CrossMark

# Untangling the Tangled Bank: A Novel Method for Partitioning the Effects of Phylogenies and Traits on Ecological Networks

**Vinicius A. G. Bastazini**[1] · **Pedro M. A. Ferreira**[2] · **Bethânia O. Azambuja**[1] ·
**Grasiela Casas**[1] · **Vanderlei J. Debastiani**[1] · **Paulo R. Guimarães Jr.**[3] ·
**Valério D. Pillar**[4]

**Abstract**  Understanding how evolutionary and ecological processes shape species interaction networks remains as one of the main challenges in eco-evolutionary studies. Here, we present an integrative analytical framework to partition the effects of phylogenies and functional traits on the structure of ecological networks. The method combines fuzzy set theory and matrix correlation, implemented under a Monte Carlo framework. We designed a simulation study in order to estimate the accuracy of the methods proposed here, measuring Type I Error rates. The simulation study shows that the method is accurate, i.e., incorrectly rejecting a true null hypothesis in ~5% of the cases and falling within the confidence interval. We illustrate our framework using data from a seed dispersal network from southern Brazil. Our analyses suggest that birds must have specific traits in order to consume their plant resources, and that phylogenetic resemblance has no explanatory power for species traits and species interactions in this seed-dispersal network.

**Keywords**  Eco-evolutionary dynamics ·
Ecophylogenetics · Fuzzy sets · Phylogenetic comparative method · Species interaction · Trophic association

✉ Vinicius A. G. Bastazini
bastazini.vinicius@gmail.com

1   Graduate Program in Ecology, Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves 9500, 91501-970 Porto Alegre, RS, Brazil

2   Departamento de Biodiversidade e Ecologia, Pontifícia Universidade Católica do Rio Grande do Sul, Av. Ipiranga, 6681, Porto Alegre, RS, Brazil

3   Department of Ecology, Instituto de Biociências, Universidade de São Paulo, Rua do Matão, Travessa 14, 05508-900 São Paulo, SP, Brazil

4   Department of Ecology, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

## Introduction

Species interactions are central to understanding ecological and evolutionary dynamics, yet are among the most challenging attributes to quantify and relate to other aspects of biodiversity (Strauss and Irwin 2004; Pascual and Dunne 2006; Bersier and Kehrli 2008; Gómez et al. 2010). Darwin (1859) envisaged the complex network of interacting organisms as an "entangled bank" shaped by deterministic "laws" that drive the associations among species. We have now come to recognize that these "laws" likely reflect different ecological and evolutionary processes (Strauss and Irwin 2004). Some of the well-recognized factors shaping the architecture of ecological networks are environmental gradients (Danielson 1991; Tylianakis et al. 2007), species abundance (Krishna et al. 2008; Vázquez et al. 2009; Verdú and Valiente-Banuet 2011; Ibanez 2012), sequential evolution (Jermy 1976), cocladogenesis or cospeciation (Page and Charleston 1998; Page 2003; Thompson 2014), coevolution (Thompson 1994; Gross and Blasius 2008; Guimarães et al. 2011) and species functional traits (Cohen 1977; Rezende et al. 2007; Rossberg et al. 2010; Ibanez 2012; Vizentin-Bugoni et al. 2014), defined here as any measurable organismal eto-ecological characteristic or phenotype associated with a biotic interaction or an ecosystem function of interest (Schmitz et al. 2015). Nonetheless, knowledge of the independent role of these factors on species

interactions at the community level is still poorly appreciated (Mitter et al. 1991; Lewinsohn et al. 2005; Novotny and Basset 2005), especially when it comes to measuring the importance of functional traits. As closely related species are likely to retain their ancestral traits, discriminating the effects of species traits and long-term evolutionary history is troublesome (Felsenstein 1985; Losos 2008).

Over the past recent decades, ecologists have started to incorporate phylogenetic-based methods into ecological research, which gave rise to a novel and prolific field of inquiry, the so-called Ecophylogenetics (Webb 2000; Webb et al. 2002; Mouquet et al. 2012). These methods have allowed ecologists to overcome some of the difficulties inherent to the study of evolutionary phenomena, and have been very useful in studies trying to understand how historical and evolutionary contingencies affect patterns at different ecological scales (Mouquet et al. 2012). With the ready availability of extensive phylogenies, community ecologists have started to integrate ecological and evolutionary processes in order to describe patterns of ecological community assembly and species diversity (Webb 2000; Webb et al. 2002; Johnson and Stinchcombe 2007; Cavender-Bares et al. 2009; Pillar and Duarte 2010).

Despite the advances that phylogenetic data have brought to our understanding of patterns of community structure, the same has not been achieved for the ecological knowledge on interacting species yet. In this respect, approaches that relate phylogenetic data to species interactions have been used in order to verify the degree of phylogenetic clustering or over dispersion in the specificity of plant-herbivore interactions, especially in tropical forests (e.g., Novotny and Basset 2005; Weiblen et al. 2006). Such an approach has also been used to study coevolution within small clades (e.g., Pedron et al. 2012), to investigate the importance of traits and phylogenies within a single trophic level (e.g., Novotny and Basset 2005; Weiblen et al. 2006) and to evaluate the effect of phylogenetic structure on coextinction cascades within networks (Rezende et al. 2007). However, the only ecological information incorporated in these models is the relationship between abundance and each taxa, and important ecological factors such as habitat structure and species traits are usually omitted (but see, for instance, Pearse and Hipp 2009; Vázquez et al. 2009; Rafferty and Ives 2013). The interplay between phylogenetic history and the degree of phylogenetic conservatism in species traits are likely to affect some of the patterns found in empirical ecological networks (Fig. 1).

Insofar, few integrative methods have been proposed and evaluated in detail. For instance, Legendre et al. (2002) and Hommola et al. (2009) developed permutation methods to measure coevolution and cospeciation between hosts and their parasites. Cruz et al. (2012) developed a method to test the influence of phylogenies in the arrangement of bipartite networks (i.e., a network which has two sets of species, such as plants and pollinators, with edges between and not within sets). These methods use only phylogenetic data, and recent analyses have demonstrated that some of them are likely to result in high rates of Type I Error, i.e., incorrectly rejecting a true null hypothesis (Hadfield et al. 2014). Model based approaches, such as generalized linear mixed models or phylogenetic regressions (Ives and Godfray 2006; Hadfield et al. 2014) have also been developed in order to describe phylogenetic patterns in bipartite networks, and have recently been extended to include trait data (Rafferty and Ives 2013). However, these model-based approaches rely on distribution assumptions, which may be unrealistic and complicated to parameterize. Randomization tests offer a simpler solution and are much less restricted by assumptions, as they rely on distribution free statistics and on algorithms that use systematic or random data permutations to generate alternative outcomes for the chosen test statistic under a true null hypothesis (Pillar 2013).

Here, we develop a statistical framework that integrates phylogenetic and trait data from bipartite networks, although the method can be extended to unipartite networks (networks that have only one set of species, such as a co-occurrence network), based on fuzzy set theory (Zadeh 1965) and randomization tests. Our approach enables the detection of: (1) phylogenetic signal in traits accounting for species interactions; (2) correlation among traits accounting for species interactions, removing the effect of phylogenies; and (3) correlation between phylogenies, which is a necessary but not sufficient condition to be considered a sign of cospeciation or cocladogenesis. We provide a simulation study that shows that our methodology provides adequate rates of Type I Error. We also showcase our framework with an analysis of a mutualistic network from southern Brazil.

## Input Data

Our framework requires data organized into, at least, five matrices (Fig. 2a): (i) A matrix containing the observed pairwise species interactions (matrix $\mathbf{N_{AB}}$; containing either binary or frequency data). For simplicity we will assume that this matrix contains two different trophic levels, with prey (A) and predators (B), but it can accommodate any kind of antagonist or mutualistic interaction, and it can also contain data from unipartite networks. The phylogenetic pairwise resemblances between species: matrix $\mathbf{A_A}$ for the lower trophic level and $\mathbf{A_B}$ for the higher trophic level (ii–iii). Species trait data: matrices $\mathbf{B_A}$ (*traits × species*) for the lower trophic level and $\mathbf{B_B}$ for the higher trophic level (iv–v), with traits of any type. An additional binary matrix

**Fig. 1** Patterns of ecological interactions arising from the interplay between species functional traits and their evolutionary history. *Black squares* denote realized species interactions. **a** Species interactions are a result of labile species traits that are not phylogenetic conserved; **b** traits accounting for species interaction are phylogenetic conserved; **c** interactions are mediated by shared phylogenetic history, independent of species traits; **d** functional traits are partially phylogenetic conserved
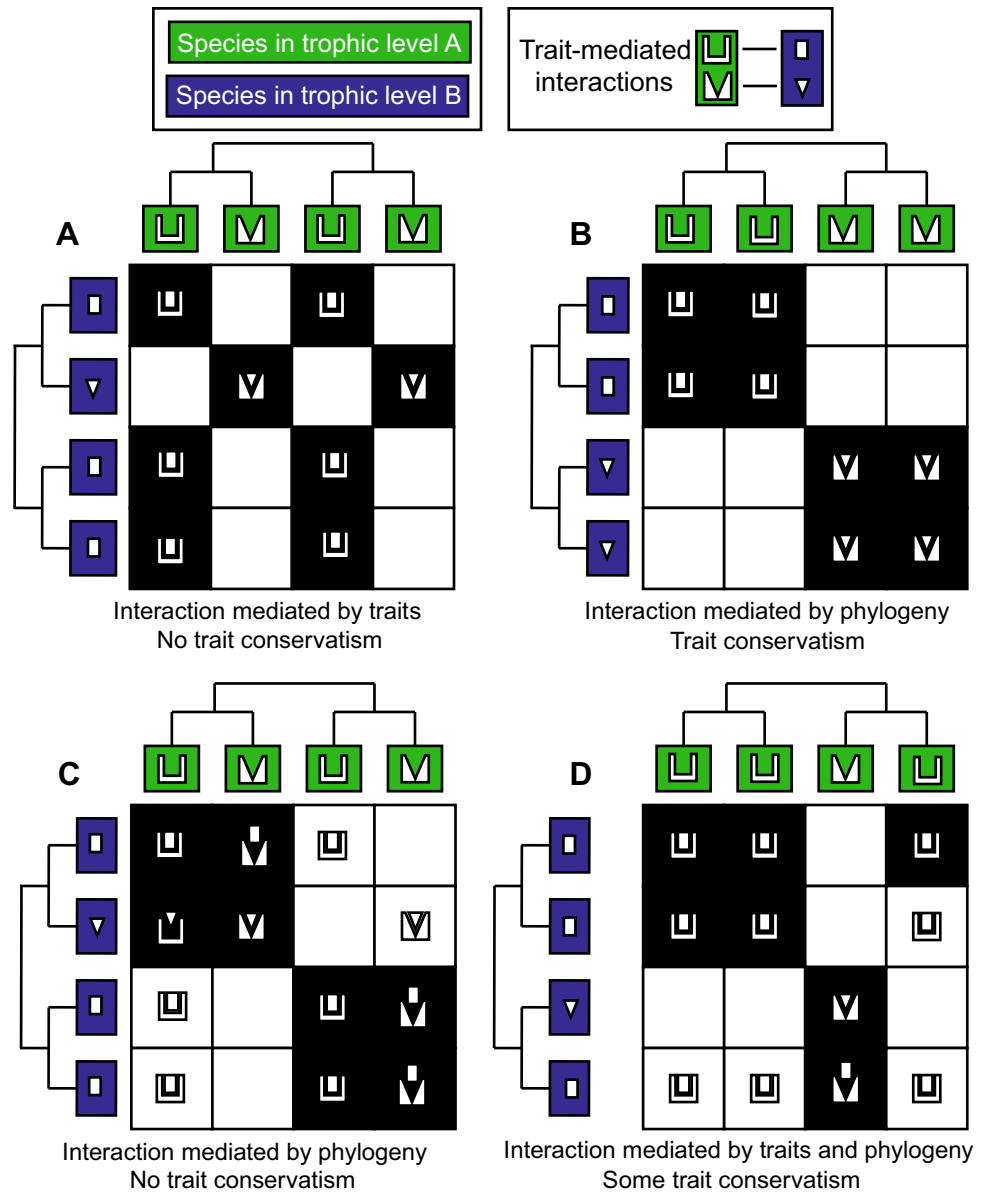


(**F**), containing species pairwise co-occurrence (either temporal or spatial), coded as 0 (the pair of species do not co-occur) or 1 (the pair of species co-occur), can also be used (Fig. 2c). Matrix **F** is not required for the analysis, but may provide additional information and control for situations where phenological or spatial decoupling, for instance, is common.

## Scaling Up of Traits and Phylogenies to the Network Level

Trait data in matrices $\mathbf{B_A}$ and $\mathbf{B_B}$, and phylogenetic data in matrices $\mathbf{A_A}$ and $\mathbf{A_B}$ (Fig. 2a) are scaled up to the network level using similar methods previously described in Pillar

and Orlóci (1991), Pillar et al. (2009) and Pillar and Duarte (2010), which are extended and explained here in detail. The first step is to transform trait matrices ($\mathbf{B_A}$ and $\mathbf{B_B}$), into a symmetric similarity matrix (Matrices $\mathbf{S_A}$ and $\mathbf{S_B}$; Fig. 2b), using an appropriate measure. For this, we have used Gower's general similarity coefficient in the range 0–1, which is defined as:

$$G_{ij} = \frac{\sum_{j=1}^{p} w_{ijh} s_{ijh}}{\sum_{j=1}^{p} w_{ijh}} \tag{1}$$

where $s_{ijh}$ is the partial similarity for trait $h$ in the range 0–1 between any species $i$ and $j$ of the same trophic level, and $p$ is the number of traits. This partial similarity can be applied to different types of traits (binary, qualitative,
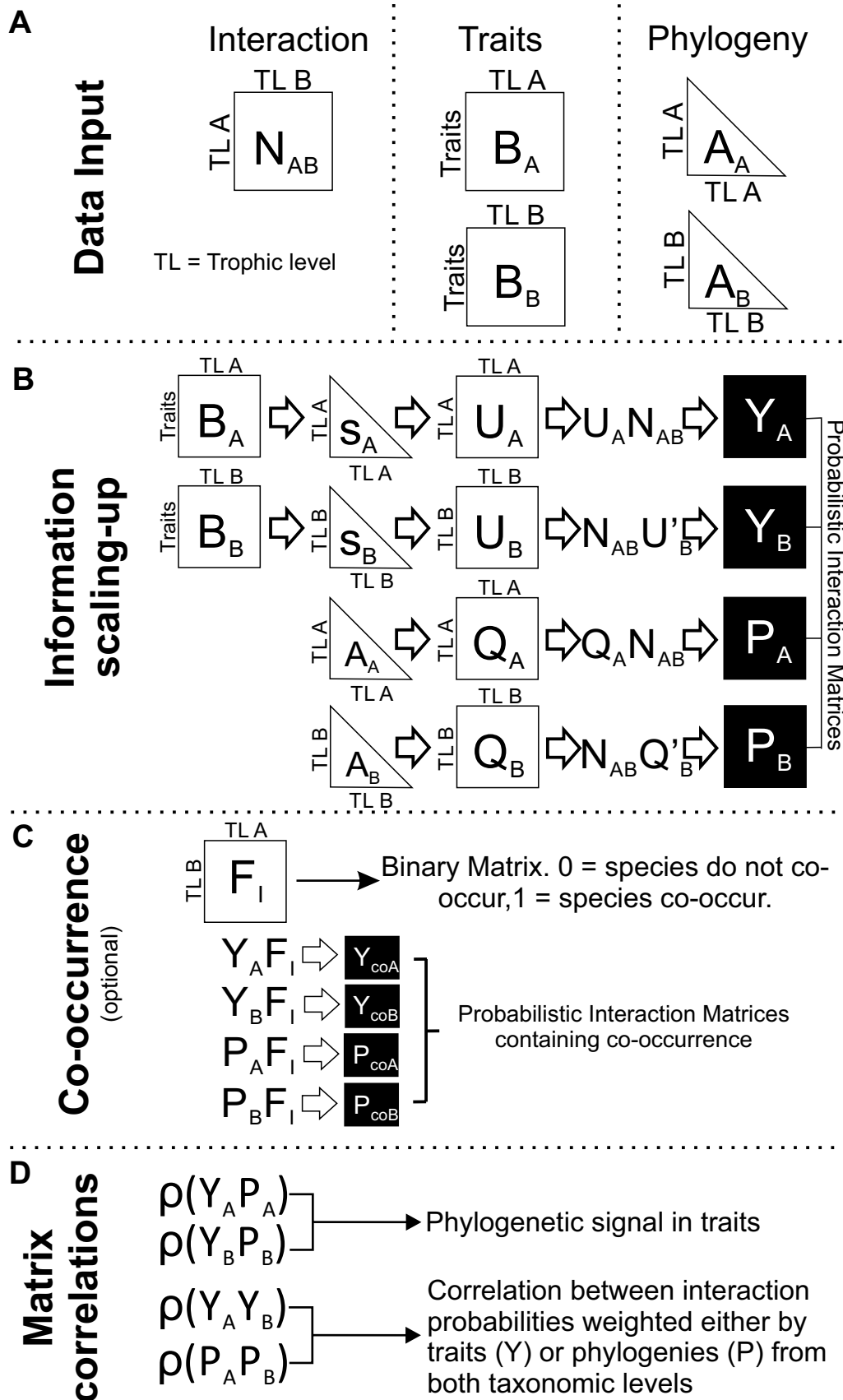
◄**Fig. 2** An overall scheme of our analytical framework. **a** Input data; **b** scaling-up of phylogeny and trait-based data to the network level; **c** fuzzy-weighted interaction matrices are multiplied, using Hadamard element-wise multiplier, by a (temporal or spatial) co-ocurrence matrix; **d** eco-evolutionary patterns revealed by matrix correlations

semiquantitative, quantitative and mixed variables) each one according to its own mathematical type (Podani 2000; Legendre and Legendre 2012). $w_{ijh}$ is the Kronecker's delta, describing the presence of information: $w_{ijh} = 0$ when the information about trait $h$ is missing for species $i$ and/ or $j$; $w_{ijh} = 1$ when information is present for both species (Podani 2000; Legendre and Legendre 2012). $w_{ijh}$ may also be used as weights (between 0 and 1), instead of being used as Kronecker's delta (Legendre and Legendre 2012).

Based on their functional similarity, every species $i$ among $s$ species in each trophic level specifies a fuzzy set to which every species $j$ ($j = 1$ to $s$ species, including species $i$) in the same trophic level belongs to with a certain degree of belonging ($u_{ij}$), taking values in a unit interval. The degrees of belonging of each species are the elements comprised in the new matrices $\mathbf{U_A}$ or $\mathbf{U_B}$ (Fig. 2b) and are standardized to unit column total by

$$u_{ij} = \frac{s_{ij}}{\sum_{k=1}^{s} s_{kj}} \tag{2}$$

where $s_{ij}$ is the pairwise similarity in the range 0–1, in our definition, Gower's similarity.

Using matrix multiplication, $\mathbf{Y_A} = \mathbf{U_A N_{AB}}$ and $\mathbf{Y_B} = \mathbf{N_{AB} U'_B}$, we can estimate the probability of a species in one trophic level interacting with a species in the other trophic level, weighted by their trait similarities (Fig. 2b). Using the same procedure, matrices of phylogenetic similarities $\mathbf{A_A}$ and $\mathbf{A_B}$ are transformed into fuzzy set matrices ($\mathbf{Q_A}$ and $\mathbf{Q_B}$; Fig. 2b), and then scaled up to the network level, generating matrices $\mathbf{P_B}$ and $\mathbf{P_A}$, corresponding to the probability of species interactions weighted by species phylogenetic resemblance in each trophic level (Fig. 2b). Matrices $\mathbf{Y_A}$, $\mathbf{Y_B}$, $\mathbf{P_A}$ and $\mathbf{P_B}$ contain the probability of interaction between species that are originally described in matrix $\mathbf{N_{AB}}$. However, the probability value for the same interaction will differ between matrices, since in each matrix the values are constrained by a different subset of traits (matrices $\mathbf{Y_A}$ and $\mathbf{Y_B}$) and phylogenetic distances (matrices $\mathbf{P_A}$ and $\mathbf{P_B}$).

## Dealing with Co-Occurrence

Ecologists usually assemble networks that result from sampling schemes that encompass different spatial (Winemiller 1990; Dáttilo et al. 2013) and temporal scales (Winemiller 1990; Díaz-Castelazo et al. 2010; Johansson et al. 2015).

Nonetheless, in the resulting network some interactions may never happen because a pair of species may not co-occur either in time (e.g., phenological differences) or space (e.g., species inhabit different habitat patches), despite their phylogenetic- or trait-similarity. To illustrate the point, imagine two predator species that have similar probabilities of interaction with a specific prey, either because they are functionally or phylogenetically similar, but one of them may not consume the same resource species because they never occur at the same place and time (e.g., spatial and phenological decoupling), either by chance or by biological reasons. Matrix $\mathbf{F}$ may be used to take this into account or to reveal spatial and temporal patterns. Using Hadamard element-wise multiplier, we can multiply matrices $\mathbf{Y_A}$, $\mathbf{Y_B}$, $\mathbf{P_A}$ and $\mathbf{P_B}$ by matrix $\mathbf{F}$ (Fig. 2c). This multiplication will result in four new matrices (Fig. 2c), which contain interaction probabilities, either weighted by trait- or phylogenetic resemblance, but corrected by either spatial or temporal mismatch between species from different trophic levels.

## Matrix Correlations and Null Model

The probabilistic matrices defined before can be used in a series of matrix correlation tests in order to detect (i) the phylogenetic component of interactions, (ii) phylogenetic signal in traits in both trophic levels and (iii) the functional component of interactions (Fig. 2d). The statistical significance of each matrix correlation is evaluated by permutation against a null model, which is explained in details bellow. Using Mantel matrix correlation, the correlation $\rho(\mathbf{P_A P_B})$ reveals patterns of the phylogenetic component of interactions which might be an evidence of cospeciation or, more generally, cocladogenesis (Fig. 2d), in which two or more interacting lineages undergo matched branching events during their phylogenetic history (Thompson 2014). Phylogenetic signal in traits are evaluated through the matrix correlations $\rho(\mathbf{Y_A P_A})$ and $\rho(\mathbf{Y_B P_B})$ (Fig. 2d). When the matrix correlation $\rho(\mathbf{Y_A Y_B})$ is high, we may infer that the observed interactions are largely due to functional traits of species in both trophic levels, e.g. plant and animal species. However, as we pointed out earlier, traits may present a significant degree of phylogenetic autocorrelation (Felsenstein 1985; Losos 2008), which makes necessary to consider the influence of phylogeny into this trait-based analysis in order to appreciate the independent effect of traits on species interaction. At first, we used higher-order partial matrix correlation to remove the effect of phylogenies and estimate the independent effect of traits, but this approach resulted in tests that were inaccurate in terms of Type I Error rates (see "Simulation Study to Assess Type I Error Rates" section for more details). Thus, to assess the independent trait component of interactions, we used

a slightly different approach removing the phylogenetic component of interactions by computing linear regressions using the vector that contains species pairwise phylogenetic similarities as the predictor variable and the vector containing species pairwise functional similarities as response variable (Fig. 3). These vectors are obtained from the original phylogenetic similarity matrices and by calculating similarity matrices from the original trait matrices from each trophic level (Fig. 3). The residuals of these regressions are then used in the scaling up process, resulting in two new matrices, $U_{resA}$ and $U_{resB}$ (Fig. 3). By matrix multiplication,

we obtain new trait-weighted interaction matrices removing the effect of phylogenies: $R_A = U_{resA}N_{AB}$ and $R_B = N_{AB}U'_{resB}$. The matrix correlation $\rho(R_A R_B)$ will express the correlation between these trait-weighted interaction matrices independent of the effect of phylogenies (Fig. 3).

The statistical significance of each zero-order matrix correlation, based on distance matrices, is then evaluated by permutation against a null model in which the interaction matrix is independent from the interacting species traits or phylogeny. This is achieved by random permutation among the column vectors in one of the fuzzy weighted matrices

**Fig. 3** Method used to assess the functional component of interactions removing the effect of phylogeny. **a** Scaling up of traits to the network level, removing the phylogenetic signal, based on linear models (LM). **b** Matrix correlation model used to test the effect of traits in species interactions
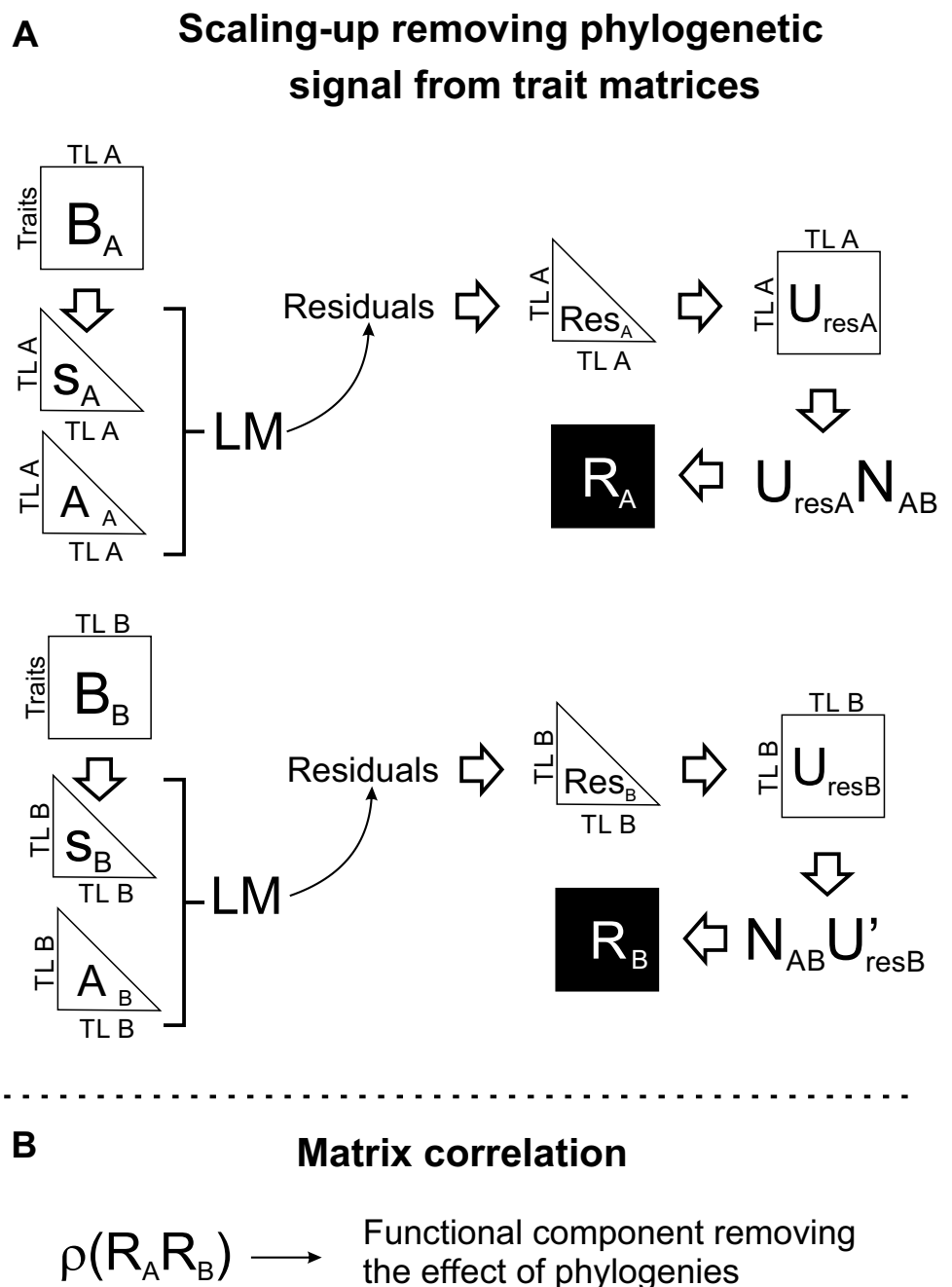
($\mathbf{U_A}$, $\mathbf{U_B}$, $\mathbf{Q_A}$ or $\mathbf{Q_B}$ and related matrices; Figs. 2b, 3a). The column vectors in these matrices represent the fuzzy sets defined by each and every species, which are kept intact when permuted, generating a new random matrix. This random rearrangement is repeated many times and new values of matrix correlations are computed and compared to the observed value; P values for the test will be the proportion of matrix correlations statistics ($\rho$) larger than or equal to the observed $\rho$ in a large number of random permutations ($\mathrm{P}[\rho^{null} \geq \rho^{observed}]$).

## Computer Software

The methods described here have been implemented in the R (R Core Team 2012) package SYNCSA (Debastiani and Pillar 2012) and the code is available in the Online Resource 1.

## Simulation Study to Assess Type I Error Rates

We evaluated the accuracy of the test by investigating the rates of Type I Error (the incorrect rejection of a true null hypothesis) using large sets of simulated data (50,000 iterations in each test). Simulations were implemented as to represent realistic biological scenarios, with interaction networks varying in size, number of functional traits and the number of interactions, and with species traits evolving according to distinct evolutionary models. However, as we were interested in the accuracy of the test, a pure stochastic process was used to determine species interactions. Simulations were implemented in several steps, which are described in detail below:

a. We first produced simulated phylogenetic trees of different sizes, from 5 to 20 species for each trophic level, which gave us networks (see step *d*) that varied in size, from 10 to 40 species. The size of the networks used in our simulations is consistent with the size of sampled ecological networks, which are usually small, comprising, on average, 23 species as suggested by Jonsson (2014);

b. We simulated the evolution of a small number of continuous characters (1–8 traits) along with their phylogenetic trees, under an Ornstein–Uhlenbeck model, which simulates different levels of constraints, such as stabilizing selection towards an optimum trait value for all species in each phylogenetic lineage (Hansen 1997; Blomberg et al. 2003). The choice of the number of traits is also congruent with evidence that suggests that very few traits are necessary to account for species interaction (Cohen 1977; Allesina et al. 2008;

Eklöf et al. 2013). The Ornstein–Uhlenbeck model is governed by one special parameter ($\alpha$), which directs the strength of the stabilizing selection towards an optimum trait value; larger values of $\alpha$ indicate that trait values will be closer to the optimum value, thus showing less phylogenetic signal (Blomberg et al. 2003; Harmon et al. 2010). We simulated different sets of parameter values, in order to determine the strength of character evolution towards the optimum trait value. The range of parameter choices used in the simulations allowed trait evolution to either follow a Brownian process ($\alpha \rightarrow 0$), as if evolution of traits followed a random walk through evolutionary time, or to follow a strong stabilizing selection towards an optimum trait value ($\alpha \rightarrow 1$).

c. As we were interested in estimating Type I Error rates, we simulated qualitative interaction matrices based on a random process, using a binomial distribution, where each cell of each simulated network has a probability of success, i.e., that a pair of species from different trophic levels interact, equal to 0.5. Therefore, the resulting network is determined solely by chance, without any regard to traits or phylogenies.

d. We then tested three sets of matrix correlations, in order to assess the following eco-evolutionary patterns (see Figs. 2d, 3b): (1) phylogenetic component, i.e., the association between phylogenies; and the functional component of interactions (2) removing the effect of phylogeny and (3) without removing it. The significance level to reject the null hypothesis was set a priori at $\alpha = 0.05$. Each test was deemed accurate if Type I Error was within 99.9% confidence limits ($0.05 \pm 0.0032$ for 50,000 simulated datasets). The R code for the simulation study is available in the Online Resource 2.

The results of our simulation study show that our framework is accurate, with Type I Error within 99.9% confidence limits for P = 0.05 (Table 1). All tests behave adequately in terms of Type I Error, rejecting a true null hypothesis in ~5% of the simulated datasets, independent

**Table 1** Simulation results for Type I Error of the matrix correlation tests proposed in this paper, based on 50,000 simulated data sets. Tests were considered accurate in terms of Type I Error if the rejection rates fell within the 99.9% confidence limits ($0.05 \pm 0.0032$)

|  | Proportion of tests |
| --- | --- |
| Phylogenetic component | 0.0495 |
| Functional component removing phylogenetic signal | 0.0496 |
| Functional component without removing phylogenetic signal | 0.0496 |

of network size, number of traits and the strength of stabilizing selection. The results for each subset of parameters can be seen in the Online Resource 3. In a first simulation study, we tried to estimate the functional component of interactions by removing the phylogenetic signal using simple first order matrix correlations (akin to partial Mantel test) using the phylogenetic and trait matrices after they have been scaled up to the network level (matrices $Y_A$, $Y_B$, $P_A$ and $P_B$ in Fig. 2). However, the results of this first simulation study showed these higher order matrix correlations were inaccurate, with very low rates of Type I Error, in which true null hypotheses were rejected in 0.3% of the simulated datasets. In order to circumvent this problem, we used the procedure described above (see section "Matrix Correlations and Null Model" and Fig. 3) by scaling up the residual matrices from phylogenetic and trait matrices. This procedure proved to be accurate in terms of Type I Error (Table 1).

## An Example from a Mutualistic Network

In order to illustrate our approach, we used data from a seed dispersal network from a grassland-forest mosaic from southern Brazil (30°25′S; 52°21′W). Birds were captured with mist-nets, which were placed in grasslands near to forest patches. Mist nets were set up monthly for eight consecutive days, between July 2007 and June 2008. Birds captured in mist nets were then placed into fabric bags for 20 min, and their feces were collected from bags. The seeds found in fecal samples were identified to species, and with this information at hand it was possible to build an interaction matrix between birds and the plant they consumed, with the frequencies of interaction (i.e., the number of times a specific bird species consumed a specific plant species). This interaction matrix consists of 128 interactions between 22 plant species, and 12 bird species (see Fig. 4).

We measured five plant traits (dispersule diameter, maximum plant height, aril presence, diaspore shape and color) and four bird traits (body length, beak length, beak height and the hand-wing index; for further details, see Online Resource 4). Using the APG III megatree (R20100701. new; Angiosperm Phylogeny Group 2009), we constructed a base family-level phylogeny. Due to the lack of within-family resolution, species and genera were represented as soft polytomies within genera and families, respectively. Branch lengths were added to the phylogeny using the BLADJ algorithm (Webb et al. 2008), which constrains the age of nodes included in the sample according to the dated molecular phylogeny of Wikström et al. (2001). For birds, we used the phylogenetic trees available at birdtree.org (Jetz et al. 2012). We then built a strict consensus tree using Mesquite 2.75 (Maddison and Maddison 2007). As

for plant species, bird taxa were also represented as soft polytomies. Branch lengths of terminal taxa were arbitrary lengthened in the consensus tree using the function "ultra-metricize" in Mesquite. Phylogenetic distances between all pairs of species were then calculated using the PHYDIST procedure of Phylocom (Webb et al. 2008). The data are available at the Interaction Web Database (https://www.nceas.ucsb.edu/interactionweb/).

We employed our framework to scale phylogenetic and trait data to the network level. As sampling lasted nearly 1 year, we used Hadamard multiplier to multiply all four probabilistic matrices by the F matrix, which is an incidence matrix, where 1 s mean that at least once, during the sampling period, a bird species occurred in the study site when a specific plant species was fructifying, and 0 s mean otherwise. We then tested four matrix correlation models, using Euclidean distance as a resemblance measure: phylogenetic signal in traits in each trophic level; correlation among traits accounting for species interactions, removing the effect of phylogenies and; the correlation between phylogenies, in order to detect patterns of cocladogenesis.

The results are shown in Table 2. Among the four tested models, the only significant correlation was the one between bird and plant traits, showing that birds must have specific traits in order to consume their plant resources, and that phylogenetic resemblance has no explanatory power for the evolution of species traits and species interactions in this seed-dispersal network. As the benefits of plant-animal disperser interactions are mutual, ecologists and evolutionary biologists have long pondered whether this sort of mutualistic interaction is constrained by strong evolutionary forces (Snow 1971; Morton 1978; Howe 1993; Muller-Landau and Hardesty 2005). Theoretical models and empirical data suggest that plant-seed disperser interactions are a system of low specialization (Herrera 1995; Muller-Landau and Hardesty 2005), in which a species from one trophic level can interact with many species of the other level. Moreover, it has been suggested that frugivory is a foraging mode that does not require special adaptation in traits involved in the interaction (Jordano 2000). Our results seem to corroborate the idea that plant-seed disperser networks exhibit low specialization, in terms of interacting species, and are not governed by strong evolutionary constraints. Alternatively, the lack of correlation between phylogenies may indicate selection imposed by mutualistic interactions strongly affecting trait evolution, as expected in mutualisms among free-living species and based on trait-matching (Thompson 2006; Guimarães et al. 2011). In fact, our results also demonstrate that some level of trait complementarity is relevant for determining interaction patterns, despite the high plasticity of the foraging spectrum of birds. Although one might expect trait evolution to be influenced by the phylogenetic relationship among species, it is now
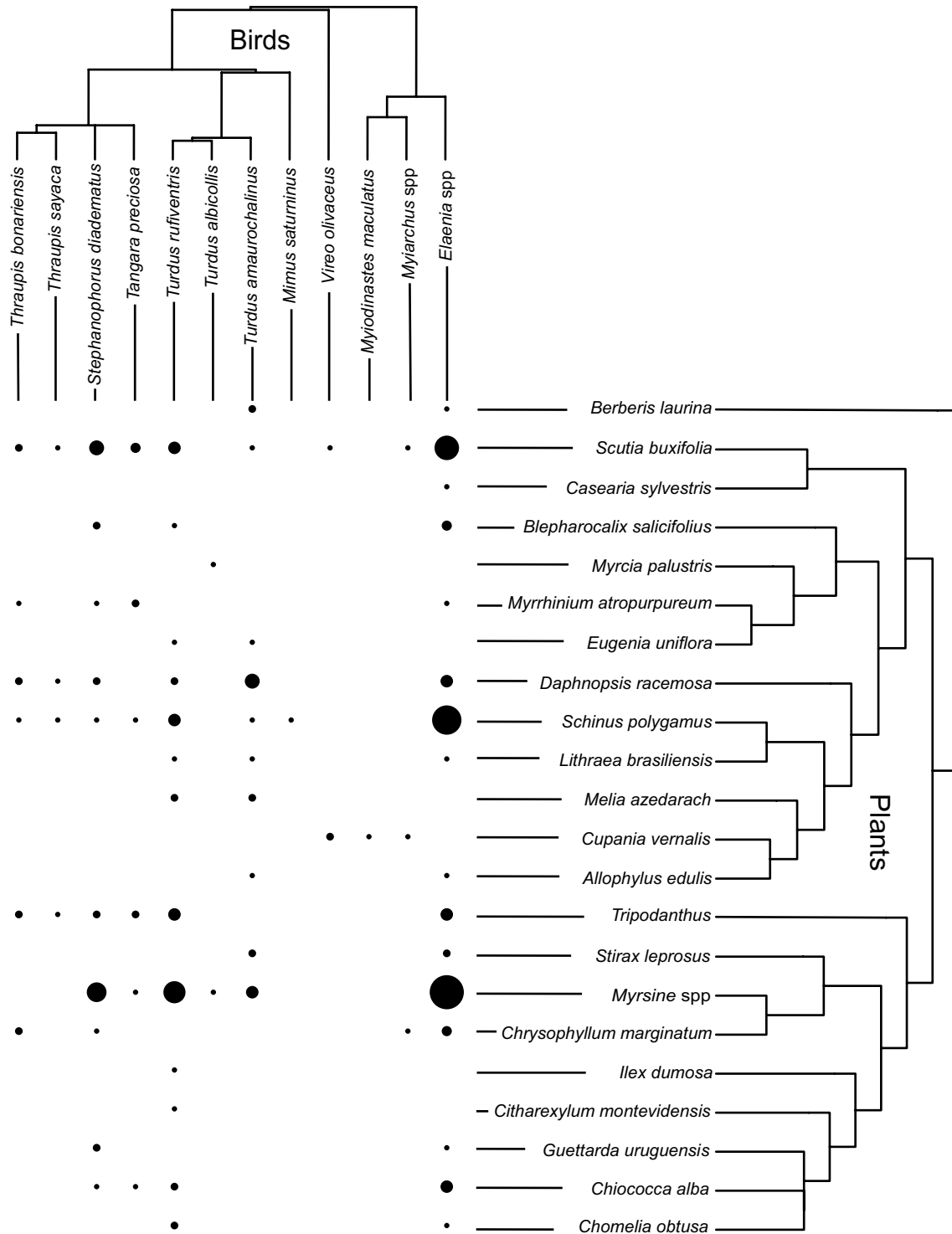
**Fig. 4** Seed-dispersal interactions in southern Brazil, with the phylogeny of each trophic level represented in the margins. *Circle* sizes are proportional to the frequency of the interaction

widely recognized that traits vary tremendously in their phylogenetic pattern (Blomberg et al. 2003), and recent evidence suggest that this relationship should be expected only under some very specific conditions that are seldom

expected to hold in mutualisms among free-living species (Thompson 1994). For example, Nuismer and Harmon (2015) have demonstrated that mutualistic networks, such as our example, should exhibit less phylogenetic signal in

**Table 2** Summary statistics for four models for the seed dispersal network from southern Brazil

|  | $\rho$ | P value |
| --- | --- | --- |
| Phylogenetic component | 0.06 | 0.84 |
| Functional component, removing phylogenetic signal | 0.49 | <0.001* |
| Phylogenetic signal in plant species | 0.28 | 0.99 |
| Phylogenetic signal in bird species | 0.25 | 0.94 |

*Results deemed significant at $\alpha = 0.05$

rates of interactions than competitive ones, and that if interactions within a network depend on a mechanism of phenotype differences, phylogenetic information has little predictive power for trait evolution and interaction rates.

## Discussion and Final Remarks

Understanding how species interact and how the topology of ecological networks influences the dynamics of populations and communities is at the core of ecological inquiry and has been challenging ecologists for over a century now (Camerano 1880; Paine 1980; Polis and Strong 1996; Schleuning et al. 2015). Despite the long tradition, the study of complex ecological networks has dramatically increased in the past two decades (Miranda et al. 2013). Nonetheless, most empirical studies still focus on the description of network properties, such as size (i.e., species richness within the network), connectivity and degree distribution (Miranda et al. 2013). Only recently, there has been a growing interest of moving beyond the description of topological patterns and the integration of other types of data, such as phylogenetic and trait data, to reveal patterns of species interactions (e.g., Rezende et al. 2007; Rafferty and Ives 2013; Hadfield et al. 2014; Vizentin-Bugoni et al. 2014). Here, we developed a statistical framework based on fuzzy set theory and Monte Carlo statistics capable of identifying eco-evolutionary patterns that might suit some of this interest. More specifically, our method is intended to answer the following ecological questions: Is there phylogenetic signal in traits accounting for species interaction? Are species interactions driven by trait resemblance? Are species interactions driven by phylogenetic resemblance (i.e., is there evidence of cocladogenesis?)? At this point, however, it is important to note that interpreting cophylogenetic analyses can be extremely difficult and finding correlated phylogenies does not necessarily indicate a history of coevolution or cospeciation (Banks and Paterson 2005; Thompson 2005). Accordingly, coevolution does not imply in cospeciation, cocladogenesis, or in correlated phylogenies (Thompson 1994, 2005). For instance, in host-parasite

networks, gradual host-switching by parasite species, followed by speciation events may lead to "false" congruent phylogenies (Banks and Paterson 2005). Thus, results should be interpreted with caution. However, our cophylogenetic analysis can help ecologists to understand network patterns and bring important information to applied matters, such as choosing target species to conservation or restoration plans, as it allows us to identify pair of species with higher chance of interacting, and the role of phylogenetic- and/or trait-resemblance, even if different processes may lead to the same apparent pattern.

Our method is less restricted by statistical assumptions than model-based approaches, it is relatively easy to apply and can be used with both qualitative and quantitative species interaction matrices. Other advantage of the method is that it is suitable for a range of data types, allowing investigators to use qualitative (expanded as dummy variables) and/or quantitative traits. Also it does not require the availability of dated phylogenetic trees, as the method is based on resemblance matrices. In situations where phylogenetic trees are not available, distances can be computed from raw data such as sequences in databases. Moreover, our method could be easily expanded to include other kind of ecological information, such as environmental variables. By using higher order matrix correlations, one could access the effects of phylogeny, traits and the underlying environmental gradient on the architecture of networks.

Hadfield et al. (2014) evaluated the accuracy of the cophylogenetic permutation tests proposed by Legendre et al. (2002), Homola et al. (2009) and the model-based approach proposed by Ives and Godfray (2006). Their results show that in the presence of phylogenetic signal, and when evolutionary interactions effects are present, these three methods give severely inflated Type I Error rates when trying to estimate the correlation between phylogenies in a bipartite network. Contrary to these methods and despite its resemblance with the Mantel and partial Mantel approach (Smouse et al. 1986; Legendre and Fortin 2010; Debastiani and Duarte 2016), which has been criticized by its low power, and under some circumstances, inflated Type I Error rates (Lapointe and Legendre 1995; Harmon and Glor 2010; Legendre and Fortin 2010; but see Debastiani and Duarte 2016), our statistical framework gave accurate rejection rates, rejecting a true null hypothesis in ~5% of the simulated datasets, even in the presence of strong phylogenetic signal, as shown in our simulation study.

We additionally point out another possible application of the framework developed here. The method could be used to construct an interaction probability matrix for unobserved realizations, and predict possible interactions. Using the scaling up process we presented here, one could predict the interaction probability of species that are known to occur in the study site, but were not observed interacting

with other species. These species could be included in the interaction matrix, with zero values across their rows or columns. Then, one could estimate the probability of interaction with each and every species in the other trophic level, based on species trait- of phylogenetic- resemblance. Although this possibility has yet to be tested, it is a potential feature of the method, which can account for sampling problems or be used in other applied matters, when one has to estimate the probability of interactions that were not observed directly in the field.

To conclude we stress that the statistical method we developed here and its extension can de used to tackle a wide range of eco-evolutionary questions. We hope that further developments broaden the range of possible applications and our understanding of underlying mechanisms driving patterns in ecological networks.

**Compliance with Ethical Standards**

# References

Allesina, S., Alonso, D., & Pascual, M. (2008). A general model for food web structure. *Science, 320*, 658–661.

Angiosperm Phylogeny Group. (2009). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society, 161*, 105–121.

Banks, J. C., & Paterson, A. M. (2005). Multi-host parasite species in cophylogenetic studies. *International Journal for Parasitology, 35*, 741–746.

Bersier, L. F., & Kehrli, P. (2008). The signature of phylogenetic constraints on food-web structure. *Ecological Complexity, 5*, 132–139.

Blomberg, S. P., Garland, T., & Ives, A. R. (2003). Testing for phylogenetic signal in comparative data: Behavioral traits are more labile. *Evolution, 57*, 717–745.

Camerano, L. (1880). Dell'equilibrio dei viventi merc la reciproca distruzione. *Atti della Reale Accademia delle Scienze di Torino, 15*, 393–414.

Cavender-Bares, J., Kozak, K. H., Fine, P. V., & Kembel, S. W. (2009). The merging of community ecology and phylogenetic biology. *Ecology Letters, 12*, 693–715.

Cohen, J. E. (1977). Food webs and the dimensionality of trophic niche space. *Proceedings of the National Academy of Sciences, 74*, 4533–4536.

Cruz, C. P., Fonseca, C. R., & Corso, G. (2012). Ecological interaction and phylogeny, studying functionality on composed networks. *Physica A: Statistical Mechanics and its Applications, 391*, 673–679.

Danielson, B. J. (1991). Communities in a landscape: The influence of habitat heterogeneity on the interactions between species. *The American Naturalist, 138*, 1105–1120.

Darwin, C. (1859). *On the origin of species by means of natural selection*. London: J. Murray.

Dáttilo, W., Guimarães, P. R. & Izzo, T. J. (2013). Spatial structure of ant–plant mutualistic networks. *Oikos, 122*, 1643–1648.

Debastiani, V. J., & Duarte, L. S. (2016). Evolutionary models and Phylogenetic signal assessment via mantel test. *Evolutionary Biology*, 1–9.

Debastiani, V. J., & Pillar, V.D. P. (2012). SYNCSA: R tool for analysis of metacommunities based on functional traits and phylogeny of the community components. *Bioinformatics (Oxford, England), 28*, 2067–2068.

Díaz-Castelazo, C., Guimarães, P. R., Jordano, P., Thompson, J. N., Marquis, R. J., & Rico-Gray, V. (2010). Changes of a mutualistic network over time: Reanalysis over a 10-year period. *Ecology, 91*, 793–801.

Eklöf, A., Jacob, U., Kopp, J., et al. (2013). The dimensionality of ecological networks. *Ecology Letters, 16*, 577–583.

Felsenstein, J. (1985). Phylogenies and the comparative method. *The American Naturalist, 125*, 1–15.

Gómez, J. M., Verdú, M., & Perfectti, F. (2010). Ecological interactions are evolutionarily conserved across the entire tree of life. *Nature, 465*, 918–921.

Gross, T., & Blasius, B. (2008). Adaptive coevolutionary networks: A review. *Journal of the Royal Society Interface, 5*, 259–271.

Guimarães, P. R. Jr., Jordano, P., & Thompson, J. N. (2011). Evolution and coevolution in mutualistic networks. *Ecology Letters, 14*, 877–885.

Hadfield, J. D., Krasnov, B. R., Poulin, R., & Nakagawa, S. (2014). A tale of two phylogenies: Comparative analyses of ecological interactions. *The American Naturalist, 183*, 174–187.

Hansen, T. F. (1997). Stabilizing selection and the comparative analysis of adaptation. *Evolution, 51*, 1341–1351.

Harmon, L. J., & Glor, R. E. (2010). Poor statistical performance of the Mantel test in phylogenetic comparative analyses. *Evolution, 64*(7), 2173–2178.

Harmon, L. J., Losos, J. B., Jonathan Davies, T., et al. (2010). Early bursts of body size and shape evolution are rare in comparative data. *Evolution, 64*, 2385–2396.

Herrera, C. M. (1995). Plant-vertebrate seed dispersal systems in the Mediterranean: Ecological, evolutionary, and historical determinants. *Annual Review of Ecology and Systematics, 26*, 705–727.

Hommola, K., Smith, J. E., Qiu, Y., & Gilks, W. R. (2009). A permutation test of host–parasite cospeciation. *Molecular Biology and Evolution, 26*, 1457–1468.

Howe, H. F. (1993). Specialized and generalized dispersal systems: Where does the paradigm stand? *Vegetatio, 107*, 3–13.

Ibanez, S. (2012). Optimizing size thresholds in a plant–pollinator interaction web: Towards a mechanistic understanding of ecological networks. *Oecologia, 170*, 233–242.

Ives, A. R., & Godfray, H.C.J. (2006). Phylogenetic analysis of trophic associations. *The American Naturalist, 168*, E1–E14.

Jermy, T. (1976). Insect–host-plant relationship: Co-evolution or sequential evolution? In Jermy (Ed.), *The host-plant in relation to insect behaviour and reproduction* (pp. 109–113). New York: Springer.

Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., & Mooers, A. O. (2012). The global diversity of birds in space and time. *Nature, 49*, 444–448.

Johansson, J., Nilsson, J. Å., & Jonzén, N. (2015). Phenological change and ecological interactions: An introduction. *Oikos, 124*, 1–3.

Johnson, M.T.J., & Stinchcombe, J. R. (2007). An emerging synthesis between community ecology and evolutionary biology. *Trends in Ecology and Evolution, 22*, 250–257.

Jonsson, T. (2014). Trophic links and the relationship between predator and prey body sizes in food webs. *Community Ecology, 15*, 54–64.

Jordano, P. (2000). Fruits and frugivory. In Fenner, P. (Ed.), *Seeds: The ecology of regeneration in plant communities* (pp. 125–162). Wallingford: CABI Publishing.

Krishna, A., Guimaraes, P. R. Jr., Jordano, P., & Bascompte, J. (2008). A neutral-niche theory of nestedness in mutualistic networks. *Oikos, 117*, 1609–1618.

Lapointe, F. J., & Legendre, P. (1995). Comparison tests for dendrograms: A comparative evaluation. *Journal of Classification, 12*(2), 265–282.

Legendre, P., Desdevises, Y., & Bazin, E. (2002). A statistical test for host–parasite coevolution. *Systematic biology, 51*, 217–234.

Legendre, P., & Fortin, M. J. (2010). Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources, 10*(5), 831–844.

Legendre, P., & Legendre, L. F. (2012). *Numerical ecology*. Amsterdam: Elsevier.

Lewinsohn, T. M., Novotny, V., & Basset, Y. (2005). Insects on plants: Diversity of herbivore assemblages revisited. *Annual Review of Ecology, Evolution, and Systematics, 36*, 597–620.

Losos, B. J. (2008). Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecology Letters, 11*, 995–1007.

Maddison, W. P., & Maddison, D. R. (2007). *Mesquite: a modular system for evolutionary analysis*. Version 2.75. 2011. URL http://mesquiteproject.org.

Miranda, M., Parrini, F., & Dalerum, F. (2013). A categorization of recent network approaches to analyse trophic interactions. *Methods in Ecology and Evolution, 4*, 897–905.

Mitter, C., Farrell, B., & Futuyma, D. J. (1991). Phylogenetic studies of insect-plant interactions: Insights into the genesis of diversity. *Trends in Ecology and Evolution, 6*, 290–293.

Morton, E. S. (1973). On the evolutionary advantages and disadvantages of fruit eating in tropical birds. *The American Naturalist, 107*, 8–22.

Mouquet, N., Devictor, V., Meynard, C. N., et al. (2012). Ecophylogenetics: Advances and perspectives. *Biological Reviews, 87*, 769–785.

Muller-Landau, H. C., & Hardesty, B. D. (2005). Seed dispersal of woody plants in tropical forests: Concepts, examples and future directions. In D. F. Burslem et al. (Eds.), *Biotic interactions in the tropics: Their role in the maintenance of species diversity* (pp. 267–309). Cambridge: Cambridge University Press.

Novotny, V., & Basset, Y. (2005). Host specificity of insect herbivores in tropical forests. *Proceedings of the Royal Society of London B: Biological Sciences, 272*, 1083–1090.

Nuismer, S. L., & Harmon, L. J. (2015). Predicting rates of interspecific interaction from phylogenetic trees. *Ecology Letters, 18*, 17–27.

Page, R. D. (2003). *Tangled trees: Phylogeny, cospeciation, and coevolution*. Chicago: University of Chicago Press.

Page, R. D., & Charleston, M. A. (1998). Trees within trees: Phylogeny and historical associations. *Trends in Ecology and Evolution, 13*, 356–359.

Paine, R. T. (1980). Food webs: Linkage, interaction strength and community infrastructure. *Journal of Animal Ecology, 49*, 667–685.

Pascual, M., & Dunne, J. A. (2006). *Ecological networks: Linking structure to dynamics in food webs*. Oxford: Oxford University Press.

Pearse, I. S., & Hipp, A. L. (2009). Phylogenetic and trait similarity to a native species predict herbivory on non-native oaks. *Proceedings of the National Academy of Sciences, 106*, 18097–18102.

Pedron, M., Buzatto, C. R., Singer, R. B., Batista, J. A., & Moser, A. (2012). Pollination biology of four sympatric species of Habenaria (Orchidaceae: Orchidinae) from southern Brazil. *Botanical Journal of the Linnean Society, 170*, 141–156.

Pillar, V. D. (2013). How accurate and powerful are randomization tests in multivariate analysis of variance? *Community Ecology, 14*, 153–163.

Pillar, V. D., & Duarte, L. D.S. (2010). A framework for metacommunity analysis of phylogenetic structure. *Ecology Letters, 13*(5), 587–596.

Pillar, V. D. P., Duarte, L. D.S., Sosinski, E. E., & Joner, F. (2009). Discriminating trait-convergence and trait-divergence assembly patterns in ecological community gradients. *Journal of Vegetation Science, 20*, 334–348.

Pillar, V. D. P., & Orlóci, L. (1991). Fuzzy components in community level comparisons. In E. Feoli & L. Orlóci (Eds.), *Computer assisted vegetation analysis* (pp. 87–93). Berlim: Kluwer Academic Publishers.

Podani, J. (2000). *Introduction to the exploration of multivariate biological data*. Leiden: Backhuys Publishers.

Polis, G. A., & Strong, D. R. (1996). Food web complexity and community dynamics. *The American Naturalist, 147*, 813–846.

R Core Team. (2012). *R: A language and environment for statistical computing. R foundation for statistical computing*, Vienna, URL http://www.R-project.org/.

Rafferty, N. E., & Ives, A. R. (2013). Phylogenetic trait-based analyses of ecological networks. *Ecology, 94*, 2321–2333.

Rezende, E. L., Lavabre, J. E., Guimarães, P. R., Jordano, P., & Bascompte, J. (2007). Non-random coextinctions in phylogenetically structured mutualistic networks. *Nature, 448*, 925–928.

Rossberg, A. G. Brännström, Å., & Dieckmann, U. (2010). How trophic interaction strength depends on traits. *Theoretical Ecology, 3*, 13–24.

Schleuning, M., Fründ, J., & García, D. (2015). Predicting ecosystem functions from biodiversity and mutualistic networks: An extension of trait-based concepts to plant–animal interactions. *Ecography, 38*, 380–392.

Schmitz, O. J., Buchkowski, R. W., Burghardt, K. T., & Donihue, C. M. (2015). Functional traits and trait-mediated interactions: Connecting community-level interactions with ecosystem functioning. *Advances in Ecological Research, 52*, 319–343.

Smouse, P. E., Long, J. C., & Sokal, R. R. (1986). Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Systematic Zoology, 35*(4), 627–632.

Snow, D. W. (1971). Evolutionary aspects of fruit-eating by birds. *Ibis, 113*, 194–202.

Strauss, S. Y., & Irwin, R. E. (2004). Ecological and evolutionary consequences of multispecies plant–animal interactions. *Annual Review in Ecology, Evolution and Systematics, 35*, 435–466.

Thompson, J. N. (1994). *The coevolutionary process*. Chicago: University of Chicago Press.

Thompson, J. N. (2005). *The geographic mosaic of coevolution*. Chicago: University of Chicago Press.

Thompson, J. N. (2006). Mutualistic webs of species. *Science, 312,* 372–373.

Thompson, J. N. (2014). Coevolution and speciation. In J. Losos. (Ed.), *The Princeton guide to evolution* (pp. 535–548). Princeton: Princeton University Press.

Tylianakis, J. M., Tscharntke, T., & Lewis, O. T. (2007). Habitat modification alters the structure of tropical host–parasitoid food webs. *Nature, 445,* 202–205.

Vázquez, D. P., Blüthgen, N., Cagnolo, L., & Chacoff, N. P. (2009). Uniting pattern and process in plant–animal mutualistic networks: A review. *Annals of Botany, 103,* 1445–1457.

Verdú, M., & Valiente-Banuet, A. (2011). The relative contribution of abundance and phylogeny to the structure of plant facilitation networks. *Oikos, 120,* 1351–1356.

Vizentin-Bugoni, J., Maruyama, P. K., & Sazima, M. (2014). Processes entangling interactions in communities: Forbidden links are more important than abundance in a hummingbird–plant network. *Proceedings of the Royal Society of London B: Biological Sciences, 281,* 20132397.

Webb, C. O. (2000). Exploring the phylogenetic structure of ecological communities: An example for rain forest trees. *The American Naturalist, 156,* 145–155.

Webb, C. O., Ackerly, D. D., & Kembel, S. W. (2008). Phylocom: Software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics (Oxford, England), 24,* 2098–2100.

Webb, C. O., Ackerly, D. D., McPeek, M. A., & Donoghue, M. J. (2002). Phylogenies and community ecology. *Annual Review of Ecology and Systematics, 33,* 475–505.

Weiblen, Webb, C, O., Novotny, V., Basset, Y., & Miller, S. E. (2006). Phylogenetic dispersion of host use in a tropical insect herbivore community. *Ecology, 87,* S62–S75.

Wikström, N., Savolainen, V., & Chase, M. W. (2001). Evolution of the angiosperms: Calibrating the family tree. *Proceedings of the Royal Society of London B: Biological Sciences, 268,* 2211–2220.

Winemiller, K. O. (1990). Spatial and temporal variation in tropical fish trophic networks. *Ecological Monographs, 60,* 331–367.

Zadeh, L. A. (1965). Fuzzy sets. *Information Control, 8,* 338–353.