

The Forest, the Trees, and the Phylo-diversity Jungle

Florent Mazel^{1*}, Caroline M. Tucker², Marc W. Cadotte^{3,4}, Silvia B. Carvalho⁵, T. Jonathan Davies^{6,7}, Susanne A. Fritz^{8,9}, Rich Grenyer¹⁰, Matthew R. Helmus^{11,12}, Arne Ø. Mooers¹³, Sandrine Pavoine¹⁴, Oliver Purschke^{15,16,17}, Dan. F. Rosauer¹⁸, Marten Winter¹⁵

¹*Univ. Grenoble Alpes & CNRS, Laboratoire d'Écologie Alpine, F-38000 Grenoble.*

²*Centre d'Ecologie Fonctionnelle et Evolutive, CNRS, 1919 Route de Mende, 34293 Montpellier Cedex 5, France.*

³*Biological Sciences, University of Toronto-Scarborough, 1265 Military Trail, Scarborough, Ontario, Canada M1C 1A4.*

⁴*Stake Key Laboratory of Biocontrol, Key Laboratory of Biodiversity Dynamics and Conservation of Guangdong, Higher Education Institutes, College of Ecology and Evolution, Sun Yat-sen University, Xin Gang Xi Lu No. 135, Haizhu District, Guangzhou, PR China*

⁵*CIBIO/ InBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos da Universidade do Porto, R. Padre Armando Quintas, 4485-661 Vairão, Portugal.*

⁶*Department of Biology, McGill University, 1205 Avenue du Docteur-Penfield, Montréal, Quebec, Canada H3A 1B1.*

⁷*African Centre for DNA Barcoding, University of Johannesburg, APK Campus, PO Box 524, Auckland Park 2006, Johannesburg, South Africa*

⁸*Senckenberg Biodiversity & Climate Research Centre (BiK-F), Senckenberganlage 25, 60325 Frankfurt am Main, Germany.*

⁹*Institute of Ecology, Evolution and Diversity, Goethe University, Max-von-Laue-Straße 9, 60438 Frankfurt, Germany*

¹⁰*School of Geography and the Environment, University of Oxford, South Parks Road, Oxford, OX1 3QY United Kingdom.*

¹¹*Department of Ecological Sciences - Animal Ecology, Vrije Universiteit, 1081 HV Amsterdam, Netherlands.*

¹²*Center for Biodiversity, Department of Biology, Temple University, Philadelphia, PA 19122 USA*

¹³*Department of Biology, Simon Fraser University, 8888 University Drive, Burnaby, BC, Canada V5A 1S6.*

¹⁴*Centre of Ecology and Conservation Sciences (UMR 7204 CESCO), Museum National d'Histoire Naturelle, 61 rue Buffon, Paris, France.*

¹⁵*Biodiversity Synthesis Research Group, German Centre of Integrative Biodiversity Research, Halle-Jena-Leipzig, Deutscher Platz 5e, DE-04103 Leipzig, Germany*

¹⁶*Department of Computer Science, Martin-Luther-University, Halle-Wittenberg, Von-Seckendorff-Platz 1, DE-06120 Halle (Saale), Germany*

¹⁷*Geobotany and Botanical Garden, Institute of Biology, Martin Luther University, Halle-Wittenberg, Am Kirchtor 1, DE-06108 Halle (Saale), Germany*

¹⁸*Research School of Biology, Australian National University, Acton ACT 2601 Australia*

* *Author for correspondence (E-mail : flo.mazel@gmail.com).*

Abstract

The joint use of phylogenetic trees and ecological data has proven useful for many aspects of ecology. However, there are a multitude of phylo-diversity metrics with complex interdependencies and mathematical redundancies (the so-called ‘jungle’ of metrics). Several recent papers have been trying to ‘map’ this jungle but appear at a first glance to contradict each other. We suggest that these contradictory results are in fact complementary and reflect two approaches to understand diversity metrics: the first focuses on general mathematical properties, the second focuses on assessing metric performance in relation to particular questions. In this manuscript, we discuss the complementarity of the two approaches and in particular how recent papers fit into this categorisation.

Main text

The joint use of phylogenetic trees and ecological data has proven useful for understanding the assembly of local communities (Webb et al., 2002; Bryant et al., 2008), for exploring large scale diversity patterns (Graham & Fine, 2008), and even for developing target conservation priorities (Isaac et al., 2007). However, the last decades have seen a proliferation of phylo-diversity metrics; we counted more than 70 (Tucker et al. 2016). Winter et al. (2013) refer to the ever-increasing portfolio of phylo-diversity indices as a “jungle”, alluding to both the multitude of metrics and their complex interdependencies and mathematical redundancies. Several groups (e.g. Vellend et al. 2010, Pavoine and Bonsall 2011, Pearse et al. 2014, Tucker et al. 2016) have tried to navigate through this jungle by exploring metric relationships, complementarity and utility. In a recent paper, Miller et al. (2016) contribute an important discipline-specific perspective that further resolves the emerging map.

Metrics can be analysed in two ways: (1) by grouping them based on their underlying properties (e.g. by comparing mathematical formulations); and (2) by assessing context-dependent behaviour (e.g. by comparing metric performance in relation to particular questions).

The first approach requires theoretical and cross-disciplinary studies to summarize the main dimensions along which phylo-diversity metrics vary, while the second provides a field-specific perspective to quantify the ability of a particular metric to test a particular hypothesis. These two approaches have different aims, and their results are not *a priori* expected to be identical.

Miller et al. (2016) carry out this second approach within the discipline of community ecology by testing the ability of 32 phylo-diversity metrics and nine null models in discriminating between two ecological processes: habitat filtering and competition (see e.g. Hardy 2008).

The authors first simulated communities under three main assembly rules: competitive exclusion leading to species being less related than expected by chance, habitat filtering leading to species being more closely related than expected by chance, and neutral assembly. They then tested *a posteriori* which combination of metrics and null models yielded the best statistical performance. Surprisingly, only a fraction of phylo-diversity metrics and null models exhibited the ideal statistical properties of high statistical power coupled with low Type I error rate, leading Miller et al. to conclude that some metrics and null models proposed in the literature should be avoided when asking if filtering and competition play an important role in structuring communities. This is an important finding.

But before theoreticians run off to find new metrics, more multidisciplinary work is in order. One reason there are so many metrics is that they have been pooled across community ecology, macroecology and conservation biology. The questions typically asked by conservationists and macroecologists, for example, differ from those of community ecologists. Different metrics might perform better or worse for different types of problems. One solution would be to explicitly simulate the processes of interest for a given research question (e.g. vicariance or diversification processes in macroecological research), and select the most appropriate metric for the task. The R package presented by Miller *et al.*, as well as others (e.g. Pearse et al. 2015) help facilitate this approach. Of course, applying all possible metrics to the appropriate simulations and null models for a given hypothesis, is complex, time consuming and inefficient. To the extent this is true, this motivates the other approach to navigating the jungle, the unified framework.

We (Tucker et al., 2016) recently took the first approach to metric analysis and classified the 70 phylo-diversity metrics along three broad dimensions: *richness*, *divergence* and *regularity*-

-the sum, mean and variance of phylogenetic distances among species of assemblages, respectively-. Building upon a previous phylo-diversity classification system (Pavoine & Bonsall, 2011), which itself is based on a system for classifying taxonomic and functional diversity metrics (e.g. Ricotta 2007, Villéger et al. 2008), the Tucker et al. (2016) classification offers clear theoretical linkages between phylogenetic and functional approaches in ecology. We then used simulations to corroborate this classification system. Although they conclude differently, we feel that Miller et al. actually provide independent support of this tri-partite classification system. Indeed, the vast majority of metrics used by Miller et al. on their simulated communities group according to this richness-divergence-regularity classification system (see Figure 1 of Miller et al.). And metrics like H_{AED} and E_{ED} , which stem from a mathematical combination of richness and regularity dimensions, are expected to sometimes cluster with *richness* (as observed by Miller et al.), and sometimes with *regularity* (see the specific discussion on these hybrid metrics in Tucker et al., 2016).

So, while the synthesis presented in Tucker et al. (2016) takes a broad perspective across fields and across most metrics by providing an objective conceptual classification system, more focussed analyses, such as that by Miller et al., offer a detailed description of metric performance relevant to a given biological question. Both approaches have utility, and importantly, both approaches benefit each other. On one hand, detailed analyses of metric performance offer a valuable test of the broader classification system, using alternative simulations and codes. On the other hand, broad syntheses offer a conceptual framework within which results of more focussed analyses may be interpreted. For example, Miller et al. find that, in some situations, the metrics with the best statistical performances are Rao's quadratic entropy and IntraMPD. These are hybrid *richness* plus *divergence* metrics. And indeed, the finding that metrics closely aligned with only a single dimension (called 'anchor' metrics in Tucker et al. 2016) do not act as the best indicators of community assembly algorithms, provides strong support for the idea that the community assembly processes modelled by Miller et al. actually involve multiple phylo-diversity dimensions. This is a characteristic that has also been recognized in the functional trait literature (see, e.g. Botta-Dukát and Czúcz 2016).

In summary, Miller et al.'s community ecology study offers an excellent complement to broad-scale syntheses such as that provided by Tucker et al. (2016). We call on researchers to

continue to hack away at the jungle of phylo-diversity metrics in their own fields, in the hopes that the combination of in-depth understanding of what the metrics are (e.g., how they capture richness, divergence, regularity and combinations thereof), and how they perform under particular ecological and evolutionary processes will allow us a clearer view of our respective fields. Machetes up!

References

- Botta-Dukát Z. & Czúcz B. (2016) Testing the ability of functional diversity indices to detect trait convergence and divergence using individual-based simulation. *Methods in Ecology and Evolution*, **7**, 114–126.
- Bryant J.A., Lamanna C., Morlon H., Kerkhoff A.J., Enquist B.J., & Green J.L. (2008) Microbes on mountainsides: contrasting elevational patterns of bacterial and plant diversity. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 11505–11.
- Graham C.H. & Fine P.V.A. (2008) Phylogenetic beta diversity: linking ecological and evolutionary processes across space in time. *Ecology letters*, **11**, 1265–1277.
- Hardy O. (2008) Testing the spatial phylogenetic structure of local communities: statistical performances of different null models and test statistics on a locally neutral community. *Journal of ecology*, **96**, 914–926.
- Isaac N.J.B., Turvey S.T., Collen B., Waterman C., & Baillie J.E.M. (2007) Mammals on the EDGE: conservation priorities based on threat and phylogeny. *PLoS One*, **2**, e296.
- Miller E.T., Farine D.R., & Trisos C.H. (2016) Phylogenetic community structure metrics and null models: a review with new methods and software. *Ecography*
- Pavoine S. & Bonsall M.B. (2011) Measuring biodiversity to explain community assembly: a

- unified approach. *Biological Reviews*, **86**, 792–812.
- Pearse W.D., Cadotte M.W., Cavender-Bares J., Ives A.R., Tucker C.M., Walker S.C., & Helmus M.R. (2015) *pez*: phylogenetics for the environmental sciences. *Bioinformatics*, **31**, 2888–2890.
- Pearse W.D., Purvis A., Cavender-Bares J., & Helmus M.R. (2014) Metrics and Models of Community Phylogenetics. *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology* pp. 451–464. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Ricotta C. (2007) A semantic taxonomy for diversity measures. *Acta biotheoretica*, **55**, 23–33.
- Tucker C.M., Cadotte M.W., Carvalho S.B., Davies T.J., Ferrier S., Fritz S.A., Grenyer R., Helmus M.R., Jin L.S., Mooers A.Ø., Pavoine S., Purschke O., Redding D.W., Rosauer D.F., Winter M., & Mazel F. (2016) A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological reviews of the Cambridge Philosophical Society*.
- Vellend M., Cornwell W.K., Magnuson-Ford K., & Mooers A.Ø. (2010) Measuring phylogenetic biodiversity. *Biological diversity: frontiers in measurement and assessment*. (ed. by A.E.M.B.J. McGill), pp. 193–206. Oxford University Press,
- Villéger S., Mason N., & Mouillot D. (2008) New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology*, **89**, 2290–2301.
- Webb C.O., Ackerly D.D., McPeck M.A., & Donoghue M.J. (2002) Phylogenies and Community Ecology. *Annual Review of Ecology, Evolution, and Systematics*, **33**, 475–505.
- Winter M., Devictor V., & Schweiger O. (2013) Phylogenetic diversity and nature conservation:

where are we? *Trends in Ecology and Evolution*, **28**, 199–204.