Deconstructing the relationships between phylogenetic diversity and ecology: a case study on ecosystem functioning

T. Jonathan Davies,1,2,7 Mark C. Urban,3 Bronwyn Rayfield,1,4 Marc W. Cadotte,5 and Pedro R. Peres-Neto6

1Department of Biology, McGill University, Montreal, Quebec H3G 0B1 Canada
2African Centre for DNA Barcoding, University of Johannesburg, PO Box 524, Auckland Park, 2006, Johannesburg, South Africa
3Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, Connecticut 06269-3043 USA
4Département des sciences naturelles, Institut des sciences de la forêt tempérée, Université du Québec en Outaouais, 58 rue Principale, Ripon, Quebec J0V 1V0 Canada
5Biological Sciences, University of Toronto–Scarborough, Toronto, Ontario M1C 1A4 Canada
6Canada Research Chair in Spatial Modelling and Biodiversity, Département des sciences biologiques, Université du Québec à Montréal, Montréal, Quebec H2X 1Y4 Canada

Abstract. Recent studies have supported a link between phylogenetic diversity and various ecological properties including ecosystem function. However, such studies typically assume that phylogenetic branches of equivalent length are more or less interchangeable. Here we suggest that there is a need to consider not only branch lengths but also their placement on the phylogeny. We demonstrate how two common indices of network centrality can be used to describe the evolutionary distinctiveness of network elements (nodes and branches) on a phylogeny. If phylogenetic diversity enhances ecosystem function via complementarity and the representation of functional diversity, we would predict a correlation between evolutionary distinctiveness of network elements and their contribution to ecosystem process. In contrast, if one or a few evolutionary innovations play key roles in ecosystem function, the relationship between evolutionary distinctiveness and functional contribution may be weak or absent. We illustrate how network elements associated with high functional contribution can be identified from regressions between phylogenetic diversity and productivity using a well-known empirical data set on plant productivity from the Cedar Creek Long-Term Ecological Research. We find no association between evolutionary distinctiveness and ecosystem functioning, but we are able to identify phylogenetic elements associated with species of known high functional contribution within the Fabaceae. Our perspective provides a useful guide in the search for ecological traits linking diversity and ecosystem function, and suggests a more nuanced consideration of phylogenetic diversity is required in the conservation and biodiversity–ecosystem-function literature.

Key words: complementarity; ecosystem function; key innovations; phylogenetic diversity; productivity; sampling effect.

INTRODUCTION

Ecologists have long been interested in establishing a link between species diversity and various ecosystem properties including community productivity and stability (Loreau et al. 2001, Tilman et al. 2006, Naeem et al. 2012). In the shadow of the current extinction crisis, the drive to understand the relationship between biodiversity and ecosystem function has acquired new urgency (Hooper et al. 2012). Accumulating evidence suggests that ecosystem function and stability is generally enhanced within more biologically diverse communities (Cardinale et al. 2012). However, the mechanisms underlying the relationship between biodiversity and ecosystem function have been strongly debated (e.g., Aarsen 1997, Huston 1997). Carefully designed experiments along with new statistical approaches that were able to separate species-specific effects from biodiversity effects (Loreau and Hector 2001) helped move the field toward consensus (Loreau 2010). Two broad classes of mechanisms are
now recognized: the complementarity effect, in which ecosystem function is enhanced through functional complementarity among species and effect size is inversely proportional to niche overlap, and the selection effect, in which ecosystem function is enhanced in more diverse communities because of the increased likelihood of sampling species with large impact (Loreau 2000).

More recently, researchers have suggested that our understanding of diversity effects on ecosystem function might be furthered by incorporating information on the phylogenetic distances separating species (Maherali and Klironomos 2007, Cadotte et al. 2008, 2009) as proxies for unmeasured species characteristics that may be shared through common ancestry. One frequent practice has been to use phylogenetic diversity, the sum of the phylogenetic branch lengths connecting a species set (Faith 1992), as a surrogate for functional diversity (Vane-Wright et al. 1991, Crozier 1997). Using this framework, communities with greater phylogenetic diversity are assumed to capture greater trait diversity, and hence represent more potential niche complementarity.

In an analysis of biodiversity experiments, Cadotte and coworkers indicated that phylogenetic diversity provided a better predictor of ecosystem productivity than either species richness or functional diversity alone (Cadotte et al. 2008, see also Flynn et al. 2011). The linkage between productivity and phylogenetic diversity has subsequently been questioned by Venail et al. (2015), among others, and is further complicated by the strong collinearity between species richness and phylogenetic diversity. A reanalysis of the data presented in Venail et al. (2015) by Cadotte (2015) again suggested that phylogenetic diversity was a better predictor of productivity than either species richness, although the debate is ongoing (Cardinale et al. 2015). Why might phylogenetic diversity correlate more strongly with productivity than species richness or directly measured functional traits? The complementarity effect is predicated upon niche differentiation among coexisting species, allowing for complementary resource use. However, not all species are equal (or equally distinct), as some may be more similar than others as a product of their shared evolutionary history. Under most models of trait evolution, more closely related species are predicted to occupy similar niches (phylogenetic niche conservatism sensu Harvey and Pagel 1991), and thus show greater functional redundancy (Fig. 1).

The predictive power of phylogenetic diversity might nonetheless appear surprising. First, if we could measure the trait(s) relevant to ecosystem function, this would obviously provide a more direct index of niche complementarity, yet we find that phylogenetic diversity can outperform indices of functional diversity (Cadotte et al. 2008), although not always (e.g., Flynn et al. 2011, Paquette and Messier 2011). Second, most traits are likely to depart to some extent from a simple model of neutral (i.e., Brownian motion) evolution (Hansen 1997), and traits relevant to ecosystem function are likely to be under selection, such that evolutionary distances might not capture accurately species functional differences. Third, even for traits demonstrating tight phylogenetic conservatism, phylogenetic distances can still be a poor proxy for functional differences (Letten and Cornwell 2014). In part, the poor performance of functional trait indices likely reflects the difficulties associated with identifying functionally relevant traits and obtaining useful data (see Petchey and Gaston [2006] for further discussion). Frequently, trait indices are determined as much by data availability as by knowledge of organismal biology and ecology, although increased access to large trait databases, such as the TRY global database of plant traits (Kattge et al. 2011), might help bridge this gap. Phylogeny, however, might integrate across unmeasured traits that contribute to productivity differences (Flynn et al. 2011). Critically, even if individual traits evolve somewhat idiosyncratically, phylogenetic distance may still provide a reasonable approximation of species functional differences, assuming function is an emergent property reflecting the synergistic effects of multiple independent traits.

We suggest that if ecosystem function is enhanced through niche complementarity and location in niche space is determined by multiple traits, then species or clades that are the most evolutionarily distinct in a community should have the largest effect on community function (Fig. 1), providing a short cut to identifying functionally important species or lineages. Recent discussions have debated whether evolutionarily distinct lineages are also functionally distinct (Rosauer and Mooers 2013, Winter et al. 2013). In one cited example, *Welwitschia mirabilis*, the sole surviving member of the family Welwitschiaceae and sometimes referred to as a living fossil, is both phylogenetically and functionally distinct. In addition, experimental manipulations of plant communities have shown that species tend to perform better when

![Fig. 1](image-url)
grown with more distant relatives (e.g., Burns and Strauss 2011, Cadotte 2013), suggesting increased niche partitioning and functional complementarity with phylogenetic distance (see Parker et al. 2015 for an alternative explanation for this relationship based upon reduced pathogen sharing among more distant relatives). However, the evidence remains mixed (e.g., Cahill et al. 2008, Connolly et al. 2011, Narwani et al. 2013, Venail et al. 2015), and it is as yet unclear whether such relationships can be more widely generalized (Winter et al. 2013). For example, if community function is determined by particular key evolutionary innovations, then evolutionarily distinct clades, i.e., those with a large contribution to phylogenetic diversity, might not necessarily match to those with a large contribution to community function. In the latter case, it is not the summed length of the phylogenetic branches (phylogenetic diversity) that is important for ecosystem function, but rather their identity (placement on the phylogenetic tree), which is linked to the inheritance of key evolutionary innovations from a common ancestor. If the possession of a particular key innovation is more important than the integration of multiple phenotypic differences we might predict a disconnection between phylogenetic diversity and ecosystem function.

Here we demonstrate how two common indices of network centrality can be used to describe the evolutionary distinctiveness of network elements (nodes and branches) on a phylogeny, and present a simple method to identify phylogenetic elements associated with a large functional contribution. Our framework is general enough that it can be used to identify phylogenetic elements important in any analysis involving phylogenetic biodiversity (e.g., phylogenetic diversity across environmental gradients; Graham et al. 2009). We focus on ecosystem function as there has been much recent interest in the biodiversity–ecosystem-function relationship, data is easily available, and it provides a simple illustration of how our framework is useful in exploring the links between phylogenetic diversity and a key ecological process. A strong correlation between evolutionary distinctiveness and functional contribution could suggest that multiple (possibly interacting) traits are important for determining function, and that the loss of evolutionarily distinct species or clades would have a large impact on ecosystem function, providing further justification for conservation initiatives, such as the Evolutionarily Distinct and Globally Endangered (EDGE) of existence program (http://www.edgeofexistence.org/conservation/) that focuses upon EDGE species (Isaac et al. 2007). Alternatively, if the relationship between evolutionary distinctiveness and functional contribution is weak or absent, it may indicate that ecosystem function is instead regulated by only one or a few key evolutionary innovations because these would tend to be only poorly represented by phylogenetic relationships (even if they demonstrate strong phylogenetic conservatism, see above). However, identifying the phylogenetic elements with large functional contribution would then help guide the search for those ecological traits key to maintaining ecosystem function. We demonstrate our approach using an empirical data set on plant productivity from the Long-Term Ecological Research (LTER) experiments at Cedar Creek, Minnesota, for which a strong species richness-productivity relationship has been well documented and for which we know nitrogen fixers have large effect on community function (Tilman et al. 1996).

**Measuring the Evolutionary Distinctiveness of Phylogenetic Elements**

The field of ecological phylogenetics has expanded rapidly over the past decade (Webb 2000, Webb et al. 2002, Cavender-Bares et al. 2009, Vamosi et al. 2009), and various metrics have been described that capture the phylogenetic structure of ecological assemblages (e.g., Kemel 2009, Cadotte et al. 2010, Vellend et al. 2011, Peres-Neto et al. 2012). Choosing among them depends upon the question of interest. For example, tree topology can capture the process of speciation, extinction (Mooers and Heard 1997), and immigration (Davies and Buckley 2012), whereas the distances connecting taxa may provide information on the process of community assembly (Webb et al. 2002). Methods to quantify the evolutionary distinctiveness of species are already well developed, typically derived from some function of phylogenetic distance between neighboring nodes (e.g., Vane-Wright et al. 1991, Pavoine et al. 2005, Vellend et al. 2011). Some recent metrics now also allow for the identification of functionally distinct clades (e.g., the KSI index of Cornwell et al. [2014]) and nodes subtending clades with unusual biogeographical distributions (e.g., Borregaard et al. 2014). Here, we extend this thinking and use metrics more commonly applied to describe network connectivity (Newman 2010) to derive indices of evolutionary distinctiveness that can be applied to both tip species and internal nodes in the phylogenetic tree.

We consider two well-known metrics that describe the connectedness or centrality of nodes and edges within a network (Freeman 1979, Newman 2010): closeness centrality and betweenness centrality (Fig. 2). These network metrics capture connectivity patterns among non-neighboring nodes thereby identifying internal nodes that enhance or diminish the connectivity between network elements. We assume phylogeny can be represented as a node-based tree (Wiley and Lieberman 2011) where internal nodes capture evolutionary splits and external nodes represent species. Nodes are connected by weighted edges (phylogenetic branch lengths) representing evolutionary distances associated with parent–child relationships, for example, measured in number of substitutions. To be consistent with conventional phylogenetic terminology, we hereafter refer to edges as branches. Phylogenetic networks belong to a class of networks that exhibit dendritic topologies, characterized by hierarchical arrangements of nodes and branches (Peterson et al. 2013). Network centrality metrics, including betweenness
Measuring the Functional Contribution of Phylogenetic Elements

Phylogenetic diversity can be quantified as the sum of the branch lengths connecting species within a community or geographic location (Faith 1992). Given multiple communities or locations, phylogenetic diversity can then be correlated with some index of ecosystem function, such as productivity, using simple linear regression (e.g., see...
Cadotte et al. [2008]). Here we illustrate a novel framework that allows us to quantify the relative influence of different parts of the phylogenetic tree in determining the strength of the correlation between phylogenetic diversity and ecosystem function. First, we fit the regression of phylogenetic diversity against the ecological variable of interest (productivity in our example) and extract a metric of statistical leverage (DFBETA) for each community from the regression. The DFBETA measures the influence of each data point on the estimated regression coefficient (Neter et al. 1996). A data point with a high DFBETA signifies a point that alters the regression coefficient appreciably when removed, and is calculated as the difference between the regression coefficient calculated on all of the data and the regression coefficient calculated with the data point deleted, scaled by the standard error calculated with the data point deleted (Chen et al. 2003). Second, for each community in turn, we identify the branches in the phylogeny of the complete species pool that are represented in the community sub-tree, and assign each of the shared branches a weight equal to the respective DFBETAs from the regression for that community. Last, we calculate the mean weight across all community sub-trees for each branch on the original phylogeny of the species pool (see Fig. 3). Because of the nested architecture of phylogenetic trees, branches deeper in the tree will be included more times in our regression plots, and will thus tend toward the mean.

**A Case Study: The Cedar Creek LTER**

We illustrate our approach using a data set on productivity (annual aboveground biomass production) across 160 experimental plots from the Cedar Creek LTER diversity experiment. We use the molecular phylogenetic tree constructed by Cadotte et al. (2009) with branch lengths/weights proportional to number of substitutions using maximum likelihood (Newick tree file provided as supplemental Data S1). The complete phylogeny includes 18 species (excluding trees and species with insufficient plot-level data to evaluate correlations with biomass), and spans the monocot–eudicot split. Our purpose here was to evaluate our approach using a well-studied data set on ecosystem function (quantified here using primary productivity) showing both strong links with plant species richness and functional traits (nitrogen fixation) (Tilman et al. 1996, 1997). First, we quantify functional contribution by regressing phylogenetic diversity against biomass and calculate DFBETAs for each of the 160 Cedar Creek LTER plots. Our regression differs from that by Cadotte et al. (2008) in that we included root distance in our estimates of phylogenetic diversity so that single species plots had non-zero phylogenetic diversity. Second, we quantify evolutionary distinctiveness of nodes and branches, using the inverse of the two metrics of centrality described in Measuring the Evolutionary Distinctiveness of Phylogenetic Elements in the R library igraph (Csárdi and Nepusz 2006).

Our regression model explains 46% of the variation in biomass (Fig. 4, $P < 0.001$, $t = 11.54$), and is marginally preferred to a model with species richness (the change in the Akaike information criterion, $\Delta AIC = 3.56$), although...
the increase in explanatory power was small (~2%). By mapping the functional contribution (DFBETAs) of each branch on the phylogeny, we can reveal the relative contribution of each phylogenetic element to community-wide productivity (Fig. 5). To contrast with functional contribution, we present the equivalent phylogenetic tree with branches shaded by the two centrality indices (Fig. 6). For ease of comparison, we assign node closeness centralities to the parent branches.

In the Cedar Creek phylogeny, the branch subtending the clade containing Dalea candida, Dalea purpurea, Amorpha canadensis, Lespedeza capitata, and Lupinus perennis (i.e., the legume family, Fabaceae) is easily identified as an obvious functional outlier (Fig. 5). By contrast, Asclepias tuberosa, Schizachyrium scoparium, Pascopyrum smithii, and Elymus canadensis, all returned little additional biomass relative to their contribution to total phylogenetic diversity. Some branches receive similar ranks for both evolutionary distinctiveness and functional contribution (cf. Figs. 5 and 6). For example, the branches subtending from A. tuberosa (milkweed) and S. scoparium (little bluestem), receive low values for both functional (low DFBETAs) and evolutionary indices (high betweenness and closeness). However, differences in weightings are common. Notably, the clade including L. perennis (sundial lupine) and L. capitata (bush clover), has high positive functional influence (high DFBETAs), but contributes little to phylogenetic structure. To evaluate more formally whether clades with large influence on phylogenetic structure also had large influence on ecosystem function, we regressed functional weights against the two evolutionary indices and found no significant correlation ($P > 0.1$ for both regressions).

**Discussion**

There is growing appreciation that evolutionary relationships might be important in shaping the functional ecology of communities (Maherali and Klironomos 2007, Cadotte et al. 2008, 2009, Cavender-Bares et al. 2009, Wiens et al. 2010, Flynn et al. 2011, Mouquet et al. 2012, Srivastava et al. 2012) because phylogeny may capture the integrated phenotypic differences among taxa (Harvey and Pagel 1991, Vane-Wright et al. 1991, Faith 1992, Crozier 1997) and thereby niche differences and similarities (but see Silvertown et al. 2006, Glücksman et al. 2010). While the precise relationship between ecosystem function and phylogenetic diversity is not without controversy (e.g., see Cadotte 2015 and Venail et al. 2015), there is accumulating evidence to suggest that in
Fig. 6. Regional phylogeny for the Cedar Creek biodiversity experimental plots, with branch lengths shaded by their contributions to community phylogenetic structure: (a) undirected, weighted closeness and (b) directed, unweighted betweenness. Color scheme from red (low) to white (high). See online version for color image.
some cases phylogenetic diversity might be a better predictor of community productivity than simple measures of species richness or functional diversity estimated from a few select traits (Cadotte et al. 2008, Flynn et al. 2011, Cadotte 2013).

Here, we have presented a simple heuristic for quantifying the contribution of lineages to ecosystem function and phylogenetic community structure. We illustrated our approach using data from the Cedar Creek LTER biodiversity experiments. Despite demonstrating a highly significant relationship between phylogenetic diversity and productivity (see also Cadotte et al. 2008), we find no evidence for a correlation between the evolutionary distinctiveness of lineages and their contribution to community productivity. Clades important in explaining variation in biomass are different from those that structure the phylogenetic architecture of the community, at least for the indices examined here. Nonetheless, our method is able to capture the large functional contribution of species within the nitrogen-fixing clade.

Why did we fail to detect any relationship between functional contribution and evolutionary distinctiveness, yet still find a strong correlation between phylogenetic diversity and productivity? Phylogenetic diversity might be a poor surrogate for functional diversity when only one or a few traits determine functional responses and/or they evolve rapidly along the branches of the phylogenetic tree (Flynn et al. 2011, Peres-Neto et al. 2012). Nonetheless, we might still expect to find an overall positive relationship between phylogenetic diversity and biomass via the selection effect: more phylogenetically diverse plots are more likely to sample lineages incorporating key evolutionary innovations (Gravel et al. 2012).

For the Cedar Creek experimental plots, we expected a priori that biomass production would be greater than predicted from phylogenetic diversity for plots including a nitrogen fixer (i.e., species within the Fabaceae family; see Fig. 4). The ability to fix nitrogen is a well-recognized key innovation (Vitousek and Hooper 1993) and perhaps the single most important functional trait explaining variation in biomass productivity within these plant communities (Tilman et al. 2001). Importantly, it is not that nitrogen fixers themselves contribute more biomass, but rather they facilitate the growth of co-occurring species that would otherwise be nitrogen limited. Hence, it is the interaction between species functional groups that maintains ecosystem functioning. This single key trait, the ability to fix nitrogen, therefore explains much of the variation in the phylogenetic-diversity–productivity relationship (Fig. 4). Cadotte et al. (2009) were able to show that a significant, albeit weaker, correlation between productivity and phylogenetic diversity remained even when comparing only plots including legumes (see also Fig. 4). Although our approach does not allow us to generate the equivalent model (legumes would always be identified as the lineage with large effect if included in the community sample), we reevaluated the relationship between productivity and evolutionary distinctiveness on plots excluding legumes and we still failed to detect any relationship ($P > 0.05$ for correlations between functional weights and both our centrality indices after excluding legumes), although it should be noted that the sample size of plots excluding legumes is much reduced.

Implicit within most phylogenetic diversity metrics (e.g., Faith’s [1992] phylogenetic diversity, which equals the sum of the branch lengths connecting taxa) is the assumption that phylogenetic branches of equivalent length are more or less interchangeable, such that the phylogenetic diversity represented by a terminal branch subtending from a tip species has an equal contribution as an internal branch of the same length. Here we have questioned this assumption. We showed that some phylogenetic branches may have disproportionate or keystone effects on community function. By locating the position of these branches, we can help isolate the evolutionary transitions that might be most important in determining community function and stability. Such efforts might be particularly worthwhile where the key evolutionary innovation is not as obvious as for the Cedar Creek example.

Our example emphasizes the importance of functional traits in moderating ecosystem processes, and indicates that in some cases simple metrics of phylogenetic diversity, such as Faith’s phylogenetic diversity, might fail to capture key species traits relevant to ecosystem properties. Nonetheless, we show that phylogeny is still informative even in the absence of detailed trait data. First, when the biodiversity–ecosystem-function relationship is driven by species complementarity, phylogenetic diversity might better capture species’ niche differences (see Cadotte 2013). Second, by identifying branches on the phylogeny as likely sites of key evolutionary transitions, phylogenies help direct the search for functional traits when prior knowledge of function is lacking and selection effects are thought to be important. Although we did not explore them here, other facets of diversity likely have important additional contributions to ecosystem function, such as the relative density and evenness in abundances of constituent species. For example, it is possible that shifts in density could drive differences in productivity (Marquard et al. 2009), and help explain the observation that more diverse ecosystems sometimes are also more stable over time (Tilman et al. 2006). While density might also demonstrate phylogenetic patterning, for example, mediated by shared pathogens (Parker et al. 2015), our method may be less sensitive to such processes.

We have illustrated our approach using two network metrics of centrality, betweenness and closeness, that allowed us to quantify the evolutionary distinctiveness of internal elements on a phylogenetic tree. For tip branches, we have shown that closeness centrality is tightly correlated to more commonly used metrics of phylogenetic distinctiveness. However, we believe that our approach may be easily extended to consider other aspects of community structure, and we provide one example of how lineages or clades with large influence on phylogenetic
tree structure can be identified using metrics from the community phylogenetics literature (Appendix S2).

We believe there is much scope for the application of graph theory and network analysis (Freeman 1979, Newman 2010) within phylogenetic studies beyond the reconstruction of evolutionary relationships (Foulds et al. 1979, Huson and Bryant 2006). We highlighted a single example using data on plant productivity and attempted to identify individual network elements with high contribution. A next step could be to explore combinations of elements that maximally complement each other, emphasizing the importance of key interactions in shaping ecosystem properties (Lambers et al. 2004). Paralleling advances in the field of evolutionary biology that have allowed us to identify clades and nodes associated with shifts in evolutionary rates, several recent efforts have developed indices that identify clades with unusual functional or biogeographical distributions (e.g., Leibold et al. 2010, Borregaard et al. 2014, Cornwell et al. 2014). We suggest also that such approaches might naturally be combined with network theory to characterize features of tree topology associated with unusual ecological features. The application of network theory in phylogenetic ecology has been little explored, and more work is needed to evaluate the behavior of network metrics applied to phylogenetic trees (see Appendix S4). Nonetheless, our analyses indicate that the merging of these fields holds much promise.

Acksnowledgments

This paper was initiated in a workshop co-funded by the Canada Research Chair in Spatial Modelling and Biodiversity and by the Quebec Centre for Biodiversity Science. We would like to thank David Tilman for access to recent Cedar Creek data and Mark Vellend and two anonymous reviewers for valuable comments on an earlier version of this manuscript. We also acknowledge the U.S. National Science Foundation Long-Term Ecological Research (NSF-LTER) funding that supported the data used.

Literature Cited


**Supporting Information**

Additional supporting information may be found in the online version of this article at http://onlinelibrary.wiley.com/doi/10.1002/ecty.1507/suppinfo