Title Page

pez: Phylogenetics for the Environmental Sciences

Short Structured Abstract

Summary

Pez is an R package that calculates ¿30 community phylogenetic metrics, statistically models phylogenetic, trait, and community data, and simulates community structure. It provides a common programmatic standard for the management and manipulation of eco-phylogenetic data in R.

Availability

' Pez is released under the GPL v3 open-source license, available on the Internet from CRAN (`http://cran.r-project.org`). The package is under active development, and the authors welcome contributions (see `http://github.com/willpearse/pez`).

Contact

William D. Pearse (`wdpearse@umn.edu`) Department of Ecology, Evolution, and Behavior, University of Minnesota, 100 Ecology Building, 1987 Upper Buford Circle, Saint Paul, Minnesota, 55108, USA.
Marc Cadotte XXX
Caroline Tucker XXX
Steve Walker XXX
Matthew R. Helmus (`mrhelmus@gmail.com`) Department of Animal Ecology, Vrije Universiteit, 1081 HV, Amsterdam, The Netherlands

Supplmentary information

A vignette (available with the package) showing some of the package's functions, and briefly describing the classification of community phylogenetic metrics.

Text

Introduction

Community phylogenetics is a subfield of ecology which links ecological phenomena (e.g., competition, ecosystem functioning) with the evolutionary processes that generate species and their traits (**Webb2002**; **Cavender-Bares2**). This growing field has generated a number of statistical tools, mostly metrics of phylogenetic biodiversity, and software code and packages to implement such statistics (**Webb2008**; **Regetz2009**; **Kembel2010**; **Orme2013**; **Eastman2013**). The code is disparate and handles data differently making routine data analyses with different statistcs by the end user challenging. In many cases, the statistics are also not formally implmeneted in a particular software package and are available only as supplementary materials to papers which are not maintained or have public source code. Without active maintenance, these valuable statistics are effectively lost to the scientific community.

Pez provides an R class (*comparative.comm*) that integrates phylogenetic, community, environmental and trait data in as a single object. From one of these eco-phylogenetic data objects, one can calculates a large body of phylogenetic biodiversity metrics (many previously unavailable in R). The implemented metrics are thematically organized according to the framework outlined in **Pearse2014review** allowing the user to better understand the actual aspects of phylogenetic biodiversity they are calculating with any particular metric. As the structure of the package has been written to be flexible and intuative, newly published metrics will continue to be incorporated in *Pez*.

Further, *Pez* implements and extends the regression framework presented by (**Cavender-Bares2004**) (**Cavender-Bares2006**). Finally, it provides functions that simulate ecological communities to exhibit varying degrees of phylogenetic and trait community structure. Together, the functions we have implemented in *Pez* will facilitate and speed the use of analysis of eco-phylogenetic datasets, and provide a common R framework for community phylogenetics.

Function	Description
comparative.comm	Stores community, phylogenetic, environmental, and species trait data
shape	Calcuates PSV^* (**Helmus2007**), PSR^* (**Helmus2007**), PD^* & MPD^* (**Faith1992**), Colless' Index* (**Colless1982**), γ^* (**Pybus2000**), Δ^* (**Warwick1995**), E_{ED} & H_{ed} (**Cadotte2010**), phylo-eigenvectors* (**Diniz-Filho2011**)
evenness	Calculates Δ (**Warwick1995**), Phylogenetic Entropy (**Allen2009**), PAE, IAC, H_{aed}, & E_{aed} (**Cadotte2010**), Rao's quadratic entropy* (**Rao1982a**), λ (**Pagel1999**), δ (**Pagel1999**), κ (**Pagel1999**), Simpson's XXX (XXX)
dispersion	SES_{MPD}/NRI^* (**Webb2000**; **Webb2002**; **Kembel2009**), SES_{MNTD}/NTI^* (**Webb2000**; **Webb2002**; **Kembel2009**), $INND$ (**Ness2011**), D^* (**Fritz2010**)
dissimilarity	UniFrac (**Lozupone2005**), PCD (**Helmus2010**), PhyloSor (**Bryant2008**), Rao's Q (**Rao1982a**)
fingerprint.regression	Compares phylogenetic, community co-existence, and trait similarity matrices using Mantel tests, quantile regressions, or linear models following (**Cavender-Bares2004**; **Cavender-Bares2006**). Completely re-implements all features previously available in *EcoPhyl* (XXX)
scape	Simulates community phylogenetic structure across a landscape, simulating phylogenetic repulsion, attraction, niche width, and range size (Helmus2012)
trait.asm	Calculate optimal traitgram that explains community data, following XXX
Others?	classic.phy.signal (worth keeping?) - some stuff from willeerd?

Table 1: Some of the functions available in *pez*. * indicates a metric whose code is based on existing R code. Note that calculation of Pagel's λ (**Pagel1999**), δ (**Pagel1999**), κ (**Pagel1999**) are calculated using the *pgls* function in *caper* (**Orme2013**).

Description

Data manipulation and storage

Pez extends the data manipulation and import/export features of R, and contains a unified class to contain all eco-phyogenetic data. It simplifyies the process of ensuring community, trait, environmental, and phylogenetic data are compatible by overloading R's standard row and column operators. This makes it simple to make consistent changes across datasets; for example, removing a species from the community data automatically trims it from the phylogeny, and removes its associated trait data. *Pez* was designed to be backwards-compatible with *caper*'s (**Orme2013**) *comparative.data* class, seamlessly permitting detailed comparative analysis of species trait data.

Metrics

Following the classification of **Pearse2014review** *pez* simplifies the calculation and comparison of a number of metrics by grouping them into four categories: *shape*, *evenness*, *dispersion*, and *dissimilarity*. Shape metrics measure the structure of an community phylogeny. Evenness metrics are shape metrics that incorporate a weighting with species' abundances. Dispersion metrics calculate the phylogenetic biodiversity expected given a source pool. Finally, dissimilarity differ from the former as they directly calculate the pairwise difference in phylogenetic biodiversity between assemblages. Table 1 gives an overview of the functions contained within *pez*, *trait.asm* permits the direct comparison of phylogenetic and trait structure, following the *traitgram* framework of XXX. Speeding and easing the comparison of community phylogenetic metrics is important, since each can reveal radically different aspects of the structure of ecological assemblages (**Cadotte2010**).

Models

Model-based approaches are, in many ways, an advance over metrics that simply quantify phylogenetic pattern in ecological assemblages. Models explain phylogenetic structure across a number of assemblages simultaneously, maximising statistical power, and can incorporate phylogenetic, environmental, trait, and other information. *Pez* contains a complete re-implentation of all routines previously available in *EcoPhyl* (XXX) that regress species co-existence matrices against environmental and, trait, and phylogenetic distance matrices

(**Cavender-Bares2004**; **Cavender-Bares2006**). It extends these routines to include, for example, more modern measures of phylogenetic signal (*e.g.*, λ, *delta*, κ, D), and to examine differences in relationships among clades.

Simulation

Finally, *pez* contains a number of functions that will simulate ecological communities under various null models of community assembly and/or evolution. *scape* is intended to model community assembly over a region, and models both species niche width and range. *metapop.sim* fits models of trait evolution and dispersal ability to a static landscape, evolving both a phylogeny, species traits, and the resulting distribution of species across the landscape. While these are works in progress, we are confident that they at least form a basis for further model-based exploration of data.

Acknowledgements