

SYNCSA—R tool for analysis of metacommunities based on functional traits and phylogeny of the community components

Vanderlei J. Debastiani* and Valério D. Pillar

Programa de Pós-Graduação em Ecologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS 91501-970, Brazil

Associate Editor: David Posada

ABSTRACT

Summary: SYNCSA is an R package for the analysis of metacommunities based on functional traits and phylogeny of the community components. It offers tools to calculate several matrix correlations that express trait-convergence assembly patterns, trait-divergence assembly patterns and phylogenetic signal in functional traits at the species pool level and at the metacommunity level.

Availability and implementation: SYNCSA is a package for the R environment, under a GPL-2 open-source license and freely available on CRAN official web server for R (<http://cran.r-project.org>).

Contact: vanderleidebastiani@yahoo.com.br

Received on March 5, 2012; revised on May 29, 2012; accepted on May 30, 2012

1 INTRODUCTION

Comprehending how biological communities are structured in both time and space is a major challenge for ecologists. It is well known that communities are not random samples from the species pool. Niche theory predicts that in order to persist in a local community, each species has to have a particular set of traits that allow it to overcome environmental filtering and to coexist with other species (Keddy and Weiher, 1999). Nonetheless, considering species traits in community ecology studies may be complicated as traits are not independent among species (Felsenstein, 1985). In the last decades, functional traits and phylogeny information have been incorporated in ecological models in order to determine the factors responsible for creating and maintaining community structure (Cavender-Bares *et al.*, 2009; Fine and Kembel, 2011; Kembel and Hubbell, 2006; Webb *et al.*, 2002). Currently, information about traits and phylogeny are available for several groups of organisms (see APG, 2009; Bininda-Emonds *et al.*, 2007; Jones *et al.*, 2008; Kattge *et al.*, 2011). Furthermore, recent developments made possible novel perspectives in functional and phylogenetic analyses of ecological data (Pillar and Duarte, 2010; Pillar *et al.*, 2009).

SYNCSA is an R package (R Development Core Team, <http://www.r-project.org/>) released under open-source license, and freely available from CRAN (<http://cran.r-project.org>), that implements the methods described in Pillar and Duarte (2010) and Pillar *et al.* (2009) for the comparison of ecological communities based on functional traits and the phylogeny of the community

components, the detection of phylogenetic signal at the species pool and metacommunity levels and the testing of phylogenetic niche conservatism. SYNCSA allows the identification of optimal trait subsets expressing trait-convergence assembly patterns or trait-divergence assembly patterns in metacommunities. The functions are based on the abovementioned methods and are complementary to other functions available in R for the analysis of communities, which take into account traits and phylogenetic information. SYNCSA for R implements the same methods available in recent versions of SYNCSA application written in C++ (<http://ecoqua.ecologia.ufgrs.br/ecoqua/SYNCSA>), which was developed by V. Pillar and evolved from methods described in Pillar and Orlóci (1993).

2 DESCRIPTION

2.1 Data manipulation

SYNCSA uses the same data manipulation as conventionally used in R. Four basic input data matrices (matrix notation follows Pillar and Duarte, 2010) may be entered in order to run all SYNCSA analyses, but at least two matrices are necessary: the community matrix and a matrix with either traits or phylogenetic information. Data matrices should be organized as follows: (i) community matrix, arranged with species as columns and sampling units as rows, which may contain either species presences/absences or abundances (matrix **W**). (ii) Trait matrix, with traits as columns and species as rows, which should contain a set of functional traits describing the species (matrix **B**). Both binary and/or quantitative traits may be included. (iii) The species phylogenetic pairwise dissimilarities (matrix **D_F**), such as the ones obtained in the software Mesquite (W.P. Maddison and D.R. Maddison, <http://mesquiteproject.org>) or the R package Ape (Paradis, 2004). (iv) The environmental data of interest, which may be one or more community drivers or ecosystem effects (matrix **E**), with variables in columns and communities in rows. When labels for both row and columns are available for all matrices, the function *organize.syncsa* will rearrange all matrices to be in the same order.

2.2 Fuzzy types, scaling-up and assembly patterns

SYNCSA offers 10 functions that can be used for data manipulation (organization and transformations), scaling-up of trait and phylogenetic data to the metacommunity level and matrix correlation. The scaling-up allows data exploration and hypothesis testing about trait-convergence and trait-divergence assembly patterns, phylogenetic signal and their relations to environmental

*To whom correspondence should be addressed.

variation (Pillar and Duarte, 2010; Pillar *et al.*, 2009). For revealing trait-convergence assembly patterns, community weighted means of each trait will be computed by matrix multiplication ($\mathbf{T} = \mathbf{WB}$). A significant matrix correlation $\rho(\mathbf{TE})$ between community resemblances based on traits ($\mathbf{D_T}$) and on environmental variables ($\mathbf{D_E}$) is evidence of trait-convergence assembly pattern at the metacommunity level. For revealing trait-divergence assembly patterns, an essential step for the scaling-up is the weighting of species presences or abundances by the species trait similarities. Analogously, the same approach is used for the phylogenetic data, in which case the weighting is based on species phylogenetic similarities. Each element in the resulting weighted matrix (\mathbf{X} for traits and \mathbf{P} for phylogeny) will indicate the probability that a given species will occur in a community based on its trait or phylogenetic similarity to the species that were found in the same community, and thus the matrix will convey to the community level the whole trait or phylogenetic information captured by the species similarity matrix. The matrix correlation $\rho(\mathbf{XE})$ measures the strength of the association between variation in community composition weighted by the species traits and variation in community environmental variables ($\mathbf{D_E}$). The partial matrix correlation $\rho(\mathbf{XE.T})$, which removes the effect of trait-convergence in \mathbf{T} , will indicate whether a trait-divergence assembly pattern is related to the environmental variation. Furthermore, matrix correlation $\rho(\mathbf{PE})$ measures the strength of the association between community distances based on their phylogenetic structure ($\mathbf{D_P}$) and environmental distances ($\mathbf{D_E}$). Moreover, matrices \mathbf{X} and \mathbf{P} may be submitted to ordination techniques, as performed by Duarte (2011), enhancing our capability to explore patterns and test hypotheses at the metacommunity level.

2.3 Phylogenetic signal

SYNCSA allows the detection of phylogenetic signal at two levels: at the species pool and at the metacommunity. A phylogenetic signal at the species pool level is expected when the variation in species traits is strongly correlated with their evolutionary history and it is measured by the matrix correlation $\rho(\mathbf{BF})$ between species resemblances based on their traits ($\mathbf{D_B}$) and on their phylogeny ($\mathbf{D_F}$). Phylogenetic signal at the metacommunity level is expected when communities are more similar to their phylogenetic structure, and also more similar to their average trait values which is measured by the matrix correlation $\rho(\mathbf{PT})$. Phylogenetic niche conservatism can be tested using path analysis as proposed by Pillar and Duarte (2010).

2.4 Searching for optimal traits

SYNCSA offers functions for searching for an optimal trait subset unraveling trait-convergence and trait-divergence assembly patterns related to one or more environmental variables. The function *optimal* examines all possible combinations of traits in subsets with one to a maximum specified number of traits. The output will indicate which trait subset maximizes the specified matrix correlation.

3 CONCLUSION

The methods implemented in the SYNCSA package for R enable the analysis of ecological data in a novel metacommunity perspective. In this way, ecologists may assess the linkages between functional and phylogenetic structure of communities, their responses to environmental drivers and their effects on ecosystem processes. Furthermore, functional and phylogenetic mechanisms explaining both β and α niche differentiation may be empirically examined. The metacommunity approach is complementary to the analysis of trait and phylogenetic patterns within communities, widely available in the R package. Moreover, by having in R methods that were originally available only in the SYNCSA application, users will have more flexibility for modifying the methods to suit their own purposes and for developing new ways to analyze ecological data.

ACKNOWLEDGEMENTS

We thank Alexandre A. Oliveira and Marcel C. Vaz for helpful suggestions on early versions of the SYNCSA package; Leandro S. Duarte and Vinicius A. G. Bastazini provided comments that improved the reference manual and the manuscript.

Funding: VJD received MSc scholarship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Conflict of Interest: none declared.

REFERENCES

- APG. (2009) An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG III. *Bot. J. Linn. Soc.*, **161**, 105–121.
- Bininda-Emonds, O.R.P. *et al.* (2007) The delayed rise of present-day mammals. *Nature*, **446**, 507–512.
- Cavender-bares, J. *et al.* (2009) The merging of community ecology and phylogenetic biology. *Ecol. Lett.*, **12**, 693–715.
- Duarte, L.S. (2011) Phylogenetic habitat filtering influences forest nucleation in grasslands. *Oikos*, **120**, 208–215.
- Felsenstein, J. (1985) Phylogenies and the comparative method. *Am. Nat.*, **125**, 1–15.
- Fine, P.V.A. and Kembel, S.W. (2011) Phylogenetic community structure and phylogenetic turnover across space and edaphic gradients in western Amazonian tree communities. *Ecography*, **34**, 552–565.
- Jones, K.E. *et al.* (2008) PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology*, **90**, 2648.
- Kattge, J. *et al.* (2011) TRY—a global database of plant traits. *Glob. Change Biol.*, **17**, 2905–2935.
- Keddy, P. and Weiher, E. (1999) Introduction: the scope and goals of research on assembly rules. In Weiher, E. and Keddy, P. (eds.) *Ecological Assembly Rules: Perspectives, Advances, Retreats*. Cambridge University Press, Cambridge, UK, pp. 1–20.
- Kembel, S.W. and Hubbell, S.P. (2006) The phylogenetic structure of a neotropical forest tree community. *Ecology*, **87**, 86–99.
- Paradis, E. (2004) APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*, **20**, 289–290.
- Pillar, V.D. and Duarte, L.S. (2010) A framework for metacommunity analysis of phylogenetic structure. *Ecol. Lett.*, **13**, 587–596.
- Pillar, V.D. and Orłóci, L. (1993) *Character-based Community Analysis: The Theory and an Application Program*. SPB Academic Publishing, The Hague.
- Pillar, V.D. *et al.* (2009) Discriminating trait-convergence and trait-divergence assembly patterns in ecological community gradients. *J. Veg. Sci.*, **20**, 334–348.
- Webb, C.O. *et al.* (2002) Phylogenies and community ecology. *Annu. Rev. Ecol. Syst.*, **33**, 475–505.