

APPLICATION

RBrownie: an R package for testing hypotheses about rates of evolutionary change

J. Conrad Stack^{1*}, Luke J. Harmon² and Brian O'Meara³

¹Department of Biology, The Pennsylvania State University, University Park, PA 16802, USA; ²Department of Biological Sciences, University of Idaho, Moscow, ID, USA; and ³Department of Ecology and Evolutionary Biology, The University of Tennessee, Knoxville, TN, USA

Summary

1. Maximum likelihood analyses for testing hypotheses about how rates of disparification might vary across clades can provide important insight into the evolutionary process. While the Brownie phylogenetic library can perform such analyses, it does so outside of a general scripting environment.
2. We present RBrownie, an interface between the Brownie phylogenetic library and the R software environment, which provides easy access to the main methods in Brownie (see O'Meara 2008; PhD Dissertation, Nature Precedings), including discrete ancestral state reconstruction. In addition, RBrownie supplies a direct interface to Brownie, allowing advanced users to construct more complex combinations of analyses and to execute any newly added Brownie functions.
3. Overall, it is a package that features evolutionary rate analyses in a flexible and familiar environment.

Key-words: bioinformatics, comparative analysis, evolutionary biology, phylogenetics, reconstruction, software

Introduction

There are many potential factors that can affect rates of phenotypic evolution ('disparification'; Ackerly 2009) over time. For example, changes in life-history and environmental traits have both been suggested to be significant contributors to changes in evolutionary rates. A classic example is the transition in locomotion from legs to wings in Theropods, which may have accelerated the evolution of leg shape and size by removing constraints to their function (Gatesy & Middleman 1997). Methods for testing hypotheses about whether such trait changes have affected evolutionary rates are available in a C++ program called Brownie (O'Meara *et al.* 2006; O'Meara 2008) and, to some extent, in various R scripts (see Supporting Information). RBrownie makes these and other methods available in the R software environment by acting as an interface layer between R and the Brownie C++ library. It is available through the Comprehensive R Archive Network (CRAN), the package management system native to R. In addition to making Brownie methods available to the wider R audience, it is one of the first extensions of the phylogenetic package Phylobase (Hackathon *et al.* 2010). Phylobase data structures are extended to include support for character

mapping on to phylogenies, specification of taxa sets and tree weights, and novel plotting functions for viewing the results of the evolutionary rate tests (Fig. 1). Below, we describe the software in more detail and give examples of its application. Links to tutorials and further information on this software are provided in the Supporting Information.

Description

TESTING CONTINUOUS RATES

One of the major functions in the Brownie library is testing hypotheses about rates of evolution of continuously valued characters. The method for testing hypotheses about rates is discussed in detail by O'Meara *et al.* (2006). Briefly stated, these methods test for the presence of evolutionary rate shifts on phylogenetic trees. As an example, Collar *et al.* (2009) used Brownie to test the hypothesis that the development of piscivory (a discrete, binary trait) in centrarchid fishes had a limiting effect on the morphological diversification of key parts of the skull and jaw.

Brownie allows the user to first estimate the rate of continuous trait evolution for groups defined *a priori* and then test whether the inferred evolutionary rates for each group are meaningfully different. A group can be a clade, or it can be all the branches in a phylogeny for which the

*Correspondence author. E-mail: stack@psu.edu
Correspondence site: <http://www.respond2articles.com/MEE/>

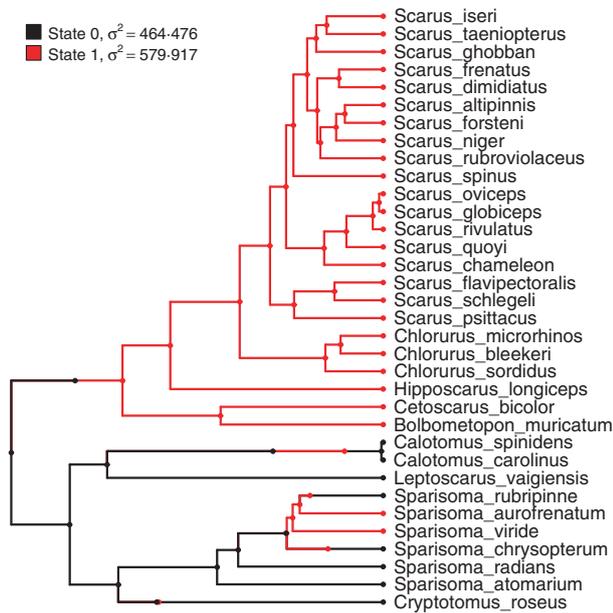


Fig. 1. Non-censored rate test applied to parrot fish jaw morphology. The legend in the upper left shows the inferred rates of jaw evolution under Brownian motion for browsers (state 0) and excavators/scrapers (state 1).

mapped character trait has a certain value (e.g. all lineages that are assumed to have the piscivory trait). Currently, there are two options for modelling the evolutionary dynamics of a continuous trait under observation: Brownian motion (BM) and the Ornstein–Uhlenbeck process (OU). Further, Brownie implements two approaches for fitting either of these continuous trait evolution models: the non-censored approach (BM and OU) and the censored approach (BM only) (in RBrownie, see the `runNonCensored` and `runCensored` commands, respectively). The two approaches differ in how they treat the branches along which group membership is inferred to have changed.

For either Brownian motion or Ornstein–Uhlenbeck processes, the null hypothesis of both the censored and the non-censored approaches constrains all branches to the same set of parameter values. The alternative hypothesis for the non-censored approach constrains the rate (BM) or mean (OU) parameter to be the same for all branches within each user-specified group, while in the censored approach, the branches on which group changes occur are deleted, and evolutionary parameter values are inferred for each resulting subtree independently. Group membership is assigned either explicitly by a character state that is mapped onto each branch (Huelsenbeck, Nielsen & Bollback 2003) or implicitly by whether or not the branch is included in a subtree which only comprises a subset of all the taxa.

In RBrownie, character states are handled as attributes assigned to each branch of a tree. When character traits change along a branch, that branch is divided into sub-branches that indicate where along the branch the change took place. Character states can be added to trees in R directly through accessor functions, or they can be read in from branch-annotated nexus

tree files (Maddison, Swofford & Maddison 1997; Bollback 2006; Supporting Information) similar to those used by mesquite (Maddison & Maddison 2010), allowing a variety of methods to be used to assign state changes to branches. Similarly, RBrownie allows users to define sets of taxa (for implicitly assigning group membership) through accessor functions, or they can be read in from an optional ASSUMPTIONS block within the nexus file. Once group membership and the approach (censored or non-censored) have been specified, maximum likelihood is used to fit a continuous time model to the data under both sets of constraints (the null and the alternative). Finally, the likelihood values that are returned may be compared.

It is important to note that RBrownie is not constrained to operate on a single tree or a single morphological character. For example, when presented with multiple trees, the user can iterate analyses over all possible combinations of trees and/or characters and the same censored or non-censored analysis can be performed on each, giving an indication of the consistency of the results with regard to phylogenetic uncertainty.

Discrete character evolution

When tips are labelled with discrete data, RBrownie can estimate the ancestral states of those data given a particular tree or set of trees (see the `runDiscrete` command) following the method described in O’Meara (2008; Chapter 3). The design is flexible allowing users to arbitrarily specify rate matrices and state frequency vectors and supports constraining certain elements of the rate matrix to be equal. It also allows state changes to occur explicitly along branches and not only at internal nodes. The number of changes that occur along a branch connecting two nodes can be given an upper limit by the user. Trees are returned with the ancestral states mapped onto the branches, using sub-branches where needed, and flexible plotting functions are provided for visualizing where changes occurred on each branch.

Conclusion

RBrownie makes the Brownie C++ library available to R users, providing access to useful analyses for addressing evolutionary rate questions. Building off of phylobase data structures (Hackathon *et al.* 2010), RBrownie adds support in R for taxa sets and character mapping on to phylogenies and provides flexible methods for visualization of these new structures. This provides a natural foundation for handling phylogenetic and morphologic data and returning results of non-censored and censored rate tests and discrete ancestral character reconstructions. While accessor functions like `runDiscrete` do not exist for each function in the Brownie library, Brownie input files that use newer Brownie methods like species delimitation (O’Meara 2010) can be constructed and executed in RBrownie and will be supported through accessors in the future. This access to the Brownie library also makes it easy for the advanced end user to construct and execute complicated combinations of analyses and makes future

expansions of the Brownie library easier to implement. RBrownie provides an important addition to R's growing phylogenetic toolkit.

Acknowledgements

J.C.S. and B.O.M. were supported by a grant from the Google Summer of Code 2010 program, which was administered by the National Evolutionary Synthesis Center. L.J.H. was supported by NSF DEB 0919499.

References

- Ackerly, D. (2009) Conservatism and diversification of plant functional traits: evolutionary rates versus phylogenetic signal. *Proceedings of the National Academy of Sciences*, **106**(Suppl. 2), 19699–19706.
- Bollback, J.P. (2006) Simmap: stochastic character mapping of discrete traits on phylogenies. *BMC Bioinformatics*, **7**, 88.
- Collar, D.C., O'Meara, B.C., Wainwright, P.C. & Near, T.J. (2009) Piscivory limits diversification of feeding morphology in centrarchid fishes. *Evolution*, **63**, 1558–5646.
- Gatesy, S.M. & Middleman, K.M. (1997) Bipedalism, flight, and the evolution of theropod locomotor diversity. *Journal of Vertebrate Paleontology*, **17**, 308–329.
- Hackathon, R., Bolker, B., Butler, M., Cowan, P., de Vienne, D., Eddelbuettel, D., Holder, M., Jombart, T., Kembel, S., Michonneau, F., Orme, D., O'Meara, B., Paradis, E., Regetz, J. & Zwickl, D. (2010) Phylobase. R package version 0.6.2. <http://CRAN.R-project.org/package=phylobase>.
- Huelsenbeck, J.P., Nielsen, R. & Bollback, J.P. (2003) Stochastic mapping of morphological characters. *Systematic Biology*, **52**, 131–158.
- Maddison, W.P. & Maddison, D. (2010) Mesquite: a modular system for evolutionary analysis. Version 2.73. <http://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.
- O'Meara, B.C. (2008) Using trees: myrmecocystus phylogeny and character evolution and new methods for investigating trait evolution and species delimitation (PhD Dissertation). Available from *Nature Precedings*.
- O'Meara, B.C. (2010) New heuristic methods for joint species delimitation and species tree inference. *Systematic Biology*, **59**, 59–73.
- O'Meara, B.C., Ané, C., Sanderson, M.J. & Wainwright, P. (2006) Testing for different rates of continuous trait evolution using likelihood. *Evolution*, **60**, 922–933.

Received 31 December 2010; accepted 15 March 2011

Handling Editor: Robert P. Freckleton

Supporting Information

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

Data S1. Information on how to obtain, install and use RBrownie.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials may be re-organized 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.