Phylogenetics

MATICCE: mapping transitions in continuous character evolution
Andrew L. Hipp1,∗ and Marcial Escudero2
1Herbarium, The Morton Arboretum, 4100 Illinois Route 53, Lisle IL 60532-1293, USA and 2Department of Molecular Biology and Biochemical Engineering, Pablo de Olavide University, Ctra. Utrera km 1, 41013 Sevilla, Spain

Received on September 15, 2009; revised on October 27, 2009; accepted on October 28, 2009
Advance Access publication October 30, 2009
Associate Editor: Martin Bishop

ABSTRACT
Summary: MATICCE is a new software package in the R language for mapping phylogenetic transitions in organismal traits that have continuous distributions. MATICCE integrates over phylogenetic and model uncertainty and provides simulation functions for visualizing evolutionary scenarios based on estimated parameter values.
Availability and Implementation: MATICCE is written in the open source R language and freely available through the Comprehensive R Archive Network (http://cran.r-project.org/web/packages/maticce).
Contact: ahipp@mortonarb.org.

1 INTRODUCTION
Inferring the evolutionary history of species traits is an issue of practical concern for any researcher testing biological hypotheses using phylogenetic data. Maximum likelihood and Bayesian approaches are commonly used to map phylogenetic transitions in the evolution of organismal traits that can be described in terms of discrete states (Huelsenbeck et al., 2003; Pagel, 1999). However, there is no standardized approach for mapping phylogenetic transitions in the distribution of continuous traits of ecological and evolutionary relevance such as body size and shape. Several models have been introduced to describe phylogenetic transitions in the evolutionary dynamics of such continuous traits, given an a priori specification of where on the phylogeny those transitions may have occurred (e.g. Butler and King, 2004; Hansen et al., 2008; O’Meara et al., 2006; Revell and Collar, 2009). A generalized framework is needed for identifying where on a phylogeny such transitions may have occurred as a guide to hypothesis testing and to evaluate whether a priori transitions in continuous character distribution are supported as strongly as other possible transitions.

In this article, we describe a new software package for mapping phylogenetic transitions in the evolution of any continuous character on a molecular or morphological phylogeny with branch lengths scaled proportional to time (i.e. ultrametric). The software, MATICCE, implements an information-theoretic approach to quantifying the statistical support for continuous character transitions. The software is written in the cross-platform R statistical environment (R Development Core Team, 2009) and integrates with the phylogenetic comparative methods in OUCH (Butler and King, 2004) and APE (Paradis et al., 2004). We describe the approach and implementation in this article. Worked examples and sample data are available in the online documentation and vignette.

2 DESCRIPTION
MATICCE models transitions in continuous character distribution according to an Ornstein–Uhlenbeck (O–U) process (Butler and King, 2004; Martins and Hansen, 1997), utilizing likelihood calculations implemented in OUCH. Under an O–U process, a continuous character is modeled as evolving stochastically toward a stationary distribution with mean θ and (at stationarity) variance σ2/2α, where α determines the rate of evolution toward the stationary distribution (Butler and King, 2004). The approach implemented in MATICCE evaluates the relative support for alternative models of continuous character distribution shifts by (i) specifying n nodes at which a change may have occurred, (ii) specifying and evaluating support for models that allow all permutations of change at those nodes up to a maximum number of nodes defined by the user, and (iii) estimating the relative support for a shift in character distribution at each node as the cumulative information criterion (e.g. small-sample Akaike information criterion, AICc, or Bayes Information criterion, BIC) weight for all models entailing a shift at that node. The method is reminiscent of stepwise AIC model selection (e.g. Alfaro et al., 2009) but, by taking a model-averaging approach, avoids sensitivity to AIC significance thresholds.

MATICCE makes it easy to generate the potentially large number of model specifications required to test character transition hypotheses and provides a set of tools for flexibly defining nodes for analysis. To relax computational limitations, users can limit the maximum number of nodes at which a character is allowed to change distribution (e.g. 30 candidate nodes could be investigated for models allowing up to a maximum of three character distribution transitions; this is a manageable 4526 models, compared with the 1.07E09 models defined by all possible permutations of 30 candidate nodes). Model specifications are stored within each MATICCE analysis object alongside analysis results and viewable in OUCH or the simulation functions of MATICCE. MATICCE performs and summarizes analyses over sets of trees, integrating over phylogenetic and model uncertainty and allowing analyses to be performed even for nodes that are not found in all trees in a set. Because sets of phylogenetic trees (e.g. bootstrap sets) are not typically identical in topology, each node is defined by the set of taxa descendent from it. MATICCE utilizes a user-defined list of taxa

∗To whom correspondence should be addressed.
MATICCE provides straightforward functions for summarizing analyses over large sets of models and trees, using BIC, AIC or AICc weights to estimate model support and model-averaged parameter values (Burnham and Anderson, 2002). Analyses are summarized for each node both over all trees and over just the trees in which the node occurs (Fig. 1A), so that the evidential support for a shift in character distribution at a given node may be interpreted as the relative support for that particular shift conditioned on the existence of the node. Based on these analyses, users will often want to evaluate the relative support for different models of character evolution at a particular node. MATICCE provides functions for evaluating and summarizing the relative support for alternative models at a single node. These functions lend themselves to customization and can easily accommodate additional models as desired.

3 CONCLUSIONS

While there has been a substantial increase in statistical analysis of continuous character evolution in the past several years, an easy-to-use general framework for reconstructing shifts in continuous character evolution across a phylogeny has been lacking. MATICCE uses existing tools to implement such a framework, providing a needed approach for exploratory analysis of continuous character evolution on a phylogeny.

ACKNOWLEDGEMENTS

We thank A. King, M. Butler, A. Platt, B. O’Meara and two anonymous reviewers for helpful comments on this work.

Funding: NESCent visiting scholar award (to A.L.H.); the National Science Foundation (0743157 to A.L.H.); the Spanish Government (FPU AP2005–3715 to M.E.).

Conflict of Interest: none declared.

REFERENCES


