off against speed for most methods. Algorithms that have proven to be reasonably robust (in my experience) are used by default. For some models, such as Mk2, Brownian motion and Ornstein–Uhlenbeck, diversitree provides alternative algorithms that perform better with large numbers of states or large trees. The possible options and algorithms are discussed in Appendix S1 section 1 and 2.

Diversitree builds on much existing software: ape (Paradis, Claude & Strimmer 2004) is used for tree loading and manipulation; the deSolve package (Soetaert, Petzoldt & Setzer 2010) and sundials library (Hindmarsh *et al.* 2005) are used for solving the systems of differential equations for the discrete trait models; and FFTW (Frigo & Johnson 2005) is used to solve the partial differential equations in QuaSSE. In addition to the R interface, Wayne Maddison has developed a wrapper around some of diversitree’s functionality to allow use from within Mesquite (Maddison & Maddison 2008), using a user-friendly point-and-click interface.

The MuSSE model

MuSSE is a straightforward extension of BiSSE to discrete traits with more than two states. Some characters are not naturally binary (e.g. mating systems, diets or count data), and MuSSE allows these to be treated naturally. This method has been used to examine the effect of diet (faunivore, folivore, frugivore) in primates (Gómez & Verdú 2012). Alternatively, MuSSE can be used to disentangle the relative importance of two or more traits to diversification.

Suppose that we have a trait that takes values $1, 2, \dots, k$ that might influence speciation and/or extinction. Using the notation and approach of Maddison, Midford & Otto (2007), let lineages in state i speciate at rate λ_i , go extinct at rate μ_i and transition to state $j \neq i$ at rate q_{ij} . For k states, there are k speciation rates, k extinction rates and $k(k-1)$ transition rates.

Derivation

Let $D_{N,i}(t)$ be the probability of a lineage in state i at time t before the present ($t = 0$) evolving into its descendant clade as observed, and let $E_i(t)$ be the probability that a lineage in state i at time t , and all of its descendants, goes extinct by the present. Under the same assumptions as Maddison, Midford & Otto (2007) and using the same approach, it is possible to derive a set of ordinary differential equations that describe the evolution of the D and E variables over time:

$$\frac{dE_i(t)}{dt} = \mu_i - (\lambda_i + \mu_i + \sum_{j \neq i} q_{ij})E_i(t) + \lambda_i E_i(t)^2 + \sum_{j \neq i} q_{ij} E_j(t) \quad \text{eqn 1a}$$

$$\frac{dD_{N,i}(t)}{dt} = - \left(\lambda_i + \mu_i + \sum_{j \neq i} q_{ij} \right) D_{N,i}(t) + 2\lambda_i E_i(t) D_{N,i}(t) + \sum_{j \neq i} q_{ij} D_{N,j}(t). \quad \text{eqn 1b}$$

For k states, there are $2k$ equations.

We can solve this system of equations numerically from the tip to base of a branch. As with BiSSE, the initial conditions for the D variables are 1 when the trait combination

is consistent with the data, and 0 otherwise, while the initial conditions for all E variables are zero. Missing trait data are allowed by setting all D values to 1 (any state is consistent with the observed data). When the phylogeny is incomplete, the initial conditions can be modified by assuming random sampling (FitzJohn, Maddison & Otto 2009).

At the node N' that joins lineages N and M , we multiply the probabilities of both daughter lineages together with the rate of speciation

$$D_{N',i}(t) = D_{N,i}(t)D_{M,i}(t)\lambda_i. \quad \text{eqn 2}$$

The equations here assume no cladogenetic change, but this can be added following the approach in Magnuson-Ford & Otto (2012) or Goldberg & Igić (in press).

As the number of parameters in MuSSE grows quadratically with the number of states, care will often be required to prevent over-fitting and pathological behaviour associated with estimation of rate parameters involving states that are rarely observed. In particular, if some state i is not observed, then the likelihood surface never has a negative slope with increasing q_{ij} ($j \neq i$) and μ_i , causing ML values for these parameters to tend to infinity, in turn causing problems for both the maximisation and likelihood calculation routines. For ordinal traits, constraining the transition rates so that $q_{ij} = 0$ for $|i-j| > 1$ may be useful.

Analysing multiple traits simultaneously

Alternatively, this method can be generalised to combinations of binary traits, following Pagel (1994); in this scheme, a discrete state would represent the combination of different binary traits; for n binary traits, there are 2^n possible states. For example, for a pair of binary traits, there are four possible state combinations: (0,0), (0,1), (1,0), (1,1). We can denote these (1,2,3,4) and use MuSSE directly. However, in this 'multitrait' model, parameters may be unintuitive to interpret, particularly as the number of traits increases. Moreover, with multiple traits, we may be explicitly interested in asking whether combinations of traits affect speciation or extinction nonadditively, and this is difficult to determine with this parametrisation.

In diversitree, an alternative parametrisation is available to facilitate interpretation and model testing. Let $\lambda_{i,j}$ be the speciation rate of a species with states $A = i$, $B = j$, for two binary traits A and B . We can use a linear modelling approach and write

$$\lambda_{i,j} = \lambda_0 + \lambda_A X_A + \lambda_B X_B + \lambda_{AB} X_A X_B, \quad \text{eqn 3}$$

where X_A and X_B are indicator variables that are 1 when trait A and B are in the '1' state (respectively), λ_0 is the 'intercept' speciation rate (if all traits are in state 0), λ_A and λ_B are the 'main effects' of traits A and B , and λ_{AB} is the interaction between these. If a combination of A and B drives speciation, then a model with λ_{AB} will fit better than a model with just the main effects. Similarly, for the extinction rate, we write

$$\mu_{i,j} = \mu_0 + \mu_A X_A + \mu_B X_B + \mu_{AB} X_A X_B. \quad \text{eqn 4}$$

The same approach can be used for the character transition rates. If we follow Pagel (1994) and allow change in only a single trait during a single point in time, then for n traits, there are only $2n$ possible 'types' of transitions (i.e. a $0 \rightarrow 1$ or $1 \rightarrow 0$ transition in one of the n traits). However, the rate at which these transitions happen may vary depending on the state of the other traits. For example, with two traits, we can write the rate of transition in trait A from 0 to 1, given that trait B is in state j , as

$$q_{A01,j} = q_{A01,0} + q_{A01,B} X_B. \quad \text{eqn 5}$$

where $q_{A01,0}$ is the intercept term and $q_{A01,B}$ is the main effect of trait B . In this scheme, if a model with $q_{A01,B}$ fits better than a model without, then the rates of $0 \rightarrow 1$ transition of trait A depends on the state of trait B .

Similar schemes can be derived for more traits; for more than two states, interaction terms will appear in the equations. For example, with three traits (A , B and C)

$$q_{A01,j,k} = q_{A01,0} + q_{A01,B} X_B + q_{A01,C} X_C + q_{A01,BC} X_B X_C \quad \text{eqn 6}$$

where $q_{A01,C}$ is the main effect of trait C on the rate of character change of trait A from 0 to 1, and $q_{A01,BC}$ is an interaction effect that specifies the level of nonadditivity of the traits B and C on character change of trait A . Of course, this parametrisation of transition rates is valid for studying character evolution in multiple binary traits without modelling its effect on diversification (as in Pagel, 1994), and this can be done with the `make.mkn.multitrait` function.

If state information is available for some traits and not the others, the initial conditions are modified to allow any trait combination consistent with the observed data. For example, if trait A is in state 0 and the state of trait B is unknown, the D variables will be 1 for the combinations (0,0) and (0,1) and zero for combinations (1,0) and (1,1).

Simulation test assessing the power of MuSSE

There are a large number of distinct ways of modelling diversification with MuSSE, and I expect that the power of the model will depend strongly on the model specification. For example, one might have an ordinal multistate trait, where transitions can only occur between adjacent states and be interested asking whether large or small values of that trait are associated with elevated rates of diversification. For a given number of states (> 2), such a model will have far fewer parameters (and greater power) than a model where the trait is purely categorical, such as diet, if all transitions are possible. The power of MuSSE will strongly depend on the number of estimated parameters (especially the character transition parameters), and I expect that for any more than four states, careful consideration of constraints in the transition parameters will be needed.

Here, I focus on a simple multitrait case where there is some number of uncorrelated binary traits that evolve at the same rate, one of which influences the rate of speciation. I investigate the ability of MuSSE to correctly identify the trait associated with elevated speciation and to rule out the association with other traits, as a function of clade size and number of possible traits.

To simulate trees, I set the intercept speciation and extinction rates (λ_0 and μ_0) to 0.1 and 0.03, respectively, and character transition rates (q_{X01} , q_{X10} , for traits $X = A, B, \dots$) to 0.01. I set $\lambda_A = 0.1$ so that when trait A is in state 1, the speciation rate is 0.2. When only a single trait is considered, these are the same parameters used by Maddison, Midford & Otto (2007) in their ‘asymmetric speciation’ case. I simulated phylogenies and character state transitions under the multitrait MuSSE model, starting at the root in one of the ‘low’ speciation states (with A in state 0), sampling randomly for the other traits. Trees were simulated to contain 50, 100, 200 or 400 species, with 1, 2, 3 or 4 traits, and with 100 replicate trees for each of the 16 combinations.

For each tree, I ran a Markov chain Monte Carlo (MCMC) analysis on a model where all speciation main effects were free to vary (but excluded interactions), fitting only intercepts for extinction and character change. For example, with two traits, this meant that the free parameters were λ_0 , λ_A , λ_B , μ_0 ,

$q_{A01,0}$, $q_{A10,0}$, $q_{B01,0}$ and $q_{B10,0}$. This model is very close to the true model, but allows for uncertainty in which trait is responsible for increased speciation (trait A or B). I used an exponential prior with a mean of twice the state-independent diversification rate for all the underlying rate parameters (Appendix S1 section 3). I ran each chain for 10 000 steps and discarded the first 500 steps as ‘burn-in’. Because the ‘dummy’ traits B , C and D are equivalent where present, I report results primarily for trait A (which increases speciation rates when in state 1) and trait B (which does not affect speciation rates).

As the size of the tree increased, the credibility intervals around the main effects on speciation decreased, and the mean estimated effect converged on the true values (Fig. 1). The uncertainty around the dummy trait, B , was not strongly affected by the number of dummy traits that were included and decreased slightly as more traits were included. For small trees (≤ 100 species), MuSSE underestimated the effect of trait A on speciation rates, especially as the number of traits increased.

Significance showed similar patterns. As tree size increased, power to correctly identify A as the trait associated with increased speciation increased (Fig. 2, blue lines), but for trees with 100 species or more, this varied only weakly with the number of included traits. The dummy trait B was significant approximately 5% of the time (based on 95% credibility

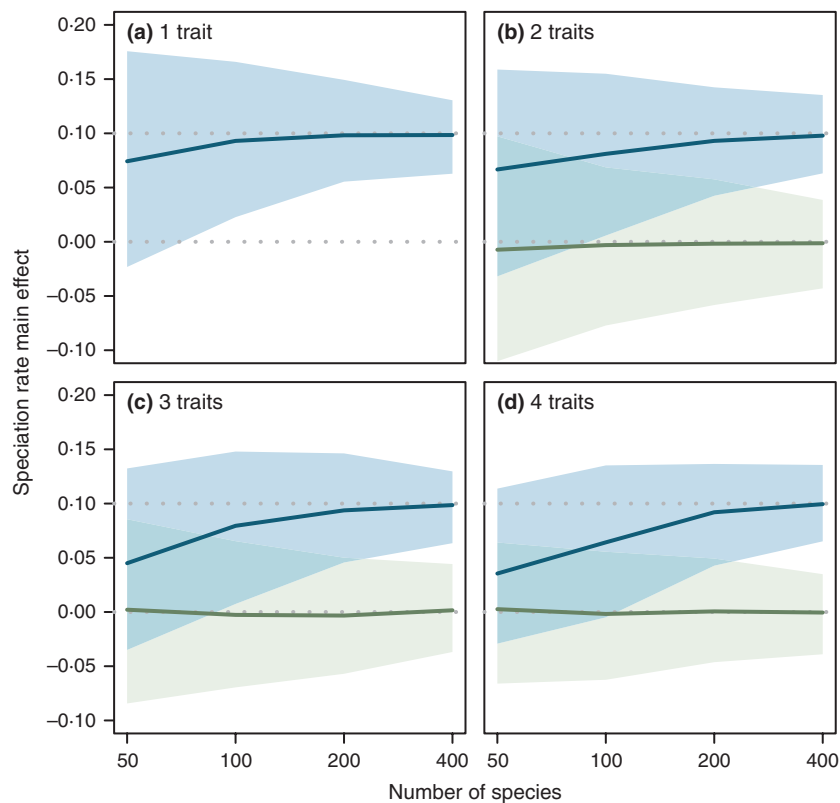


Fig. 1. Uncertainty around multitrait MuSSE parameter estimates as a function of tree size and number of traits. The solid blue line and blue region represent the mean and 95% credibility interval (CI) over 100 trees for the estimated speciation rate main effect of trait A , which increases speciation rates (true value is 0.1, indicated by the grey dotted line). The solid green line and region represent the mean and 95% CI for the speciation rate main effect for trait B , which has no effect on speciation rates (true value of zero indicated by dotted grey line). Panel (a), with one trait, is equivalent to BiSSE.

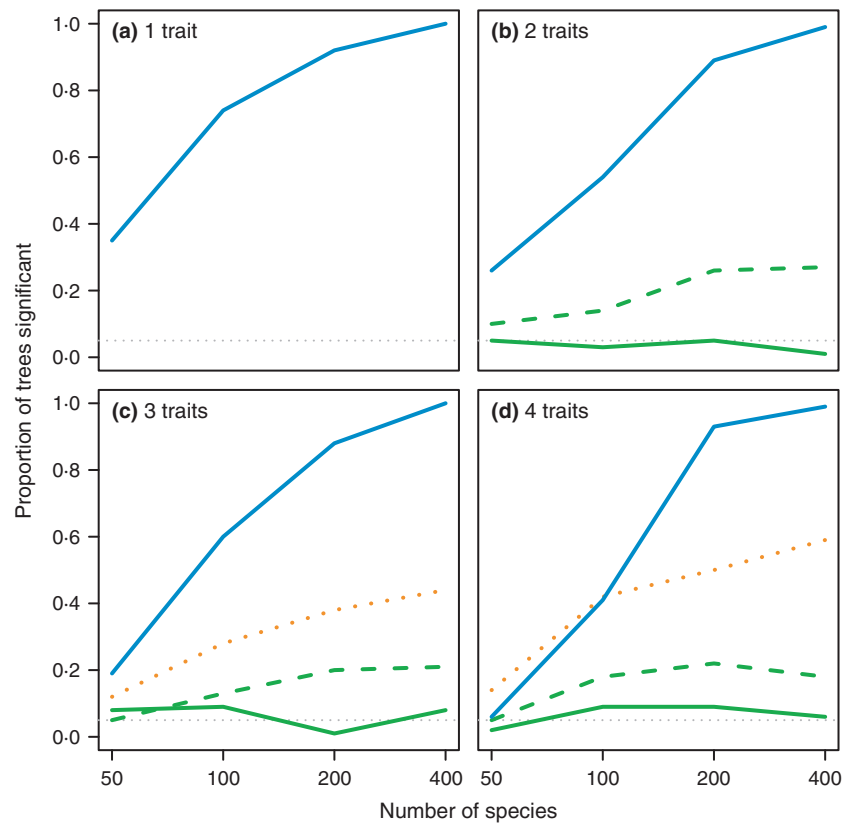


Fig. 2. Power and error rates of multitrait MuSSE, as a function of tree size. The lines are the proportion of 100 simulated trees that have 95% credibility intervals of speciation main effects that do not include zero (indicating significant state-dependent speciation). The blue line represents trait *A*, which increases speciation rates when in state 1. The solid green line represents a trait *B* with no effect on speciation. The dashed green line indicates the same trait *B*, but when trait *A* is omitted from the analysis. The dotted orange line in panels (c) and (d) is the probability of finding *any* of the dummy traits (*B*, *C*, or, where present *D*) significant in an analysis that omits trait *A*. The 5% expected type I error rate is indicated by the dotted grey line. Panel (a), with one trait, and the dashed green line in panel (b) are equivalent to BiSSE.

intervals): the rate expected because of type I error (Fig. 2, solid green lines).

To test how model misspecification would affect the results, I also reran the analyses with trait *A* omitted so that none of the analysed traits were truly associated with state-dependent diversification. The dummy trait *B* was incorrectly associated with increased speciation in up to 27% of trees (Fig. 2). While this effect was strongest when there were fewer dummy traits, the possibility of *any* trait being falsely associated with diversification increased. Indeed, where three dummy traits are included, the probability of associating any trait with increased speciation increased to 59% for the 400 species tree (Fig. 2, dotted orange lines).

These results are simultaneously encouraging and sobering. When a trait that affects speciation is included in the model, it is easily detected, and this is robust to the number of additional traits included. However, if no traits do affect speciation, as we add additional traits we risk false positives at an alarming rate. However, the rates of false positives are perhaps not surprising. The trees used do not conform well to the expectations of a constant-rate birth–death tree (there is strong phylogenetically structured variation in speciation rates), and the model is

using the only parameters it has to explain this deviation. I expect that similar problems will affect other comparative analyses such as detecting correlated trait evolution with the Mk/discrete models.

The code for this analysis is available on the diversitree github site (<http://github.com/richfitz/diversitree/tree/pub/simulations>).

Social evolution and speciation in primates

Here I give a worked example, using the trait data compiled by Redding, DeWolff & Mooers (2010), to look at social evolution in primates. Previously, Magnuson-Ford & Otto (2012) found that both monogamy and solitary behaviour in primates reduced speciation rates, although this was only marginally significant for solitariness. However, if these characters are correlated, then it is possible that the decreased speciation rates could be truly associated with just one trait. That is, the effect of one character might bias the estimated effects of the other when these are treated independently. Alternatively, it could be that an elevated (or decreased) speciation rate occurs only with some combination

of trait states (e.g. only social, polygamous taxa speciate more rapidly).

Here, I illustrate the method with R input preceded by '>', while output is upright. The full version of this analysis is presented in Appendix S1 section 3. The phylogeny is stored in NEXUS format (Maddison, Swofford & Maddison 1997) and loaded using the `read.nexus` function in `ape` as the object 'tree'. For multitrait MuSSE, the data must be stored in a data frame with species names as row labels. The two traits are 'M' (TRUE for monogamous, FALSE otherwise) and 'S' (TRUE for solitary, FALSE otherwise).

```
> head(dat)
                M          S
Allenopithecus_nigroviridis NA      FALSE
Allocebus_trichotis        TRUE      TRUE
Alouatta_belzebul          NA      FALSE
Alouatta_caraya            NA      FALSE
Alouatta_coibensis        FALSE     FALSE
Alouatta_fusca             NA      FALSE
```

Note that some of the species lack state information (i.e. have NA values). These are accommodated using the method described earlier.

The first step is to make a likelihood function with `make.musse.multitrait`. The 'depth' argument controls the number of terms to include from eqns (3–5): 0 includes only intercepts, 1 includes main effects, 2 includes interactions between two parameters and so on. If specified as a 3-element vector, the elements apply to the λ , μ and q parameters; if a scalar is given, the same depth is used for all three parameter types. To make a model with intercepts only:

```
> lik.0 <- make.musse.multitrait(tree, dat,
  depth=0)
```

This likelihood function takes a vector of parameters as its first argument. To obtain the vector of names for the parameters, use the `argnames` function:

```
> argnames(lik.0)
[ 1] "lambda0" "mu0" "qM01.0" "qM10.0"
[ 5] "qS01.0" "qS10.0"
```

This shows the six parameters: the speciation rate (`lambda0`), extinction rate (`mu0`) and four transition rates (e.g. `qM01.0` is the rate of transition of the breeding system from nonmonogamous to monogamous, and this rate does not depend on the social state `S`).

To find the maximum likelihood (ML) point, a sensible starting point must be supplied (discussed in Appendix S1 section 3); with such a point, `p.0`, we can find the ML parameters using the `find.mle` function:

```
> fit.0 <- find.mle(lik.0, p.0)
```

This returns an object (`fit.0`) that contains estimated parameters, likelihood values and other information about the fit (see the help page `?find.mle` for more information).

```
> round(coef(fit.0), 4)
lambda0 mu0 qM01.0 qM10.0 qS01.0 qS10.0
0.1912  0.1110 0.0251 0.0259 0.0009 0.0163
> fit.0$lnLik
[ 1] -786.3427
```

By default, 'subplex' (Rowan 1990) is used for the optimisation. However, different optimisation algorithms can be selected through the 'method' argument to `find.mle`.

To include state-dependent diversification, we construct a likelihood function that includes 'main effects' of the two traits on speciation and extinction. To allow this while retaining the independent model of character evolution, we change the `depth` argument:

```
> lik.1 <- make.musse.multitrait(tree, dat,
  depth=c(1, 1, 0))
> argnames(lik.1)
[ 1] "lambda0" "lambdaM" "lambdaS" "mu0"
[ 5] "muM"      "muS"      "qM01.0" "qM10.0"
[ 9] "qS01.0" "qS10.0"
```

Running an ML search from a suitable point `p.1`:

```
> fit.1 <- find.mle(lik.1, p.1)
```

These models can be compared using a likelihood ratio tests using the ANOVA function; the model with state-dependent speciation and extinction fits much better than the state-independent version ($\chi^2_4 = 24.7$, $P < 0.001$).

```
> anova(fit.1, noSDD=fit.0)
      Df lnLik AIC   ChiSq Pr(> |Chi|)
full  10 -773.97 1568.0
noSDD  6 -786.34 1584.7  24.739 5.677e-05
```

(The use of ANOVA for general model comparison is a fairly widespread convention in R packages and does not imply that an ANOVA was performed!)

We can expand the model further to allow interactions between the two traits in speciation and extinction; Is a combination of mating system and sociality associated with elevated speciation or extinction? Specifying `depth=c(2, 2, 0)` introduces the terms 'lambda.MS' and 'mu.MS' (eqns 3 and 4) to model nonadditive effects of these traits on speciation and extinction and again leaves character transitions to occur independently for the two traits.

```
> lik.2 <- make.musse.multitrait(tree, dat, depth=
  c(2, 2, 0))
> fit.2 <- find.mle(lik.2, p.2)
> anova(fit.2, addonly=fit.1)
      Df lnLik AIC ChiSq Pr(> |Chi|)
full  12 -773.73 1571.5
addonly 10 -773.97 1568.0 0.49143 0.7821
```

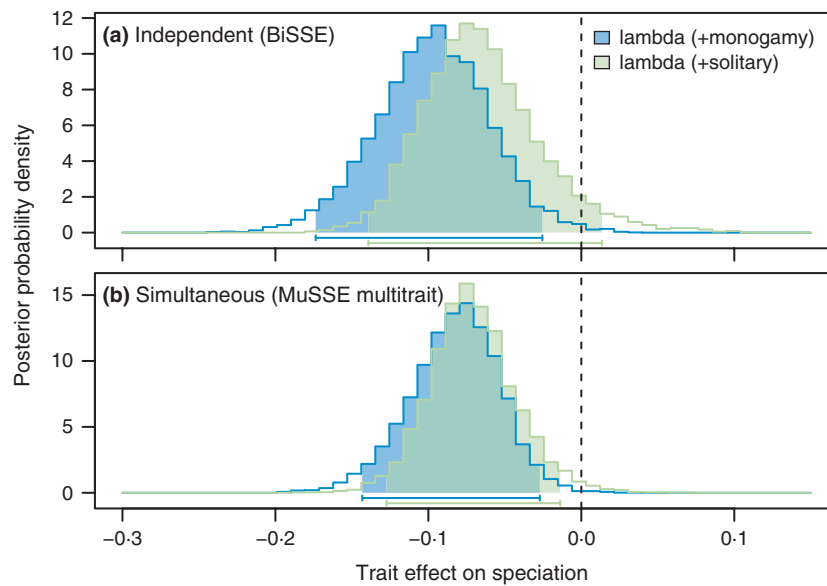


Fig. 3. Posterior probability distributions for the effects of monogamy (dark grey) and solitariness (light grey) on speciation rate. Shaded areas and bars indicate the 95% credibility intervals for each parameter. In the top panel, BiSSE was run on each character independently. In the bottom panel, the `musse.multitrait` fit the effects of both traits simultaneously. In both cases, the MCMC chain was run for 10 000 steps, and the first 500 points were dropped as burn-in.

This time the improvement is not significant, implying that there is no evidence for an interaction between these traits on speciation and extinction rates.

To test the significance of the each trait (solitariness and monogamy) in a maximum likelihood framework, we could fit models where the main effect of each trait was set to zero and compare these against the model `fit.1` using a likelihood ratio test. This approach is explored in Appendix S1 section 3. Alternatively, we might run an MCMC and examine the posterior distributions of the `lambdaM` and `lambdaS` values:

```
> samples<-mcmc(lik.1, p.1, nsteps=10000,
  w=0.5, prior=prior)
```

The prior distribution used here is exponential with respect to the underlying rates in the model (e.g. λ_{ij} , not λ_{AB} : see eqn (3) and Appendix S1 section 3), but any prior function may be specified by the user (see the main *diversitree* tutorial). The ‘slice sampling’ MCMC algorithm (Neal, 2003) is used by default and is fairly insensitive to tuning parameters. In particular, specifying a too large or too small value for the width of the proposal step (`w`) just increases the mean number of function evaluations per step, rather than the rate of mixing of the chain.

The marginal distributions of both the monogamy and sociality main effects on speciation rates are negative over the bulk of their distribution (Fig. 3). However, in contrast with treating the traits separately using BiSSE (Fig. 3a), we find that the 95% credibility intervals for both traits do not include zero (Fig. 3b). Therefore, these results support the conclusions of Magnuson-Ford & Otto (2012) that both monogamy and sociality are associated with decreased speciation rates in primates. Surprisingly, simultaneously accounting for both traits increased our confidence levels, suggesting that incorporating

additional traits can reduce noise caused by shifts in diversification because of other traits.

More comprehensive examples are included in a tutorial document on the *diversitree* website, <http://www.zoology.ubc.ca/prog/diversitree>, as well as within the online help for the package.

Closing comments

The *diversitree* package implements several methods for jointly modelling character evolution and speciation. The package is open source and designed to be fairly straightforward to extend. In particular, any model that can be expressed by moving down a tree (post-order traversal, or ‘pruning’; Felsenstein 1981) can be implemented using only a modest number of lines of R code. To facilitate the development of related methods, there is a ‘writing *diversitree* extensions’ manual available from the *diversitree* website. Stable versions of *diversitree* are available on CRAN (the Comprehensive R Archive Network) and from the website above. Development can be followed or joined on github (<http://github.com/richfitz/diversitree>).

I hope that the package will enable users to test a wide variety of macroevolutionary questions. However, I will close with a caution. All included methods are correlative only (Maddison, Midford & Otto 2007; Losos, 2011); they can merely show a statistical association between traits and speciation or extinction rates and cannot prove that the trait does affect speciation or extinction. Any unconsidered trait that is correlated with the target trait could be causal (Maddison, Midford & Otto 2007; Fig. 2). Alternatively, the associations may be spurious, perhaps driven by departures from the assumed model of cladogenesis or character evolution. There is currently no way of testing absolute goodness-of-fit with any

method, and all conclusions should be recognised as being conditional on a particular model and on that model being appropriate.

Acknowledgements

Sally Otto provided extensive support, feedback and comments on diversitree and on this manuscript. For general comments and discussions around the development of diversitree, I thank Emma Goldberg, Wayne Maddison, Karen Magnuson-Ford, Itay Mayrose, Arne Mooers, Brian O'Meara, Dan Rabosky, Stacey Smith and the users who have contacted me with comments, questions and bug reports. Karen Magnuson-Ford and Sally Otto contributed 'BiSSE-ness', Emma Goldberg contributed 'GeoSSE' and 'ClaSSE', and Wayne Maddison developed the interface with Mesquite. Emmanuel Paradis developed the ape package on which diversitree depends and uses. I thank Luke Harmon, Carl Boettiger, Graham Slater and an anonymous reviewer for suggestions that improved the manuscript. This work was supported by a University Graduate Fellowship from the University of British Columbia and a Vanier Commonwealth Graduate Scholarship from NSERC to R.G.F., and an NSERC discovery grant to Sarah P. Otto.

References

- Alfaro, M.E., Santini, F., Brock, C., Alamillo, H., Dornburg, A., Rabosky, D.L., Carnevale, G. & Harmon, L.J. (2009) Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences, USA*, **106**, 13410–13414.
- Anacker, B.L., Whittall, J.B., Goldberg, E.E. & Harrison, S.P. (2010) Origins and consequences of serpentine endemism in the California flora. *Evolution*, **65**, 365–376.
- 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.
- Bollback, J.P. (2006) SIMMAP: Stochastic character mapping of discrete traits on phylogenies. *BMC Bioinformatics*, **7**, 88.
- Burbrink, F.T., Ruane, S. & Pyron, R.A. (2012) When are adaptive radiations replicated in areas? Ecological opportunity and unexceptional diversification in West Indian dipsadine snakes (Colubridae: Alsophiini). *Journal of Biogeography*, **39**, 465–475.
- 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, Series B Biological Sciences*, **279**, 1567–1574.
- Eastman, J.M., Alfaro, M.E., Joyce, P., Hipp, A.L. & Harmon, L.J. (2011) A novel comparative method for identifying shifts in the rate of character evolution on trees. *Evolution*, **65**, 3578–3589.
- Felsenstein, J. (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution*, **17**, 368–376.
- Felsenstein, J. (1985) Phylogenies and the comparative method. *The American Naturalist*, **125**, 1–15.
- FitzJohn, R.G. (2010) Quantitative traits and diversification. *Systematic Biology*, **59**, 619–633.
- FitzJohn, R.G., Maddison, W.P. & Otto, S.P. (2009) Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Systematic Biology*, **58**, 595–611.
- Frijo, M. & Johnson, S.G. (2005) The design and implementation of FFTW3. *Proceedings of the IEEE*, **93**, 216–231.
- Goldberg, E.E. & Igić, B. (in press) Tempo and mode in plant breeding system evolution. *Evolution*, DOI: 10.1111/j.1558-5646.2012.01730.x.
- Goldberg, E.E., Lancaster, L.T. & Ree, R.H. (2011) Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Systematic Biology*, **60**, 451–465.
- Gómez, J.M. & Verdú, M. (2012) Mutualism with plants drives Primate diversification. *Systematic Biology*, **61**, 567–577.
- Hindmarsh, A.C., Brown, P.N., Grant, K.E., Lee, S.L., Serban, R., Shumaker, D.E. & Woodward, C.S. (2005) SUNDIALS: Suite of Nonlinear and Differential/Algebraic Equation Solvers. *ACM Transactions on Mathematical Software*, **31**, 363–396.
- Johnson, M.T.J., FitzJohn, R.G., Smith, S.D., Rausher, M.D. & Otto, S.P. (2011) Loss of sexual recombination and segregation is associated with increased diversification in evening primroses. *Evolution*, **65**, 3230–3240.
- Kass, R.E. & Raftery, A.E. (1995) Bayes factors. *Journal of the American Statistical Association*, **90**, 773–795.
- Lewis, P.O. (2001) A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology*, **50**, 913–925.
- Losos, J. (2011) Seeing the forest for the trees: the limitations of phylogenies in comparative biology. *The American Naturalist*, **177**, 709–727.
- Maddison, W.P. (2006) Confounding asymmetries in evolutionary diversification and character change. *Evolution*, **60**, 1743–1746.
- Maddison, W.P. & Maddison, D.R. (2008) Mesquite: a modular system for evolutionary analysis. Version 2.5. <http://www.mesquiteproject.org>.
- Maddison, D.R., Swofford, D.L. & Maddison, W.P. (1997) NEXUS: an extensible file format for systematic information. *Systematic Biology*, **46**, 590–621.
- 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.
- Magnuson-Ford, K. & Otto, S.P. (2012) Linking the investigations of character evolution and species diversification. *American Naturalist*, DOI: 10.1086/666649.
- Neal, R.M. (2003) Slice sampling. *The Annals of Statistics*, **31**, 705–767.
- Nee, S., May, R.M. & Harvey, P.H. (1994) The reconstructed evolutionary process. *Philosophical Transactions of the Royal Society of London Series B*, **344**, 305–311.
- Pagel, M. (1994) Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proceedings of the Royal Society of London Series B*, **255**, 37–45.
- Paradis, E., Claude, J. & Strimmer, K. (2004) APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*, **20**, 289–290.
- R Development Core Team. (2012) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Rabosky, D.L. & Glor, R.E. (2010) Equilibrium speciation dynamics in a model adaptive radiation of island lizards. *Proceedings of the National Academy of Sciences, USA*, **51**, 22178–22183.
- Read, A.F. & Nee, S. (1995) Inference from binary comparative data. *Journal of Theoretical Biology*, **173**, 99–108.
- Redding, D.W., DeWolf, C. & Mooers, A.Ø. (2010) Evolutionary distinctiveness, threat status and ecological oddity in primates. *Conservation Biology*, **24**, 1052–1058.
- Rowan, T. (1990) *Functional Stability Analysis of Numerical Algorithms*. PhD thesis, University of Texas at Austin.
- Schluter, D. (2009) Evidence for ecological speciation and its alternative. *Science*, **323**, 737–741.
- Schluter, D., Price, T., Mooers, A.Ø. & Ludwig, D. (1997) Likelihood of ancestor states in adaptive radiation. *Evolution*, **51**, 1699–1711.
- Soetaert, K., Petzoldt, T. & Setzer, R.W. (2010) Solving differential equations in R: Package deSolve. *Journal of Statistical Software*, **33**, 1–25.
- Wilson, A.W., Binder, M. & Hibbett, D.S. (2011) Effects of gasteroid fruiting body morphology on diversification rates in three independent clades of fungi estimated using binary state speciation and extinction analysis. *Evolution*, **65**, 1305–1322.
- Winger, B.M., Lovette, I.J. & Winker, D.W. (2012) Ancestry and evolution of seasonal migration in the Parulidae. *Proceedings of the Royal Society of London Series B*, **279**, 610–618.

Received 8 March 2012; accepted 18 June 2012

Handling Editor: Luke Harmon

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Appendix S1. Tuning diversitree, fast likelihood calculations, and full code for the primate example.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials may be reorganised for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.