**Picante: R tools for integrating phylogenies and ecology**

Steven W. Kembel1,*, Peter D. Cowan2, Matthew R. Helmus3, William K. Cornwell1, Helene Morlon5, David D. Ackerly2, Simon P. Blomberg6 and Campbell O. Webb7

1Center for Ecology and Evolutionary Biology, University of Oregon, Eugene, OR, 2Department of Integrative Biology, University of California, Berkeley, CA, USA, 3Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Kunming, Yunnan, China, 4Biodiversity Research Centre, University of British Columbia, Vancouver, BC, Canada, 5Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA, USA, 6Faculty of Biological and Chemical Sciences, University of Queensland, Brisbane, Australia and 7Arnold Arboretum of Harvard University, Cambridge, MA, USA

*To whom correspondence should be addressed.

**ABSTRACT**

**Summary:** Picante is a software package that provides a comprehensive set of tools for analyzing the phylogenetic and trait diversity of ecological communities. The package calculates phylogenetic diversity metrics, performs trait comparative analyses, manipulates phenotypic and phylogenetic data, and performs tests for phylogenetic signal in trait distributions, community structure and species interactions.

**Availability:** Picante is a package for the R statistical language and environment written in R and C, released under a GPL v2 open-source license, and freely available on the web (http://picante.r-forge.r-project.org) and from CRAN (http://cran.r-project.org).

**Contact:** skembel@uoregon.edu

Received on February 16, 2010; revised on April 9, 2010; accepted on April 12, 2010

**1 INTRODUCTION**

Incorporating phylogenetic information into ecology is enlightening because it allows ecological questions to be addressed in an evolutionary context, leading to a deeper understanding of the processes that give rise to patterns of biological diversity (Webb et al., 2002). Ecophylogenetic research also has numerous applications such as designating conservation priorities, predicting species invasiveness and estimating the effects of biodiversity on ecosystem functioning (Cavender-Bares et al., 2009).

In this rapidly expanding field, two barriers have prevented ecologists from incorporating phylogenetic information into their own work. First, many ecophylogenetic methods are not available in open-source software packages, preventing end-users from modifying methods to suit their own purposes. Second, the available tools are implemented in multiple software packages, each with unique learning curves and data formatting requirements. Here, we describe a software package designed to overcome these limitations by implementing the most commonly used ecophylogenetic tools in a single open-source software package for the R statistical language and environment (R Development Core Team, 2009).

**2 DESCRIPTION**

**2.1 Data manipulation and visualization**

Picante uses the extensive data manipulation and data import/export functions included in R. Picante includes functions for reading and writing ecological data in formats used by other software packages including Phylocom (Webb et al., 2008b), allowing these tools to be integrated into data analysis workflows in the R environment. By using the basic R plotting functions along with specifically designed functions in picante and other R packages (e.g. ape: Paradis et al., 2004), phylogenetic patterns of trait evolution and community structure can easily be visualized.

**2.2 Phylogenetic diversity and community structure**

With the growing availability of phylogenies for entire ecological communities, there has been a rapid increase in the development and application of methods to quantify phylogenetic diversity. Picante contains functions for calculating the most commonly used measures of phylogenetic α-diversity (evolutionary relatedness within ecological communities) including Faith’s phylogenetic diversity (PD, the sum of all branch lengths separating taxa in a community; Faith 1992), the taxonomic distinctness index (Vane-Wright et al., 1991), the mean pairwise phylogenetic distance (MPD) and mean nearest taxon distance (MNTD) within a community (Webb et al., 2008b), the phylogenetic species diversity metrics of Helmus et al. (2007b) [phylogenetic species richness (PSR), evenness (PSE), variance (PSV), and clustering (PSC)], quadratic entropy (Rao, 1982), and species co-occurrence and phylogenetic relatedness regressions (Slingby and Verboom, 2006). Picante also calculates widely used measures of phylogenetic β-diversity (evolutionary relatedness between communities), including the phylogenetic Sørenson index (Bryant et al., 2008), the UniFrac distance metric (Lozupone and Knight, 2005), and MPD and MNTD between communities (Webb et al., 2008b). The availability of this suite of tools in a single package makes it easy to conduct comparative analyses of phylogenetic diversity on the same community dataset, which is important since different diversity metrics are sensitive to different aspects of the structure of ecological communities (Cadotte et al., 2010).

Measures of phylogenetic α- and β-diversity can be compared to null expectations of community assembly and evolution to estimate...
community structure (i.e. whether communities are composed of closely related species or distantly related species). Picante performs numerous permutation procedures not available in other software packages including generating random communities by drawing taxa from a species pool, rearrangements of the community matrix preserving sample species richnesses and species occurrence frequencies and randomizations of phylogenies. When used with the various phylogenetic diversity metrics, these permutations test different null hypotheses of community structure and allow comparisons of phylogenetic diversity among communities that vary in species richness (Kembel, 2009).

2.3 Environmental correlates of phylogenetic community structure

Picante implements several methods to test for environmental drivers of phylogenetic community structure. First, picante calculates sample-based rarefaction curves of PSR (Helmus et al., 2007b), allowing statistically valid comparisons of phylogenetic diversity among communities that differ in sampling intensity, size, area, etc. (Gotelli and Colwell, 2001). Second, picante performs species-environmental regressions (Helmus et al., 2007a) to test if closely related species respond to environmental variation in the same way and to separate determinants of phylogenetic community structure. Third, phylogenetic β-diversity metrics can be used by methods in other R packages (e.g. vegan: Oksanen et al., 2009) to perform multivariate analyses of community-environment relationships including Mantel tests, phylogenetic clustering of communities and phylo-ordinations (Webb et al., 2008a).

2.4 Phylogenetic signal in species traits and interactions

Phylogenetic signal is the tendency for taxa that share a common ancestor to resemble each other in how they look, behave and with whom they interact. Phylogenetic signal in the niches and traits of species violates the assumption that data points are statistically independent; techniques such as phylogenetically independent contrasts are available to overcome this non-independence (Paradis et al., 2004). However, methods to quantify the amount of phylogenetic signal in phenotypic traits are not widely available. Picante performs two kinds of statistical analyses of phylogenetic signal. The first is the calculation of the K statistic (Blomberg et al., 2003), a Brownian motion-based metric of the strength of phylogenetic signal. Statistical significance of $K$ is estimated by permuting trait values across the tips of a phylogenetic tree. The second implemented analysis is a phylogenetic regression of interaction strengths between species (e.g. plants and pollinators: Ives and Godfray, 2006) using an Ornstein–Uhlenbeck evolutionary interaction strengths between species (e.g. plants and pollinators: Ives and Godfray, 2006) using an Ornstein–Uhlenbeck evolutionary process to model trait evolution along the branch lengths of a phylogeny. By integrating species richness, abundance and evolutionary history, this method estimates the relative contribution of traits versus phylogeny in explaining species interactions.

**Funding:** Chinese Academy of Sciences Research Fellowship for International Young Researchers (to M.R.H.); Google Summer of Code (to P.D.C. and M.R.H.); Gordon and Betty Moore Foundation (to S.W.K.); National Center for Ecological Analysis and Synthesis (to S.W.K. and M.R.H.); Natural Sciences and Engineering Research Council of Canada (to S.W.K.); US National Science Foundation (to M.R.H. and C.O.W.).

**Conflict of Interest:** none declared.

**REFERENCES**


