

# ADVANCES AND CHALLENGES IN THE STUDY OF ECOLOGICAL NETWORKS

## A common framework for identifying linkage rules across different types of interactions

Ignasi Bartomeus<sup>\*,1</sup>, Dominique Gravel<sup>2</sup>, Jason M. Tylianakis<sup>3,4</sup>, Marcelo A. Aizen<sup>5</sup>, Ian A. Dickie<sup>6</sup> and Maud Bernard-Verdier<sup>6</sup>

<sup>1</sup>Estación Biológica de Doñana (EBD-CSIC), Avda. Américo Vespucio s/n, Isla de la Cartuja, E-41092 Sevilla, Spain;

<sup>2</sup>Département de biologie, Faculté des Sciences, Université de Sherbrooke, 2500 Boulevard Université, Sherbrooke, Quebec J1K 2R1, Canada; <sup>3</sup>Centre for Integrative Ecology, School of Biological Sciences, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand; <sup>4</sup>Department of Life Sciences, Imperial College London, Silwood Park Campus, Buckhurst Road, Ascot, Berkshire SL5 7PY, UK; <sup>5</sup>Laboratorio Ecotono-CRUB, Universidad Nacional del Comahue and INIBIOMA, Quintral 1250, 8400 San Carlos de Bariloche, Río Negro, Argentina; and <sup>6</sup>Bio-protection Research Centre, Lincoln University, PO Box 85084, Lincoln 7647, New Zealand

### Summary

1. Species interactions, ranging from antagonisms to mutualisms, form the architecture of biodiversity and determine ecosystem functioning. Understanding the rules responsible for who interacts with whom, as well as the functional consequences of these interspecific interactions, is central to predict community dynamics and stability.

2. Species traits *sensu lato* may affect different ecological processes by determining species interactions through a two-step process. First, ecological and life-history traits govern species distributions and abundance, and hence determine species co-occurrence and the potential for species to interact. Secondly, morphological or physiological traits between co-occurring potential interaction partners should match for the realization of an interaction. Here, we review recent advances on predicting interactions from species co-occurrence and develop a probabilistic model for inferring trait matching.

3. The models proposed here integrate both neutral and trait-matching constraints, while using only information about known interactions, thereby overcoming problems originating from undersampling of rare interactions (i.e. missing links). They can easily accommodate qualitative or quantitative data and can incorporate trait variation within species, such as values that vary along developmental stages or environmental gradients.

4. We use three case studies to show that the proposed models can detect strong trait matching (e.g. predator–prey system), relaxed trait matching (e.g. herbivore–plant system) and barrier trait matching (e.g. plant–pollinator systems).

5. Only by elucidating which species traits are important in each process (i.e. in determining interaction establishment and frequency), we can advance in explaining how species interact and the consequences of these interactions for ecosystem functioning.

**Key-words:** functional traits, herbivory, interaction networks, mutualisms, parasitism, pollination, predation, trait matching, trophic interactions

### Introduction

Species interactions form the architecture of biodiversity (Bascompte & Jordano 2007). There is growing recognition that community structure, stability and functioning depend not only on which species are present in a community, but

also on how they interact (Tylianakis *et al.* 2008). Complex networks of biotic interactions such as predation, parasitism and mutualism provide essential information related to conservation (Carvalho, Barbosa & Memmott 2008; Tylianakis *et al.* 2010), community stability and ecosystem functioning (Thompson *et al.* 2012; Peralta *et al.* 2014), and evolutionary processes (Jacquemyn *et al.* 2011; Fenster *et al.* 2015). These insights would be not

\*Correspondence author. E-mail: nacho.bartomeus@gmail.com

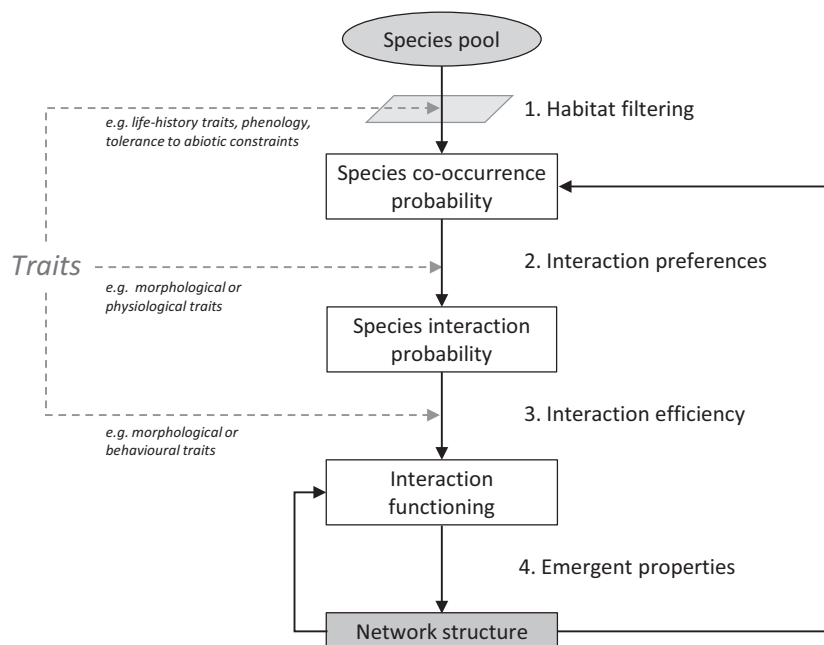
possible from simple species occurrence data or analysis of pairwise interactions. Despite the growing literature describing species interaction networks, we still have a poor understanding of how network structure comes to exist.

There are few generalizable observations of how species interactions respond to environmental changes (Tylianakis *et al.* 2008). Therefore, understanding what determines the occurrence of pairwise interactions, and, at a higher level, the structure of ecological networks, is a key challenge for ecologists. Overcoming this challenge requires the identification of the mechanisms responsible for who interacts with whom. Natural selection promotes adaptations to increase species efficiency (Castellanos, Wilson & Thomson 2003). Reciprocal trait adaptations between partners, which have positive demographic consequences, lead to increased interaction strength among co-evolved members (Sargent & Ackerly 2008). Hence, there is a great expectation that incorporating a trait-based approach can help us explain general mechanisms driving pairwise interactions. We refer here to traits in a broad sense, comprising adaptations that define organisms in terms of their ecological role, how they interact with the environment and with other species (Díaz & Cabido 2001). Most traits studied so far for predicting species interactions fall into morphological adaptations (e.g. body size), but physiological (e.g. chemical defences) or behavioural (e.g. diel) adaptations can also drive species interactions. Moreover, traits should be precise and measurable attributes of the species. Recent studies indeed suggest that ecological networks of different types (e.g. from antagonistic to mutualistic) could be described from the traits of the interacting species (Eklöf *et al.* 2013). The ability of these methods to predict novel interactions fol-

lowing species invasions or following range shifts is, however, limited.

Traits are implicated in ecological dynamics at several concatenated levels of community organization (Fig. 1) and therefore could influence the occurrence of interactions in multiple ways. Some traits determine species distributions in a multidimensional environmental space and thus impact co-occurrence in space and time. Since the occurrence of an interaction requires the presence of the two species, traits involved in phenological matching or habitat filtering could constrain interactions. Life-history traits impact demography, abundance and biomass, thereby affecting the probability of encounter. Then, provided they encounter each other in space and time, the compatibility between traits of the two species (i.e. trait-matching constraints) will also determine whether or not they interact. Finally, the intensity and the impact of an interaction will determine the functioning of the network and also feed back to determine species abundances and dynamics. How efficient an interacting species is on a per capita basis is also likely to be mediated by its behavioural or physiological traits and how these match with those of the other species. Of course, these levels interact with each other through evolutionary processes. Most work to date has focused on morphological trait matching, and little, if any, has tackled several of these stages at a time (see the review in Morales-Castilla *et al.* 2015). Our first objective here was to review what we know about each of these processes and assess their success and limitations at predicting interactions. Our second objective was to propose a way forward to evaluate trait matching in a way that is not confounded with species co-occurrences, and how this can be integrated into a larger framework, from species occurrences to ecosystem functioning.

**Fig. 1.** Species traits may influence the structure of interaction networks in three different ways. 1. Trait-based environmental filtering may determine species abundances in space and time, which will affect probability of encounter. 2. Given species co-occurrence, trait matching according to species interaction preferences will shape interaction probability. 3. Species traits might also influence the per capita efficiency and impact of an interaction, and thereby influence network functioning. In addition, 4. emergent properties inherent to the structure of the network will influence network functioning and feedback on community dynamics.



## Traits governing species encounters in space and time

Habitat filtering constrains the pool of co-occurring species in a region or microhabitat. Sharing habitat-filtering traits, like tolerance to drought or thermal preference, may hence be a prerequisite for two species to interact. Even in large and diffuse networks such as the global planktonic interactome, 18% of the variance in community composition (taxon presence and abundance) can be explained by environment alone, and these co-occurrences can be used successfully to predict interactions between taxa (Lima-Mendez *et al.* 2015). Microhabitat characteristics can also influence sessile organisms even within close proximity, as shown by interactions among mycorrhizas and plants, where rooting depth could preclude co-occurrence between shallow rooted plants and fungi restricted to lower soil horizons. In fact, the concept of 'habitat associations' as a driver of interactions has been pointed to as the sole explanation for these interactions (Zobel & Öpik 2014), suggesting that both partners interact simply because they respond independently to different environmental factors.

At broader spatial scales, species turnover along ecological gradients can also be responsible for a large fraction of network variation in space (Poisot *et al.* 2012). Range overlap determines the location and the total area over which two species can eventually interact. This can be used to better understand the consequences of range shifts on the local food web structure (Albouy *et al.* 2014). Species distribution models in combination with ecological and life-history traits (D'Amen *et al.* 2015) can be used to predict co-occurrence and potential interactions in response to global changes (Albouy *et al.* 2014; Morales-Castilla *et al.* 2015).

Similar to species distribution in space, species encounter will be determined by the synchrony of their activity periods at different temporal scales (i.e. daily, seasonal, interannual). Mismatch of phenology has been widely called to explain undetected interactions that are not possible to occur (i.e. forbidden interactions; Olesen *et al.* 2011; Encinas-Viso, Revilla & Etienne 2012; Olito & Fox 2015), that is species present in the same location that do not interact because they do not overlap in their seasonal activity periods. Phenological overlap during the season has therefore been used as a proxy for interaction probability (Bartomeus *et al.* 2013). While phenology is usually studied as the timing when species are active during a season (e.g. plant flowering period), daily fluctuations of activity can also be important for defining when interactions among partners can occur. A clear example is the distinction between crepuscular vs. diurnal species (Herrera 2000), but more subtle fluctuations of activity depending on daily temperature may be also relevant (Rader *et al.* 2013). In addition, some species may interact only with partners in a given life-history stage, for example some ectomycorrhizal fungi may require host trees to be at least

several years old and do not interact with seedlings. This highlights the importance and complexity of the temporal constraints on co-occurrence.

Given that species co-occur in space and time, their abundance also determines the frequency at which they will interact (Canard *et al.* 2014). Abundant species are simply more likely to encounter each other than rare ones. This mechanism has been called neutral because it does not rely on any niche differentiation. Thus, models that use species abundances to predict encounter probabilities have found that abundance alone can explain considerable variance in key aspects of network structure (Vázquez *et al.* 2007; Krishna *et al.* 2008; Olito & Fox 2015). Abundance is determined primarily by life-history traits (e.g. fecundity, longevity, mortality). For plant communities, there is some consensus over which traits relate to abundance or dominance in the community, such as maximal height and position along the slow-fast continuum (e.g. leaf economic spectrum; Wright *et al.* 2004). Therefore, trait distributions over environmental gradients have been used to predict plant abundance and community structure (Shipley, Vile & Garnier 2006; Laughlin *et al.* 2012). Similarly, it is possible to relate life-history traits to animal abundances. For instance, species with fast life cycles (usually small, with high reproduction rates and short longevity) tend to be more abundant than large species with slow life histories (White *et al.* 2007), and large species can decline more rapidly following habitat change (Larsen, Williams & Kremen 2005). As a result, abundance can be largely related to body size and position in the interaction network (Woodward *et al.* 2005). Overall, the relationships between traits, abundance and probability of encounter define the neutral expectations for interacting. This relationship is complex, for example because the encounter probability changes both as function of species traits (e.g. landscape use) and as a function of abundance (e.g. through density-dependent foraging).

## Trait matching

Trait matching between interacting partners has been identified for a variety of organisms. Plant corolla length and pollinator proboscis length is a classic example (Kritsky 1991). However, most pollinators are quite generalists, and while species may have specialized morphology, it does not prevent them from utilizing a diversity of resources (Waser *et al.* 1996). Bird beak size and fruit size has also been shown to be tightly related to dispersal success (Galetti *et al.* 2013). In fishes, predator mouth gap and prey size are also strong determinants of predatory interactions (Cunha & Planas 1999). More complex relationships have been found for plants too, with the role of specific leaf area in plant-plant interactions changing from facilitation to competition, depending on resource availability (Gross *et al.* 2009). Trait-matching constraints have been described for most interacting species, ranging from

arbuscular mycorrhizas and plants (Chagnon *et al.* 2013) to plants and herbivores (Deraison *et al.* 2015).

Trait matching between individuals operates in addition to neutral processes to impact pairwise interactions. Despite advances in these respective fields (e.g. null model analysis: Vázquez, Chacoff & Cagnolo 2009; trait matching analysis: Dehling *et al.* 2014; Spitz, Ridoux & Brind'A-mour 2014; Crea, Ali & Rader 2016), we still lack a common analytical framework with which to evaluate the contribution of species traits to pairwise interactions, and at the higher level to the structure of interaction networks.

Even though neutral and trait-based null models can predict the general structure of interaction networks, such models often are poor at predicting the occurrence and intensity of individual interactions (Vázquez, Chacoff & Cagnolo 2009; Olito & Fox 2015). Such models are useful because they free us from species identities and allow us to detect generalities, but there is no guaranty that synthetic network properties do not arise from the wrong reason. Another major problem that may preclude disentangling trait-based processes is that traits could influence interactions directly via trait matching, or indirectly via environmental matching. Hence, even if the variance between neutral and trait-matching components is successfully partitioned, this would ignore the fact that some of the 'neutral' variance was generated by species traits via their effect on distribution and abundance (as we outlined in the previous section). Thus, the influence of abundance vs. traits can be seen as a path diagram where traits directly affect interactions and also affect abundances, which affect interactions (Fig. 1). We propose a framework that aims to integrate, rather than separate both processes.

A significant challenge before such an analysis can be achieved is to access completely sampled networks with which to validate models. Empirical network data, however, have inherent uncertainties associated with the way in which they are sampled. Specifically, sampling completeness is rarely achieved when collecting interaction networks (Chacoff *et al.* 2012; Bartomeus 2013), and hence, some unobserved interactions may indeed occur (i.e. false absence of interactions). This would be less of a problem if the proportion of interactions that are sampled were constant, but this sampling efficiency can vary with local environmental conditions (Laliberté & Tylianakis 2010), species abundance and frequency, and of course, sampling effort. Thus, to truly understand the importance of trait matching for determining species interactions, the absence of an interaction in an empirical data set cannot be used to infer true absence of that interaction in nature. The nature of the data therefore impedes the direct evaluation of probabilistic models (e.g. Rohr *et al.* 2010; Crea, Ali & Rader 2016) and requires methods to estimate absences (Bartomeus 2013) or the development of model-fitting procedures based on observed interactions only.

Another challenge is that null models based on *a priori* rules for interactions have to be constructed using assumptions of which traits are critical for interaction establish-

ment. Constructing and interpreting biologically meaningful null models that can isolate the targeted process to be studied is not an easy task (Vázquez & Aizen 2003). As an alternative, recent attempts to understand trait matching by statistically modelling empirical data are promising (e.g. models incorporating imperfect detectability: Bartomeus 2013; fourth corner analysis: Dehling *et al.* 2014; linear models: González-Castro *et al.* 2015; Dirichlet-multinomial regression: Crea, Ali & Rader 2016), but such models are still unable to integrate the relative contribution of neutral vs. trait-based process.

A final caveat is that most models are constrained to use mean trait values at the species level, neglecting variability among individuals of the same species. However, intraspecific trait variation, which can result from life-history stage, sexual dimorphism, or stochastic, environmental, genetic or epigenetic forces (Bolnick *et al.* 2011), has been shown to affect specific interactions such as competition, as well as overall ecological dynamics (González-Suárez & Revilla 2013).

### A probabilistic method for evaluating trait matching

To overcome the limitations pointed out above, we model the probability of interaction among pairs of individuals given their traits, based on a framework developed by Gravel and colleagues (Gravel *et al.* 2013). The method also has the advantage to build directly on the established theory of ecological network structure (Williams & Martinez 2000; Eklöf *et al.* 2013), by contrast with the above-listed methods that are essentially phenomenological. We propose a method to evaluate trait-matching relationships while taking into account the abundance of the interacting partners. The fitting procedure uses information about observed interactions only, thereby overcoming problems caused by undersampling of rare interactions leading to false absences of interactions. The approach implies that sampling effort is enough to adequately describe most true interactions in trait-space and no false positives are recorded (i.e. recording interactions as true when they do not occur). A previous sensitivity analysis, however, revealed it to be robust to sampling effort (Gravel *et al.* 2013). The parameters are estimated by maximum likelihood, and the fitted model can be used to predict unobserved interactions based on species traits and abundances. Several models, corresponding to different hypotheses, can fit directly to raw data and accommodate complex trait-matching response functions to either qualitative or quantitative interaction data. Finally, they can incorporate intraspecific trait variation, avoiding the loss of realism in species with trait values that vary along developmental stages or environmental gradients. In that way, we provide a common toolbox to understand trait-matching rules across a variety of interaction types.

We are interested in evaluating from empirical data a function describing the probability of an interaction



between species  $i$  and  $j$  based on their respective sets of traits  $T_i$  and  $T_j$ . Building upon the model developed by Gravel *et al.* (2013), we aim to evaluate the parameters of a model that will relate the probability with which an interaction occurs to the set of traits of the two species:

$$P(L_{ij} = 1|T_i, T_j) \quad \text{eqn 1}$$

which reads as the probability of observing an interaction  $L$  between species  $i$  and  $j$  given the traits  $T_i$  and  $T_j$ . The function describing this probability could take any form. For the sake of the example here, we will consider a Gaussian shaped function (i.e. a function that assumes an unimodal relationship between  $T_i$  and  $T_j$ ) to represent the trait-matching interaction (also termed interaction niche; Williams, Anandanadesan & Purves 2010; see below). Other functions, such as a high-order polynomial or even regression trees, could be considered as well. The Gaussian function is, however, convenient because it is easy to integrate and further it matches the niche model of network structure (Williams, Anandanadesan & Purves 2010; Eklöf *et al.* 2013).

Equation 1 could be fitted directly to empirical data by maximum likelihood. To do so, the required data should contain information on the presence and absence of interactions (e.g. Rohr *et al.* 2010). The problem we are facing, however, is that records of the true absence of interactions are often not available in most data sets of ecological interactions, and when available, there might be considerable uncertainty in these absences (i.e. false negatives due to insufficient sampling). We therefore derive a likelihood function using Bayes theorem to fit eqn 1 indirectly, using only information about the observed interactions. Parameters are still evaluated by maximum likelihood (using simulated annealing, as described in the Supplementary Information), but one could eventually develop the method further to compute the posterior distribution of parameters.

The data contain information about the traits of species  $i$  and of species  $j$  only for observed interactions,  $L_{ij} = 1$ . We consequently revise the problem and model the probability of observing trait  $T_i$ , knowing the trait  $T_j$  and the occurrence of the interaction  $L_{ij}$ :

$$P(T_i|L_{ij} = 1, T_j) \quad \text{eqn 2}$$

which could be interpreted as the probability that we pick trait  $T_i$  from the trait distribution we model, given we know there is an interaction between species  $i$  and  $j$  and the trait  $T_j$ . This equation provides the likelihood for any observation of an interaction based on the traits of the two species. We now use Bayes' theorem,  $p(A|B)p(B) = p(B|A)p(A)$ , to decompose eqn 2, yielding the following distribution of the trait of one species, given the trait of the second species and the observation of the interaction:

$$P(T_i|L_{ij} = 1, T_j) = \frac{P(L_{ij} = 1|T_i, T_j) P(T_i)}{P(L_{ij} = 1|T_j)} \quad \text{eqn 3}$$

The first term from the numerator is the trait-matching model, described in eqn 1. It is the model for which we aim to evaluate parameters.  $P(T_i)$  is the probability density

function for the trait  $T_i$ . It corresponds to the probability of observing this trait in the regional pool. It could be weighted by abundance because the most abundant species are more likely to be sampled. The denominator is the marginal distribution of the trait-matching function, computed as the integral of the numerator over the whole distribution of the trait  $T_i$ :

$$P(L_{ij} = 1|T_j) = \int_{-\infty}^{\infty} P(L_{ij} = 1|T_i, T_j) * P(T_i) dT_i \quad \text{eqn 4}$$

As a side product, the denominator informs us of the generality of the species  $j$ . This integral might be tricky to compute analytically, depending on the form of eqn 1 and the distribution of trait  $T_i$ , but most software offer easy ways to compute it numerically.

The model given at eqn 3 should not be confounded with the more traditional use of the Bayes theorem in statistics. The resulting distribution describes the probability of observing an interaction given a trait, while in statistics, the distribution describes the probability of observing a set of parameters given the data. Here, the parameters are estimated by simulated annealing and there is only a single set of parameters yielding the maximum likelihood. True confidence intervals for parameter estimates are hard to evaluate for nonlinear models with complex likelihood surfaces, but could nonetheless be evaluated numerically. The eqn 3 could also be implemented in a Bayesian fitting procedure to obtain a posterior distribution of parameters for eqn 1 (eqn 3 being the likelihood of the Bayes theorem), but this would be out of the scope of the current study.

The model could be simplified to account only for the effect of abundance (trait distributions) to reveal the importance of the trait-matching constraint. A neutral model in this framework is found when an interaction is equally probable, irrespective of the traits of the two species involved in the interaction (i.e. eqn 1 is set as a constant). Alternatively, one could want to compare to the situation where interactions are purely determined by trait-matching constraints. In this situation, we consider the distribution of the trait  $P(T_i)$  uniform within the range of the observed traits. The eqn 3 remains the same for all three models and could be used to compute the likelihood for each of them. Equations for the pure neutral and trait-matching models, and a multitrait expansion, as well as all of the R code necessary to perform this analysis are provided in the supplementary material (see Appendix S1 in Supporting Information) and as an R package found at [https://github.com/ibartomeus/trait\\_match](https://github.com/ibartomeus/trait_match).

We re-analysed three data sets on different systems ranging from antagonistic to mutualistic interactions to illustrate the overall principle of the method. First, we use data from Barnes *et al.* (2008) on the diet of marine fish species. The traits are the individual (log transformed) body size of the predator fish species ( $M_{\text{pred}}$ ) and the individual body size of preys ( $M_{\text{prey}}$ ). We know that larger fish typically feed on smaller ones because they must catch and handle

the prey with their mouth. The frequency distribution of prey size will indeed influence the distribution of the body mass in the diet of the predator. A predator will tend to feed most often on the most abundant preys, which is a neutral component to the interaction probability. The predator does not select from that distribution randomly, however, but rather it targets only a specific range (given by eqn 1; the niche component). Both the available prey size distribution,  $P(M_{\text{prey}})$ , and the resulting prey size distribution,  $P(M_{\text{prey}}|L, M_{\text{pred}})$ , are illustrated in Fig. 2b for a given predator species. The resulting prey distribution has to be somewhere between the regional prey distribution and its preferred prey size. The model therefore integrates both neutral and trait-matching constraints.

We consider the following Gaussian function to represent the probability of an interaction given the size of the predator and the prey:

$$P(L_{ij} = 1 | M_{\text{pred}}, M_{\text{prey}}) = \exp \frac{-(\alpha_0 + \alpha_1 * M_{\text{pred}} - M_{\text{prey}})^2}{2(\beta_0 + \beta_1 * M_{\text{pred}})^2} \quad \text{eqn 5}$$

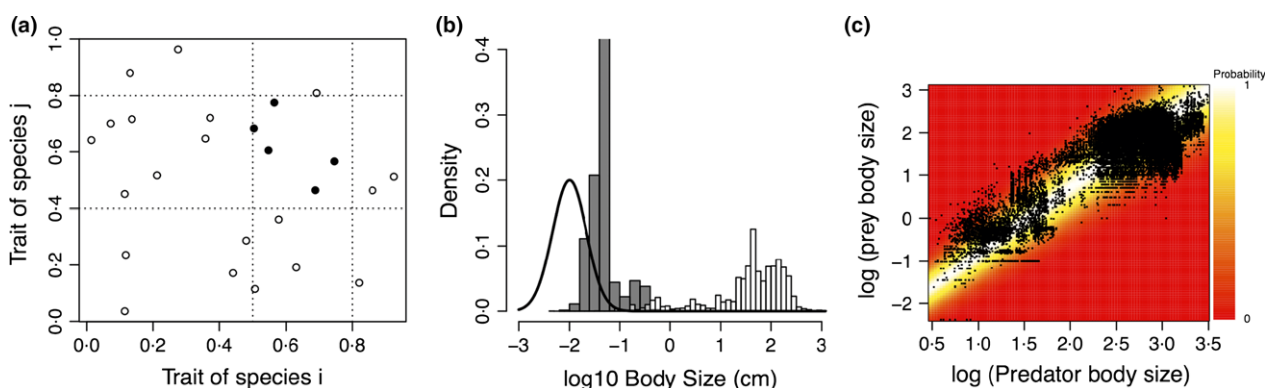
where  $\alpha_0$ ,  $\alpha_1$ ,  $\beta_0$  and  $\beta_1$  are fitted parameters describing the linear relationship between the predator size, its optimum ( $\alpha_0 + \alpha_1 M_{\text{pred}}$ ) and the range ( $\beta_0 + \beta_1 M_{\text{pred}}$ ) of its preference function. This formulation considers there is an optimal prey size for the predator, and the probability an interaction occurs reduces with any deviation from it (Williams, Anandanadesan & Purves 2010). The optimum also increases linearly with predator size. The same reasoning could also be applied to mutualistic interactions, considering there is an optimal corolla length for a pollinator of a given tongue length. One tricky issue might be to gather information about the prey trait distribution. The distribution of prey traits might be influenced by the interactions if there is a feedback of predators on prey abundance, and in the best situation, we need to tease that effect apart.

Here, we assume that the distribution of the data provides an adequate representation of the distribution of potential prey sizes because of the large number of observed interactions (>33 000) and their diversity. We thus consider a normal distribution of (log) prey size and computed the average and the standard deviation.

The predator–prey example provides a case where trait matching is a strong driver of interactions because of a strong predator–prey body size relationship (likelihood = −21223). The parameters of the fitted model can subsequently be used for predicting interactions among species that co-occur, but have not been observed to interact (e.g. due to incomplete sampling) or more interestingly, for species that currently do not co-occur but may do so in the future, for example as a consequence of range shifts under climate change (Albouy *et al.* 2014) or species invasions.

Next, we use the same models on experimental data on the relationship between grasshopper incisive strength and leaf dry matter content (Deraison *et al.* 2015). In this case, both traits are species averages. We first find weak trait matching for binary data (who eats whom at the species level; likelihood = −213; Fig. 3a). However, weighting the interactions by consumption frequency removes bias in parameter estimates and the fit of the model is considerably improved (likelihood = −5383). We thus find that strong-mandibled grasshoppers prefer plants with higher content of dry matter, as reported in the original paper (Fig. 3b).

The model could also be evaluated using traits measured at the individual level. In the last example, we related pollinator tongue length with plant nectar holder depth in visitation networks from Bartomeus, Vilà & Santamaría (2008). Individual pollinator tongue length was inferred using the allometric relationship with body intertegular span within each bee family (Cariveau *et al.* 2016), while

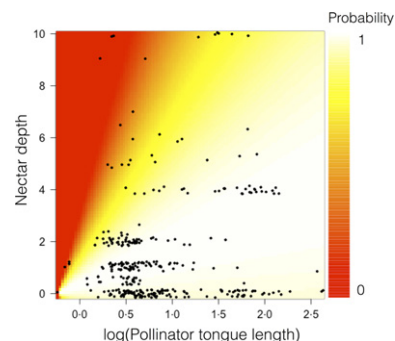


**Fig. 2.** Illustration of the quantitative framework to evaluate a trait-matching probabilistic function. (a) Conceptual representation of a trait-matching constraint. Interactions (in black) are feasible only when both species have traits that are compatible. However, we often do not have reliable information on the species that are present, but are not observed to interact (white dots). Dotted lines indicate the trait ranges of compatibility between the species (b) Representation of the density function for available body size in the (Barnes *et al.* 2008) data set (white bars), the trait-matching function (black line) and the observed distribution of prey size for the predator *Nototheniops larseni* (black bars). (c) Representation of the observed interactions (black dots) and the prediction for the maximum-likelihood estimate of the trait-matching function (from low probability in red to high probability in white).

species average flower size was considered for plants. Individual trait data for pollinators allow capture of the interindividual differences when evaluating parameters of trait-matching functions. In addition, this model uses independent information to describe the trait abundance distribution of plant species. In the past examples, abundance was inferred from the network of interactions, but in this case, independent transect measures of per cent plant cover in the site are available (Bartomeus, Vilà & Santamaría 2008). We find that the model can be interpreted as a trait-barrier, where small-tongued individuals cannot access deep flowers, but long-tongued species can access both deep and shallow flowers (likelihood =  $-705$ ; Fig. 4). However, under such weak constraints (most pollinators can access most plants), abundance is the main determinant of interaction probability. For comparison, using pollinator species trait averages instead of individual values produce a similar model, but with a worst likelihood ( $-726$ ), indicating that there is a gain from using detailed data when available.

## Discussion and conclusions

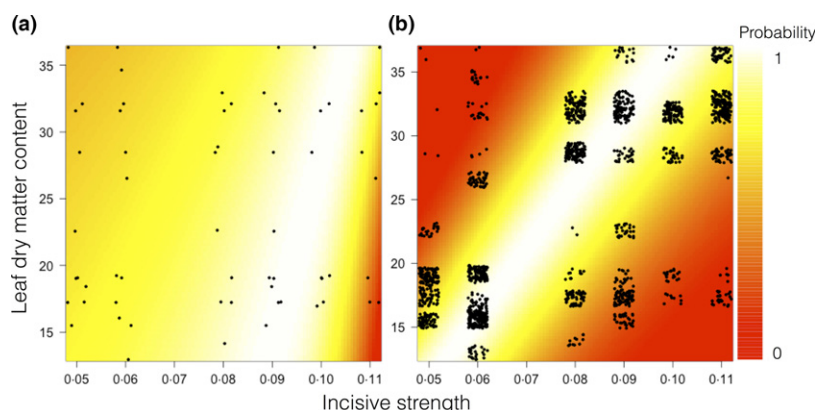
Quantifying the trait-matching relationships across species may help us to understand how networks are structured. For example, the nested structure of plant–pollinator networks may be driven from species abundance (Vázquez, Chacoff & Cagnolo 2009) or from barriers to certain interactions (Stang, Klinkhamer & van der Meijden 2006). In contrast, the strong trait matching observed in plant–herbivore interactions (e.g. plant defences limiting herbivory for all but a few tolerant species) can produce more modular networks where interactions depart more from the null expectation based solely on abundance (Thébault & Fontaine 2010). Even within plant–pollinator interactions, bird–plant networks are more specialized than insect–plant networks, which is also reflected in their degree of trait



**Fig. 4.** Representation of the fitted interaction probability plant and pollinators weighted by the frequency of interactions (from low probability in red to high probability in white). Only a few interactions among small-tongue-sized bees and long corolla depth flowers are realized (red area), while the rest of interactions are explained mainly by abundance.

matching (Maglianesi, Böhning-Gaese & Schleuning 2015). Our framework is, however, limited to pairwise interactions, and future work will have to investigate how the distribution of traits in a community constrains the emergent network properties. Moreover, trait-matching constraints describe potential interactions, but may not always reflect realized interactions (Poisot, Stouffer & Gravel 2015). The future development of a Bayesian approach to evaluate the distribution of parameters will help quantifying the uncertainty of predicted interactions.

Parameterized trait-matching functions not only provide a better understanding of the drivers of interactions, but they also allow prediction of novel interactions following deliberate introductions (e.g. of crop species or biological control agents) or unintentional invasions and range shifts (Morales-Castilla *et al.* 2015). Proxies of trait similarity, such as phylogenetic distance, have already been successfully used to predict interactions of exotic species (Pearse & Altermatt 2013), and adding traits has the



**Fig. 3.** Representation of the fitted interaction probability for grasshopper and plant interactions unweighted (a) and weighted (b) by frequency of interaction (from low probability in red to high probability in white). The probability of interaction between a grasshopper and a plant follows a positive relationship between incisive strength and plant leaf dry matter content. Note that the overlapping data in b have been jittered to appreciate the different frequencies of particular interactions. The likelihood for (a) is similar to the neutral model, while much better in (b), indicating that the frequency of interactions must be taken into account to better reveal the trait-matching constraint.

potential to enhance this approach. Species losses and gains following local and global changes are threatening most ecosystems, and it is simply impossible to measure all potential interactions in the field. Tools are consequently required to assess how the interaction network will rewire. We know that exotic species invading a community get easily integrated into the recipient network of interactions (Albrecht *et al.* 2014) and that after species turnover in a community, the remaining species reshuffle their interactions to adjust to the new composition (Kaiser-Bunbury *et al.* 2010). Our predictive ability in these situations is, however, still limited.

Careful selection of the right set of traits to run the analysis is, however, a critical step. We have seen that traits constraining interactions could potentially comprise all morphological and physiological species characteristics, and hence, are quite specific for each interaction type. A good *a priori* knowledge on the biology of the species and type of interaction involved is needed to select the right trait combinations. For example, we also explored whether body size drives host–parasite relationships using the Tylianakis, Tscharntke & Lewis (2007) data set, but in this case, all models performed poorly because the largest parasitoid is smaller than the smallest host, which allows all types of body size combinations. Alternatively, spurious trait matches could be found when some traits are correlated. For instance, traits such as body size correlate allometrically with several other morphological traits (Woodward *et al.* 2005) and might therefore provide a wrong causal explanation of the interactions. One strong limitation for some interactions, such as fungi and plants, is that the traits governing interactions remain somewhat unclear (Tedersoo *et al.* 2008; Martínez-García *et al.* 2015). The challenge for the future will be to determine and quantify the actual traits governing these interactions, including their variability among individuals or genets.

Another challenge outlined in Fig. 1 and still unresolved is inferring functioning from a network of interactions (Duffy *et al.* 2009; Thompson *et al.* 2012). Species interactions are driving several ecosystem processes and functions (e.g. animal pollination, fruit dispersion) as well as energy fluxes (e.g. predation, parasitism). Inferring the function from traits, however, requires incorporating the interaction efficiency (the per capita strength of a single interaction link; Vázquez *et al.* 2015), which in turn may be also trait- or abundance-mediated, and can depend on the extent of matching (e.g. pollinators with short tongues may be able to visit, but inefficiently pollinate long flower corollas), or on morphological, physiological or behavioural traits (e.g. large pollinators deposit more pollen; Hoehn *et al.* 2008; Fig. 1). Empirical evidence measuring interaction efficiency is still scarce.

In conclusion, different traits can inform us about how species form networks of interactions. For some interaction types, like mycorrhizal fungal interactions, traits affecting co-occurrence can be the most relevant for under-

standing the occurrence of interactions. Conversely, for other interaction types, like those between predators and prey, morphological and physiological traits may be the main determinants of who interacts with whom. Understanding which mechanisms are driving pairwise interactions is a key to predict how communities will respond to global change. Interactions regulated by co-occurrence will be more likely to be affected by climate change (e.g. changing phenologies and distributions), while changes in dominance following disturbance may redistribute the interactions in neutral-driven networks. Non-random species extinctions are also expected to affect more drastically interactions regulated by strong trait matching (Larsen, Williams & Kremen 2005). There are still too many unknowns to draw general conclusions about how communities are structured by traits and what implications this has for ecosystem functioning, but we are now armed with appropriate analytical tools to move beyond the mere description of interactions and run predictive analysis of network assembly and dynamics.

## Acknowledgements

We thank C. Barnes and H. Deraison for making their data available. We also thank T. Poisot, D. Stouffer and S. Kefi for organizing this special issue and three anonymous reviewers for appreciated and constructive comments. IB is funded by project SURVIVE\_HIREC (CGL2013-47448-P).

## Data accessibility

All data and R code to perform analysis can be found in the Supporting Information and at [https://github.com/ibartomeus/trait\\_match](https://github.com/ibartomeus/trait_match) (10.5281/zenodo.46382).

## References

- Albouy, C., Velez, L., Coll, M., Colloca, F., Loc'h, F., Mouillot, D. *et al.* (2014) From projected species distribution to food-web structure under climate change. *Global Change Biology*, **20**, 730–741.
- Albrecht, M., Padrón, B., Bartomeus, I. & Traveset, A. (2014) Consequences of plant invasions on compartmentalization and species' roles in plant–pollinator networks. *Proceedings of the Royal Society of London B: Biological Sciences*, **281**, 20140773.
- Barnes, C., Bethea, D.M., Brodeur, R.D., Spitz, J., Ridoux, V., Pusineri, C. *et al.* (2008) Predator and prey body sizes in marine food webs: ecological Archives E089-051. *Ecology*, **89**, 881.
- Bartomeus, I. (2013) Understanding linkage rules in plant–pollinator networks by using hierarchical models that incorporate pollinator detectability and plant traits. *PLoS ONE*, **8**, e69200.
- Bartomeus, I., Vilà, M. & Santamaría, L. (2008) Contrasting effects of invasive plants in plant–pollinator networks. *Oecologia*, **155**, 761–770.
- Bartomeus, I., Park, M.G., Gibbs, J., Danforth, B.N., Lakso, A.N. & Winfree, R. (2013) Biodiversity ensures plant–pollinator phenological synchrony against climate change. *Ecology Letters*, **16**, 1331–1338.
- Bascompte, J. & Jordano, P. (2007) Plant–Animal Mutualistic Networks: the architecture of biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, **38**, 567–593.
- Bolnick, D.I., Amarasekare, P., Araújo, M.S., Bürger, R., Levine, J.M., Novak, M. *et al.* (2011) Why intraspecific trait variation matters in community ecology. *Trends in Ecology & Evolution*, **26**, 183–192.
- Canard, E.F., Mouquet, N., Mouillot, D., Stanko, M., Miklisova, D. & Gravel, D. (2014) Empirical evaluation of neutral interactions in host–parasite networks. *The American Naturalist*, **183**, 468–479.
- Cariveau, D.P., Nayak, G., Bartomeus, I., Zientek, J., Ascher, J. & Winfree, R. (2016) The allometry of bee tongue length and its uses in ecology and evolution. *PLoS ONE*, **11**, e0151482.



- Carvalho, L.G., Barbosa, E.R.M. & Memmott, J. (2008) Pollinator networks, alien species and the conservation of rare plants: *Trinia glauca* as a case study. *The Journal of Applied Ecology*, **45**, 1419–1427.
- Castellanos, M.C., Wilson, P. & Thomson, J.D. (2003) Pollen transfer by hummingbirds and bumblebees, and the divergence of pollination modes in *Penstemon*. *Evolution*, **57**, 2742–2752.
- Chacoff, N.P., Vázquez, D.P., Lomáscolo, S.B., Stevani, E.L., Dorado, J. & Padrón, B. (2012) Evaluating sampling completeness in a desert plant–pollinator network. *The Journal of Animal Ecology*, **81**, 190–200.
- Chagnon, P.-L., Bradley, R.L., Maherali, H. & Klironomos, J.N. (2013) A trait-based framework to understand life history of mycorrhizal fungi. *Trends in Plant Science*, **18**, 484–491.
- Crea, C., Ali, R.A. & Rader, R. (2016) A new model for ecological networks using species-level traits. *Methods in Ecology and Evolution*, **7**, 232–241.
- Cunha, I. & Planas, M. (1999) Optimal prey size for early turbot larvae (*Scophthalmus maximus* L.) based on mouth and ingested prey size. *Aquaculture*, **175**, 103–110.
- D'Amen, M., Dubuis, A., Fernandes, R.F., Pottier, J., Pellissier, L. & Guisan, A. (2015) Using species richness and functional traits predictions to constrain assemblage predictions from stacked species distribution models. *Journal of Biogeography*, **42**, 1255–1266.
- Dehling, D.M., Töpfer, T., Schaefer, H.M., Jordano, P., Böhning-Gaese, K. & Schleuning, M. (2014) Functional relationships beyond species richness patterns: trait matching in plant–bird mutualisms across scales. *Global Ecology and Biogeography*, **23**, 1085–1093.
- Deraison, H., Badenhausser, I., Börger, L. & Gross, N. (2015) Herbivore effect traits and their impact on plant community biomass: an experimental test using grasshoppers. *Functional Ecology*, **29**, 650–661.
- Díaz, S. & Cabido, M. (2001) Vive la différence: plant functional diversity matters to ecosystem processes. *Trends in Ecology & Evolution*, **16**, 646–655.
- Duffy, J.E., Srivastava, D.S., McLaren, J., Sankaran, M., Solan, M., Griffin, J. et al. (2009) Forecasting decline in ecosystem services under realistic scenarios of extinction. *Biodiversity, Ecosystem Functioning and Human Wellbeing: An Ecological and Economic Perspective* (eds S. Naeem, D.E. Bunker, A. Hector, M. Loreau & C. Perrings), pp. 60–77. Oxford University Press, Oxford, UK.
- Eklöf, A., Jacob, U., Kopp, J., Bosch, J., Castro-Urgal, R., Chacoff, N.P. et al. (2013) The dimensionality of ecological networks. *Ecology Letters*, **16**, 577–583.
- Encinas-Viso, F., Revilla, T.A. & Etienne, R.S. (2012) Phenology drives mutualistic network structure and diversity. *Ecology Letters*, **15**, 198–208.
- Fenster, C.B., Reynolds, R.J., Williams, C.W., Makowsky, R. & Dudash, M.R. (2015) Quantifying hummingbird preference for floral trait combinations: the role of selection on trait interactions in the evolution of pollination syndromes. *Evolution*, **69**, 1113–1127.
- Galetti, M., Guevara, R., Cortés, M.C., Fadini, R., Von Matter, S., Leite, A.B. et al. (2013) Functional extinction of birds drives rapid evolutionary changes in seed size. *Science*, **340**, 1086–1090.
- González-Castro, A., Yang, S., Nogales, M. & Carlo, T.A. (2015) Relative importance of phenotypic trait matching and species' abundances in determining plant–avian seed dispersal interactions in a small insular community. *AOB Plants*, **7**, plv017.
- González-Suárez, M. & Revilla, E. (2013) Variability in life-history and ecological traits is a buffer against extinction in mammals. *Ecology Letters*, **16**, 242–251.
- Gravel, D., Poisot, T., Albouy, C., Velez, L. & Mouillot, D. (2013) Inferring food web structure from predator–prey body size relationships. *Methods in Ecology and Evolution*, **4**, 1083–1090.
- Gross, N., Kunstler, G., Liancourt, P., De Bello, F., Suding, K.N. & Lavorel, S. (2009) Linking individual response to biotic interactions with community structure: a trait-based framework. *Functional Ecology*, **23**, 1167–1178.
- Herrera, C.M. (2000) Flower-to-seedling consequences of different pollination regimes in an insect-pollinated shrub. *Ecology*, **81**, 15–29.
- Hoehn, P., Tschamtker, T., Tylianakis, J.M. & Steffan-Dewenter, I. (2008) Functional group diversity of bee pollinators increases crop yield. *Proceedings of the Royal Society of London B: Biological Sciences*, **275**, 2283–2291.
- Jacquemyn, H., Merckx, V., Brys, R., Tyteca, D., Cammue, B.P.A., Honnay, O. et al. (2011) Analysis of network architecture reveals phylogenetic constraints on mycorrhizal specificity in the genus *Orchis* (Orchidaceae). *The New Phytologist*, **192**, 518–528.
- Kaiser-Bunbury, C.N., Muff, S., Memmott, J., Müller, C.B. & Calfisch, A. (2010) The robustness of pollination networks to the loss of species and interactions: a quantitative approach incorporating pollinator behaviour. *Ecology Letters*, **13**, 442–452.
- Krishna, A., Guimaraes, P.R. Jr, Jordano, P. & Bascompte, J. (2008) A neutral-niche theory of nestedness in mutualistic networks. *Oikos*, **117**, 1609–1618.
- Kritsky, G. (1991) Darwin's Madagascan Hawk Moth prediction. *American Entomologist*, **37**, 206–210.
- Liberté, E. & Tylianakis, J.M. (2010) Deforestation homogenizes tropical parasitoid–host networks. *Ecology*, **91**, 1740–1747.
- Larsen, T.H., Williams, N.M. & Kremen, C. (2005) Extinction order and altered community structure rapidly disrupt ecosystem functioning. *Ecology Letters*, **8**, 538–547.
- Laughlin, D.C., Joshi, C., van Bodegom, P.M., Bastow, Z.A. & Fulé, P.Z. (2012) A predictive model of community assembly that incorporates intraspecific trait variation. *Ecology Letters*, **15**, 1291–1299.
- Lima-Mendez, G., Faust, K., Henry, N., Decelle, J. & Colin, S. (2015) Determinants of community structure in the global plankton interactome. *Science*, **348**, doi: 10.1126/science.1262073.
- Maglianesi, M.A., Böhning-Gaese, K. & Schleuning, M. (2015) Different foraging preferences of hummingbirds on artificial and natural flowers reveal mechanisms structuring plant–pollinator interactions. *The Journal of Animal Ecology*, **84**, 655–664.
- Martínez-García, L.B., Richardson, S.J., Tylianakis, J.M., Peltzer, D.A. & Dickie, I.A. (2015) Host identity is a dominant driver of mycorrhizal fungal community composition during ecosystem development. *The New Phytologist*, **205**, 1565–1576.
- Morales-Castilla, I., Matias, M.G., Gravel, D. & Araújo, M.B. (2015) Inferring biotic interactions from proxies. *Trends in Ecology & Evolution*, **30**, 347–356.
- Olesen, J.M., Bascompte, J., Dupont, Y.L., Elberling, H., Rasmussen, C. & Jordano, P. (2011) Missing and forbidden links in mutualistic networks. *Proceedings of the Royal Society of London B: Biological Sciences*, **278**, 725–732.
- Olito, C. & Fox, J.W. (2015) Species traits and abundances predict metrics of plant–pollinator network structure, but not pairwise interactions. *Oikos*, **124**, 428–436.
- Pearse, I.S. & Altermatt, F. (2013) Predicting novel trophic interactions in a non-native world. *Ecology Letters*, **16**, 1088–1094.
- Peralta, G., Frost, C.M., Rand, T.A., Didham, R.K. & Tylianakis, J.M. (2014) Complementarity and redundancy of interactions enhance attack rates and spatial stability in host–parasitoid food webs. *Ecology*, **95**, 1888–1896.
- Poisot, T., Stouffer, D.B. & Gravel, D. (2015) Beyond species: why ecological interaction networks vary through space and time. *Oikos*, **124**, 243–251.
- Poisot, T., Canard, E., Mouillot, D., Mouquet, N., Gravel, D. & Jordan, F. (2012) The dissimilarity of species interaction networks. *Ecology Letters*, **15**, 1353–1361.
- Rader, R., Reilly, J., Bartomeus, I. & Winfree, R. (2013) Native bees buffer the negative impact of climate warming on honey bee pollination of watermelon crops. *Global Change Biology*, **19**, 3103–3110.
- Rohr, R.P., Scherer, H., Kehrli, P., Mazza, C. & Bersier, L.-F. (2010) Modeling food webs: exploring unexplained structure using latent traits. *The American Naturalist*, **176**, 170–177.
- Sargent, R.D. & Ackerly, D.D. (2008) Plant–pollinator interactions and the assembly of plant communities. *Trends in Ecology and Evolution*, **23**, 123–130.
- Shipley, B., Vile, D. & Garnier, E. (2006) From plant traits to plant communities: a statistical mechanistic approach to biodiversity. *Science*, **314**, 812–814.
- Spitz, J., Ridoux, V. & Brind'Amour, A. (2014) Let's go beyond taxonomy in diet description: testing a trait-based approach to prey–predator relationships. *The Journal of Animal Ecology*, **83**, 1137–1148.
- Stang, M., Klinkhamer, P.G.L. & van der Meijden, E. (2006) Asymmetric specialization and extinction risk in plant–flower visitor webs: a matter of morphology or abundance? *Oecologia*, **151**, 442–453.
- Tedersoo, L., Jairus, T., Horton, B.M., Abarenkov, K., Suvi, T., Saar, I. et al. (2008) Strong host preference of ectomycorrhizal fungi in a Tasmanian wet sclerophyll forest as revealed by DNA barcoding and taxon-specific primers. *The New Phytologist*, **180**, 479–490.
- Thébault, E. & Fontaine, C. (2010) Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science*, **329**, 853–856.

- Thompson, R.M., Brose, U., Dunne, J.A., Hall, R.O. Jr, Hladyz, S., Kitching, R.L. *et al.* (2012) Food webs: reconciling the structure and function of biodiversity. *Trends in Ecology & Evolution*, **27**, 689–697.
- Tylianakis, J.M., Tschamntke, T. & Lewis, O.T. (2007) Habitat modification alters the structure of tropical host–parasitoid food webs. *Nature*, **445**, 202–205.
- Tylianakis, J.M., Didham, R.K., Bascompte, J. & Wardle, D.A. (2008) Global change and species interactions in terrestrial ecosystems. *Ecology Letters*, **11**, 1351–1363.
- Tylianakis, J.M., Laliberté, E., Nielsen, A. & Bascompte, J. (2010) Conservation of species interaction networks. *Biological Conservation*, **143**, 2270–2279.
- Vázquez, D.P. & Aizen, M.A. (2003) Null model analyses of specialization in plant–pollinator interactions. *Ecology*, **84**, 2493–2501.
- Vázquez, D.P., Chacoff, N.P. & Cagnolo, L. (2009) Evaluating multiple determinants of the structure of plant–animal mutualistic networks. *Ecology*, **90**, 2039–2046.
- Vázquez, D.P., Melián, C.J., Williams, N.M., Blüthgen, N., Krasnov, B.R. & Poulin, R. (2007) Species abundance and asymmetric interaction strength in ecological networks. *Oikos*, **116**, 1120–1127.
- Vázquez, D.P., Ramos-Jiliberto, R., Urbani, P. & Valdovinos, F.S. (2015) A conceptual framework for studying the strength of plant–animal mutualistic interactions. *Ecology Letters*, **18**, 385–400.
- Waser, N.M., Chittka, L., Price, M.V., Williams, N.M. & Ollerton, J. (1996) Generalization in pollination systems, and why it matters. *Ecology*, **77**, 1043–1060.
- White, E.P., Ernest, S.K.M., Kerkhoff, A.J. & Enquist, B.J. (2007) Relationships between body size and abundance in ecology. *Trends in Ecology & Evolution*, **22**, 323–330.
- Williams, R.J., Anandanadesan, A. & Purves, D. (2010) The probabilistic niche model reveals the niche structure and role of body size in a complex food web. *PLoS ONE*, **5**, e12092.
- Williams, R.J. & Martinez, N.D. (2000) Simple rules yield complex food webs. *Nature*, **404**, 180–183.
- Woodward, G., Ebenman, B., Emmerson, M., Montoya, J.M., Olesen, J.M., Valido, A. *et al.* (2005) Body size in ecological networks. *Trends in Ecology & Evolution*, **20**, 402–409.
- Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F. *et al.* (2004) The worldwide leaf economics spectrum. *Nature*, **428**, 821–827.
- Zobel, M. & Öpik, M. (2014) Plant and arbuscular mycorrhizal fungal (AMF) communities—which drives which? *Journal of Vegetation Science*, **25**, 1133–1140.

Received 25 August 2015; accepted 10 March 2016

Handling Editor: Timothée Poisot

## Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

**Appendix S1.** Niche, neutral and multitrait derivations.

**Appendix S2.** R code as zip.