OIKOS

Research

Species traits as drivers of food web structure

Idaline Laigle, Isabelle Aubin, Christoph Digel, Ulrich Brose, Isabelle Boulangeat and Dominique Gravel

I. Laigle (http://orcid.org/0000-0003-1422-9984) (idaline.laigle@usherbrooke.ca) and D. Gravel (http://orcid.org/0000-0002-4498-7076): Dépt de biologie, Univ. de Sherbrooke, 2500 Boulevard l'Université, Sherbrooke, QC, J1K 2R1, Canada, and: Québec Centre for Biodiversity Science. – I. Aubin, Natural Resources Canada, Canadian Forest Service, Great Lakes Forestry Centre, Sault Ste. Marie, ON, Canada. – C. Digel, J.F. Blumenbach Inst. für Zoologie und Anthropologie. Georg-August-Univ. Göttingen, Göttingen, Germany. – U. Brose (http://orcid.org/0000-0001-9156-583X), German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany. – I. Boulangeat (http://orcid.org/0000-0002-8463-7046), Dept of Bioscience - Ecoinformatics and Biodiversity, Aarhus Univ., Aarhus, Denmark.

The use of functional traits to describe community structure is a promising approach to reveal generalities across organisms and ecosystems. Plant ecologists have demonstrated the importance of traits in explaining community structure, competitive interactions as well as ecosystem functioning. The application of trait-based methods to more complex communities such as food webs is however more challenging owing to the diversity of animal characteristics and of interactions. The objective of this study was to determine how functional structure is related to food web structure. We consider that food web structure is the result of 1) the match between consumer and resource traits, which determine the occurence of a trophic interaction between them, and 2) the distribution of functional traits in the community. We implemented a statistical approach to assess whether or not 35 466 pairwise interactions between soil organisms are constrained by trait-matching and then used a Procrustes analysis to investigate correlations between functional indices and network properties across 48 sites. We found that the occurrence of trophic interactions is well predicted by matching the traits of the resource with those of the consumer. Taxonomy and body mass of both species were the most important traits for the determination of an interaction. As a consequence, functional evenness and the variance of certain traits in the community were correlated to trophic complementarity between species, while trait identity, more than diversity, was related to network topology. The analysis was however limited by trait data availability, and a coarse resolution of certain taxonomic groups in our dataset. These limitations explain the importance of taxonomy, as well as the complexity of the statistical model needed. Our results outline the important implications of trait composition on ecological networks, opening promising avenues of research into the relationship between functional diversity and ecosystem functioning in multi-trophic systems.

Introduction

The use of functional traits to describe community structure is a promising approach to reveal generalities across organisms and ecosystems (McGill et al. 2006). Plant ecologists have demonstrated the importance of traits in explaining community

NORDIC SOCIETY OIKOS

www.oikosjournal.org

structure, competitive interactions as well as ecosystem functioning (Hooper et al. 2002, Cadotte et al. 2009, Isbell et al. 2011). The diversity of traits, which can be quantified by a variety of metrics (Rao 1982, Petchey and Gaston 2002, Villéger et al. 2008), indicates complementarity and redundancy among species. High trait diversity should reduce competitive interactions between species (Loreau and Hector 2001, Kunstler et al. 2016). Several studies also revealed that trait identity may play an even more important role than diversity (Mokany et al. 2008, Minden and Kleyer 2011). Despite these advances, the applicability of a trait-based approach to animal taxa and across trophic levels remains to be validated (Lavorel et al. 2013, Gravel et al. 2016).

Animal traits have been increasingly used to assess community response to various disturbances (Moretti and Legg 2009, Aubin et al. 2013, Lavorel et al. 2013, Frainer et al. 2014, Pey et al. 2014, Gagic et al. 2015). In most cases, studies were limited to only a few taxonomic and trophic groups. In comparison with autotrophs, animal traits and functions are both more diverse and more complex (Polis et al. 1997). For instance, there are no analogs in the plant kingdom for behavioral traits, or for the drastic changes in trait values exhibited by individuals throughout ontogeny (Aubin et al. 2013). Further, the difficulty of adequately representing the functional structure of animal communities comes in part from the presence of feeding loops, cannibalistic interactions, omnivory, intra-guild predation and ontogenic functional displacement preventing the distinction of trophic levels (Polis and Holt 1992, Thébault and Loreau 2006, Ings et al. 2009). The application to animal communities of functional indices developed for plants therefore requires new conceptual developments (Gravel et al. 2016).

It is widely documented that species' traits can influence community functioning directly, but they can also do so indirectly through their impact on trophic network structure and thereby on regulation (Gravel et al. 2016). The match between the traits of a consumer with those of a resource determines pairwise interactions (Morales-Castilla et al. 2015, Bartomeus et al. 2016). Several statistical methods have been proposed to link ecological traits with pairwise interactions, some of which are derived directly from food web theory (Williams 2010, Gravel et al. 2013). For instance, Rohr et al. (2010) and Eklöf et al. (2013) used latent variables to compute the interaction probability between pairs of species and relate them a posteriori to functional traits. Spitz et al. (2014) and Krasnov et al. (2016) performed a more direct analysis, replacing the matrices conventionally used in an RLQ analysis (Dolédec et al. 1996) with matrices containing consumer and resource trait values, and their interactions. These studies led to a more mechanistic understanding of the drivers of interactions. For instance, species body size was found to be the major driver of predator-prey interactions among marine fishes (Gravel et al. 2013) and terrestrial litter invertebrates (Digel et al. 2011). Ibanez et al. (2013) and Bartomeus et al. (2016) showed that mandible traits of scavengers and herbivores determine the type of litter and plants they are able to eat. Despite some technical distinctions between the various statistical approaches, they all have in common the idea that interactions depend on the match between "foraging traits", which characterize the way a consumer attacks and handles resource, and "vulnerability traits", which characterize defense and escape from predation (Rohr et al. 2010, Gravel et al. 2016).

Trait-matching constraints determine the occurrence of interactions between pairs of species (Bartomeus et al. 2016), but it is the distribution of all traits in the community that will determine the structure of the network (Gravel et al. 2016). The functional structure of a community can be described by its functional identity (average value of each trait) and its functional diversity (the variance of each trait). Several metrics can be used to characterize these two aspects. Community weighted mean (CWM) represents the average of each trait weighted by relative abundance (Garnier et al. 2004). Functional indices can be computed to represent the multi-dimensional functional volume of communities (Villéger et al. 2008, Mouillot et al. 2013). These methods have been used to analyze the functional structure of animal communities (Mouchet et al. 2010, Dehling et al. 2014, Coulis et al. 2015), although relatively little is known about their importance for network structure. For instance, even though we know that body size strongly constrains predatory interactions (Rohr et al. 2010, Eklöf et al. 2013), less is known about the effect of average and variance of the frequency distribution of body size on network properties such as connectance and degree distribution (Albouy et al. 2014). It is as yet unknown whether any relationship between functional indices and network properties would arise once accounting for multiple traits, for which empirical distributions have many more dimensions.

Investigating the relationship between functional and network structure is particularly relevant in soil communities where taxonomic and functional diversity is high. Interest in relationships between soil food webs, their functioning and the abiotic environment has grown steadily over the past several years (Berg and Bengtsson 2007, Melguizo-Ruiz et al. 2012, de Vries et al. 2013, Moya-Laraño et al. 2014). However, the role of functional structure has not yet attracted such attention. Soil organisms play a crucial role in key ecosystem processes, they are the main agents of nutrient cycling, and thus essential for primary production in all terrestrial ecosystems (Wardle 2006). Soil communities remain, nonetheless, among the least documented ecological systems (Bardgett and Wardle 2010). This is partly due to the difficulty of observing soil organisms, particularly their ecological interactions. Formal experiments such as cafeteria style trials, gut content analysis and isotope studies provide much needed information on species interactions (Brose and Scheu 2014). However, they are not yet widely used and are possibly too time consuming to be applied extensively. The dominance of generalism in species diet and of bottom up control of population dynamic lead to organic matter breakdown favored mainly by species complementarity (Hedde et al. 2010, Digel et al. 2014). Combined together, low resource competition and opportunistic feeding promote a high diversity of organisms and a

very strong body size structure (Wardle 2006). The exceptionally high diversity of certain guilds and of functionally redundant species, as well as the difficulty of observing species dynamics due to the physical nature of soils, make the use of the functional approach particularly promising to characterize soil food web structure.

Our objective in this study is to investigate the relationship between the functional structure of soil communities and the structure of the network of trophic interactions (see Fig. 1 for a summary description of the analytical steps followed). We first investigated the extent to which functional traits determine the occurrence of pairwise trophic interactions in soil food webs. Our study includes a particularly broad range of taxa, from microbes (bacteria and fungi) to macro-arthropods. As we had no a priori trait-matching rules and wanted to account for non-linear relationships between traits, we adopted a statistical algorithm derived from machine learning techniques. We then examined the relationship between descriptors of trait structure (identity and diversity) and network structure across 48 soil food webs in Germany. We used a collinearity analysis to compare two matrices describing functional and network structure. This study is a first step toward a better understanding of the relationship between the trait structure of multi-trophic communities and their functioning.

Methods

Datasets

We used two different datasets of trophic interactions in conjunction with a literature documentation of functional traits (Fig. 1). The first dataset was obtained from the study of Digel et al. (2014), who documented the presence and absence of interactions among 881 invertebrate species that were recorded at 48 sampling sites (hereafter called the food webs dataset). The study sites of Digel et al. (2014) were located in beech and coniferous forests across three geographical areas in Germany. Interactions were monitored using a combination of methods ranging from molecular gut content analyses to cafeteria experiments, the details of which are provided in the original publication. An absence of interaction was assumed between two species if they were found together at least in one location but were not observed interacting. Each food web contained between 89 and 168 species, with the total number of trophic interactions ranging from 729 to 3344. Overall, 34 193 unique interactions were observed across the 48 food webs, and a total of 215 418 absences of interactions. Some species, particularly small ones at low trophic levels, were identified as morphospecies or even functional groups (Supplementary material Appendix 1 Fig. A1).

We compiled a second dataset from literature review (hereafter called the independent interactions dataset) in order to improve representation of interactions involving detritivores and herbivores. The literature search was
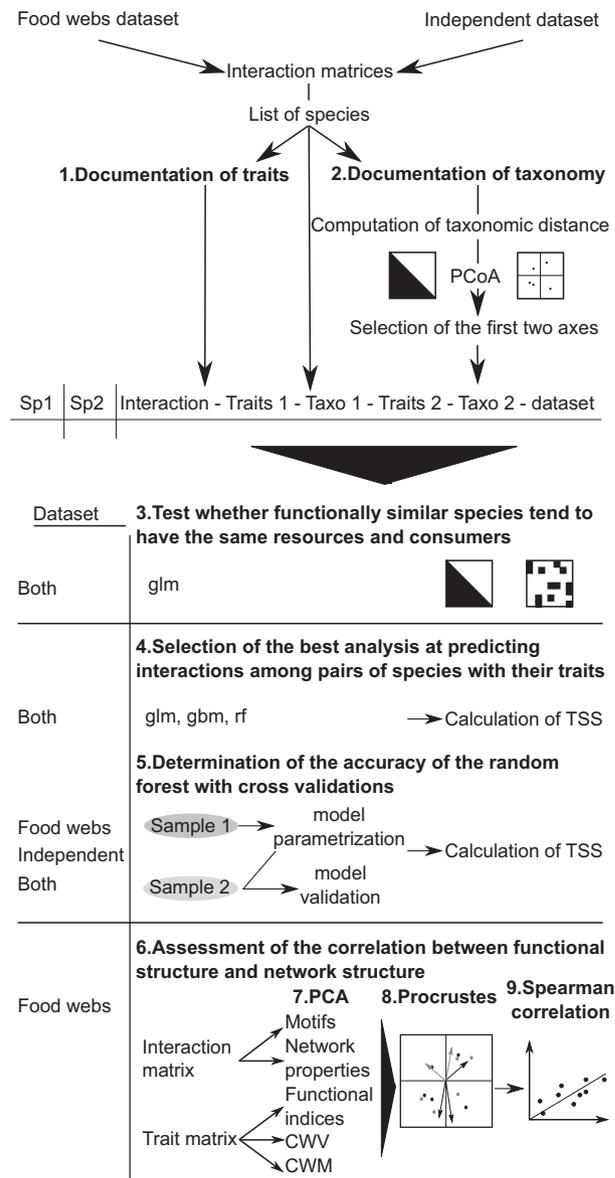

Figure 1. Schematic representation of the analyses performed. Each step is numbered chronologically and referred in the text. We used the two datasets (food webs + independent) with species traits and species interactions to ask two questions. 1) Can we predict species interactions from species traits? To answer this question, we used three types of models (generalized linear model, gradient boosted model and random forest) parametrized using several samples, and validated with different samples (cross-validation procedure). 2) Are certain aspects of food web structure related to functional structure? To answer this question we computed descriptors of network structure, functional identity and functional diversity. We linked each component of functional structure to each component of network structure with a Procrustes analysis.

carried out using Google Scholar with keywords including taxonomic group names and keywords indicating trophic interactions ('interaction', 'diet', 'consumption', 'cafeteria experiment', 'feed'). We selected articles involving interactions between terrestrial invertebrate species and any

resource, for a total of 126 studies (Supplementary material Appendix 2 Table A1). Across these, a total of 1273 interactions were recorded between 645 species. Only 88 absences of interaction were found.

We selected eight traits as potential indicators of either foraging or vulnerability attributes (Table 1, Fig. 1.1). These traits were documented for each species or taxa based on a literature review, as well as from visual assessment of images. In addition to these traits, we included two proxies for hard-to-measure traits. We considered feeding guild as a trait syndrome (co-occurring trait attributes that can be sued to synthesize a function, McIntyre et al. 1999) and a proxy of characteristics that could not be measured directly on an organism (e.g. the ability to digest plant material). Also, taxonomy was used as a latent trait (i.e. non-measured traits or syndrome that are estimated from the data, Rohr et al. 2010), following the eco-phylogenetic approach of Mouquet et al. (2012). Only a handful of studies (Rohr et al. 2010, Eklöf et al. 2012, Krasnov et al. 2016) have investigated thus far the joint contribution of traits and taxonomy to the occurrence of interactions. Here, we relied on an index of taxonomic distance to overcome the lack of time calibrated phylogenies and the challenge of accounting simultaneously for a wide range of organisms (from bacteria to arthropods). To do so, we used the R package 'ape' to obtain taxonomic distances between species and then conducted a principal coordinates analysis (PCoA) on taxonomic distances to record the score of each species from the first two axes (Fig. 1.2). This approach therefore only accounts for the topology of the tree and not the rate of evolutionary processes.

Statistical analysis

Determining interactions in soil food webs from traits

We first tested if functionally similar species tend to have the same resources and consumers (Fig. 1.3). To do so, we calculated Gower distances (Podani and Schmera 2006) between pairs of species according to their trait values. Then, we tested if the distance between a species A and B is related to the probability that both species A and B share an interaction

(or lack of) with species C. To do so, we fitted a GLM with shared interactions (0 or 1) between species A and B as a function of the distance between them. Significance was assessed with a χ^2-test.

Secondly, we investigated the role of foraging and vulnerability traits in driving the occurrence of pairwise interactions (Fig. 1.4). We considered only direct trophic interactions. We considered that the occurrence of a direct interaction between a consumer and a resource (response variable) is a binomial process conditional on their traits (explanatory variables). We compared three statistical techniques to predict the probability of interaction among species: generalized linear models (GLM), gradient boosted models (GBM), and random forest models (RF). GBM and RF are machine learning algorithms (Breiman 2001), which, using decision tree-like processes, find the best combination of variables and variable coefficients to explain the response variable. This method deals easily with non-linear relationships and is more flexible than GLMs. GLMs, however, represent the trait-matching centrality formalism developed by Rohr et al. (2010), which is grounded in food web theory and is essentially the common denominator to most trait-matching methods (Gravel et al. 2016).

We implemented different types of cross-validation to assess the model accuracy (Fig. 1.5). We first performed calibration on the food webs dataset and validated on the independent interactions dataset, and the other way around. We also performed cross-validation within the food webs dataset by calibration on two geographical areas and validation on the third area. We also performed bootstrapping cross-validation by randomly selecting 36 out of the 48 food webs for calibration and validation on the remaining 12 food webs (25% of the dataset). Random selection of food webs for cross-validation was repeated ten times. Finally, we selected two samples without any species in common, by taking randomly pairs of species inside the 48 food webs: one was used for parametrization, and the other for validation. The accuracy of the predictions of the models was calculated with the true skill statistic (TSS). The TSS quantifies the proportion of prediction success relative to false predictions and returns values ranging between 1 (perfect predictions) and –1 (inverted forecast) (Allouche et al. 2006). We evaluated the individual

Table 1. Description of traits (and proxies for other traits) used in the different models. (C): carnivorous interactions, (HFD): herbivorous–fungivorous–detritivorous interactions. *Food webs dataset and body mass measurements were provided by Digel et al. (2014). Body mass was hard to obtain from the literature; therefore, for the independent interactions dataset we assessed body mass from body size, using documentation protocols described in Andrassy (1956, Freckman 1982) and Ganihar (1997).

	Trait	Type	Description	Documentation	Model
Observed traits	prey capture strategy	boolean	web builder or not	literature	C
	poison	boolean	use of poison to kill prey or not	literature	C
	body mass	continuous	logarithm of the mass of an individual (in grams)	measurement, literature*	C, HFD
	mobility	categorical	1: immobile, 2: crawling (no legs), 3: short legs, 4: long legs, 5: jumping, 6: flight	literature	C
	toughness	categorical	0: soft (no chitin, or few lignin), 0.5: hard, 1: has a shell (or is a seed)	literature	C, HFD
Latent traits	feeding guild	boolean	carnivore: 1/0, detritivore: 1/0, fungivore: 1/0, herbivore: 1/0	literature	HFD
	taxonomy	continuous	scores on the 2 pcoa axes of the taxonomy	literature	C, HFD
	vertical stratification	boolean	below soil surface : 1/0 and/or above soil surface : 1/0	literature	C, HFD

contribution of the different traits to the occurrence of interactions. Models were fitted by removing one trait at a time and TSS of the reduced model was compared to the full model. Predictions were significantly improved by separating the model in two: one for carnivorous interactions, and one for interactions in which resources are microbes, dead matter, or producers. We report the results of these two models separately.

Relationship between descriptors of functional structure and of network structure

The next step of our analysis was to investigate the relationship between functional and network structure using the food webs dataset (Fig. 1.6). We computed CWMs of each trait for each local food web. Functional diversity was characterized using community trait variance (CWV, variance of each non-binary trait, Sonnier et al. 2010), functional richness and functional evenness (proportion of functional space filled by the species community and regularity of abundance distributions in the functional space, respectively, Villéger et al. 2008).

Standard network properties (Williams and Martinez 2000, Newman and Girvan 2004, Vermaat et al. 2009) were computed for each of the 48 food webs: number of species, number of links, connectance (number of links observed divided by the square of species richness), generality (standard deviation of normalized number of resource per species), vulnerability (standard deviation of normalized number of consumers per species), mean trophic level (weighted trophic level across the different paths from the species to the base of the food web), omnivory (fraction of species that consume resources at more than one trophic level), modularity (clustering of species into groups with which they interact more than with species in other groups), mean maximum similarity (mean of the maximum trophic similarity of each taxon to other taxa, i.e. the number of consumers and resources shared by a pair of species divided by their total number of consumers and resources). We also considered the more refined representation of network structure with the frequency distribution of four node motifs. Motifs are subset of three nodes viewed as the building blocks of diverse networks and their frequency distribution has been related to food web dynamics (Stouffer 2010). We selected the motifs representing well-known food web modules: linear food chain, apparent competition (a single consumer with two resources), resource competition (two consumers on the same resource), and omnivory (a consumer feeding simultaneously on two different trophic levels).

Basal species had a low taxonomic resolution and were therefore removed them from the computation of the different network metrics to avoid any bias. We performed a principal component analysis (PCA) on the matrices of CWM, CWV, functional indices, network properties and motif frequency in order to position the different communities in multivariate functional and network spaces (Fig.1.7). We compared these ordinations using a Procrustes analysis (Jackson 1995) and tested the hypothesis that the relative position of different food webs was related in the two ordination spaces (Fig. 1.8). The Procrustes method finds the best linear fit between two ordinations: one of the ordinations is reflected, rotated and expanded in order to position food webs as close to their counterpart as possible in the second multivariate space. Statistical significance of the correlation was assessed using a Procrustean randomization test (PROTEST), which tests whether the sum of residual deviations of the communities matching is less than expected by chance (Jackson 1995). We also computed the observed correlation between matrices to estimate the goodness of fit. We performed a Spearman correlation between pairs of metrics in order to more finely investigate univariate relationships (Fig. 1.9).

All of the analyses were performed with R (<www.r-project.org>) and the packages 'ade4' (Dray and Dufour 2007), 'gbm' (Ridgeway 2015), 'randomForest' (Liaw and Wiener 2002) and 'ape' (Paradis et al. 2004).

Data deposition

Data available from the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.1sg4j> (Laigle et al. 2017).

Results

Trait-matching predicts a large part of the variation in the occurrence of pairwise interactions. First, we found that functionally close species tend to have the same resources and consumers (comparison between the null model and the two other models, $p < 2.10^{-16}$). Secondly, we found that all three statistical analyses performed well at predicting interactions from traits. The GLM had the lowest TSS value (TSS of 0.06 for carnivorous interactions and 0.38 for other interactions), followed by the GBM (TSS = 0.40 and TSS = 0.41), and the best model was the RF (TSS = 0.98 and TSS = 0.86). Most cross-validation scenarios were satisfying (Table 2). TSS of cross-validations between food webs from the three different geographical areas ranged between 0.89 and 0.93 for carnivorous interactions, and between 0.59 and 0.61 for non carnivorous interactions. Cross-validation by removing 25% of interactions yielded TSS of 0.77 and 0.79 for carnivorous and other interactions, respectively. Trait-matching constraints were however restricted to the dataset, as revealed by the poor performance of cross-validation between datasets, with TSS between –0.16 and 0.03. Overall, we found that the most important trait or proxy to predict species interactions are body mass and taxonomy, followed by toughness of the resource, use of poison for carnivorous interactions, kingdom of the resource, and consumer diet for non-carnivorous (Fig. 2).

We evaluated the relationship between CWM and network properties (Fig. 3a) and found that the functional identity of species making up a community is related to network structure (p <1.10^{-4} for the PROTEST, R = 0.62). We present all the spearman coefficients between variables in Supplementary material Appendix 3 Table A2. Not

Table 2. Cross-validation results for the carnivorous interactions (C) and the herbivorous-fungivorous-detritivorous interactions (HFD). Random forest models were calibrated with one sub-dataset and validated on a second one. Several cross-validation scenarios were considered: between food webs from 3 different geographical areas; between randomly selected food webs (36 for parametrization, 12 for validation); between the independent interactions dataset and the food webs dataset; between samples that do not have any species in common. FW=food webs dataset, II=independent interactions dataset, a=number of observed and predicted interaction presence, b=number of observe absence and predicted presence, c=number of observed presence and predicted absence, d=number of predicted and observed absence.

Validation	Data for parametrization	Data for validation	Model	a	b	c	d	TSS
Cross areas	FW: x+y	FW: z	RF_C	12738	570	1115	86056	0.91
			RF_{HFD}	312	32	195	3384	0.60
Cross FW	38 food webs	12 food webs	RF_C	12535	271	350	83736	0.97
			RF_{HFD}	427	38	105	3429	0.79
Cross datasets	II	FW	RF_C	32283	202136	0	0	0
			RF_{HFD}	890	6150	0	0	0
	FW	II	RF_C	32	9	559	33	−0.16
			RF_{HFD}	9	0	301	32	0.13
Cross samples	FW	FW	RF_C	3299	385	958	22519	0.76
			RF_{HFD}	24	85	98	1068	0.12
	II	II	RF_C	342	14	1	0	0
			RF_{HFD}	183	4	11	15	0.73

surprisingly, mean trophic level, omnivory and connectance all increase with the number of carnivorous species (Spearman coefficients of r=0.88, r=0.74 and r=0.69, respectively). The number of links, vulnerability, generality, and also omnivory increase with the number of species that use poison to hunt (r=0.65, r=0.60, r=0.63, and r=0.60). Vulnerability also decreases with species that live belowground (r=−0.51). Finally, trophic similarity decreases with the first PCoA axis representing species taxonomy, and with species body mass (r=−0.51 and r=−0.45).

Secondly, we found that CWV is correlated to network properties (p<1.10−4 for the PROTEST, R=0.58) (Fig. 3b). Variance in body mass and in the first axis of the PCoA of the taxonomy is negatively correlated to trophic similarity (r=−0.51 and r=−0.49). Variance in toughness is negatively correlated to vulnerability, mean trophic level and omnivory (r=−0.60, r=−0.72, r=−0.79), while variance in mobility is negatively correlated to mean trophic level (r=−0.65). Similar relationships were also observed for other diversity metrics, although the relationships were weaker (with a p-value of < 1.10−4 for the PROTEST, and R=0.45), (Fig. 3c). Species richness is positively related to functional richness (r=0.71), but not functional evenness (r=0.05). Trophic similarity varies inversely to functional evenness (r=−0.48). Modularity is not correlated to any functional index.

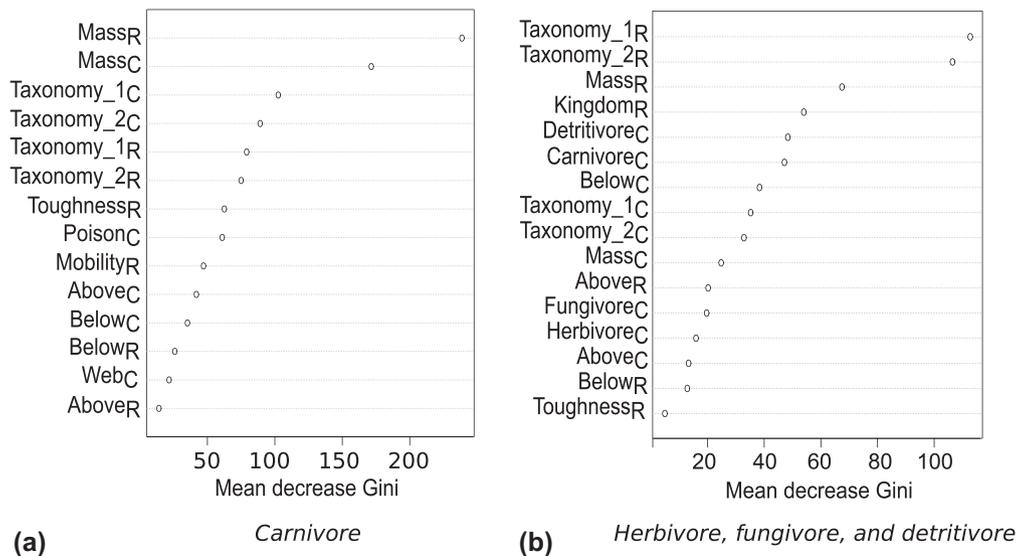(a) Carnivore

(b) Herbivore, fungivore, and detritivore

Figure 2. Importance of traits to predict carnivorous and non-carnivorous interactions according to the random forest model (see Table 2 for details on traits). The 'mean decrease Gini' (total decrease in node impurities from splitting on the variable) is proportional to the contribution of each variable in the model. High mean decrease Gini represents a high importance of the variable. C: consumer traits, and R: resource traits.
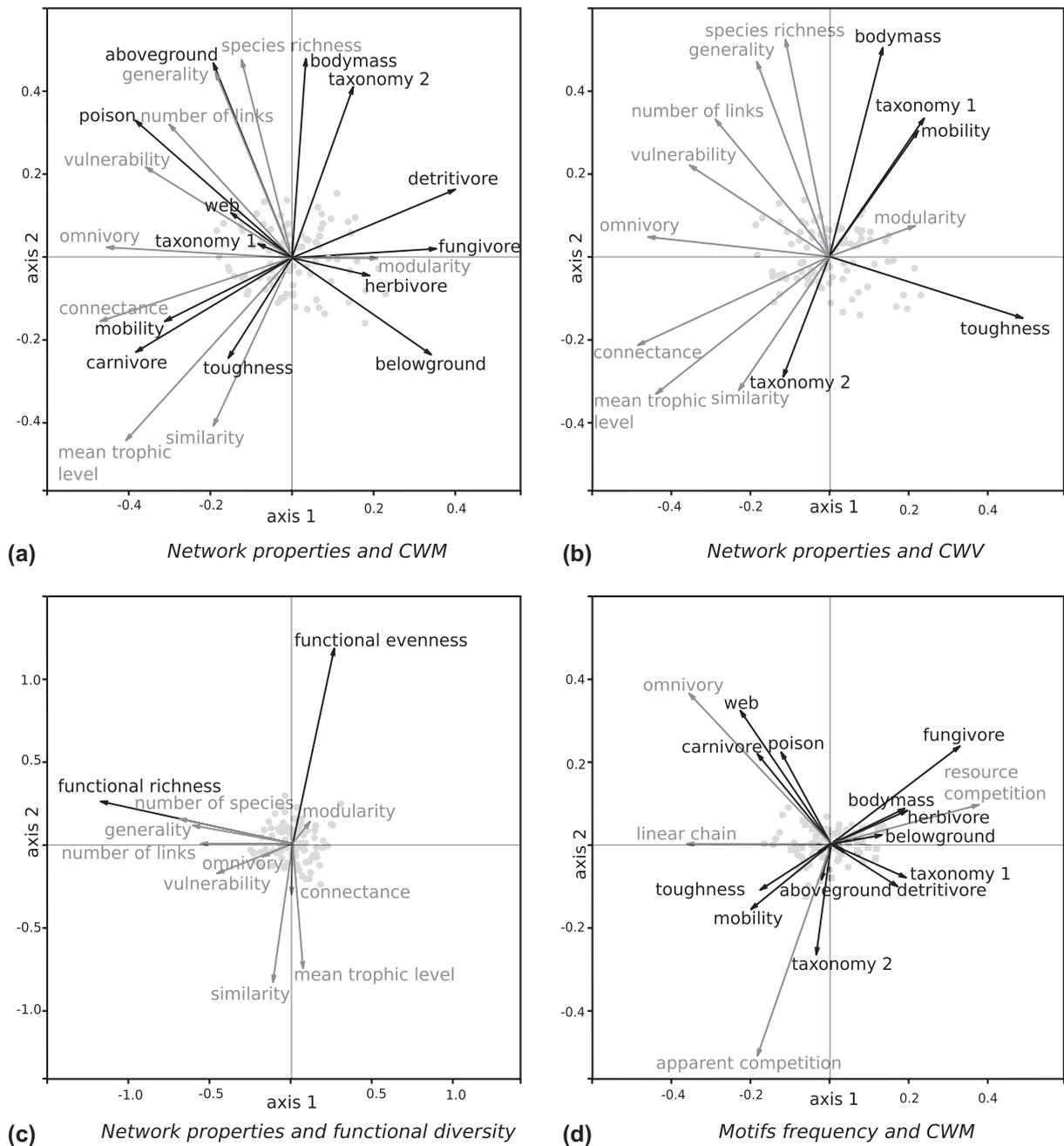
Figure 3. Procrustes representation of the correlation between the variables representing the functional structure (community weighted mean, community weighted variance, functional diversity indices) and those reflecting network structure (network properties and motifs frequency). Each dot corresponds to a local community.

The frequency of the different motifs was found to be independent of functional metrics (p-value = 0.46, R = 0.17), but correlated to the CWM (p < 1.10^{-4}, R = 0.40) (Fig. 3d). Fungivores, herbivores and detritivores increase with the frequency of motifs representing resource competition (r = 0.65, r = 0.39 and r = 0.19, respectively), while carnivores increase with the frequency of motifs related to omnivory (r = 0.37). Omnivory increases even more if carnivores use poison or webs to catch their prey (r = 0.34 and r = 0.64).

Discussion

This study is a first step toward a better understanding of the role of functional traits in the structure of trophic interaction networks (Fig. 4). Traits are known to be important drivers of assembly processes in plant communities (McGill et al. 2006, Violle et al. 2007). They also play a key role in other important ecosystem functions such as primary productivity. Building on this principle, this study shows that the structure of more complex communities such
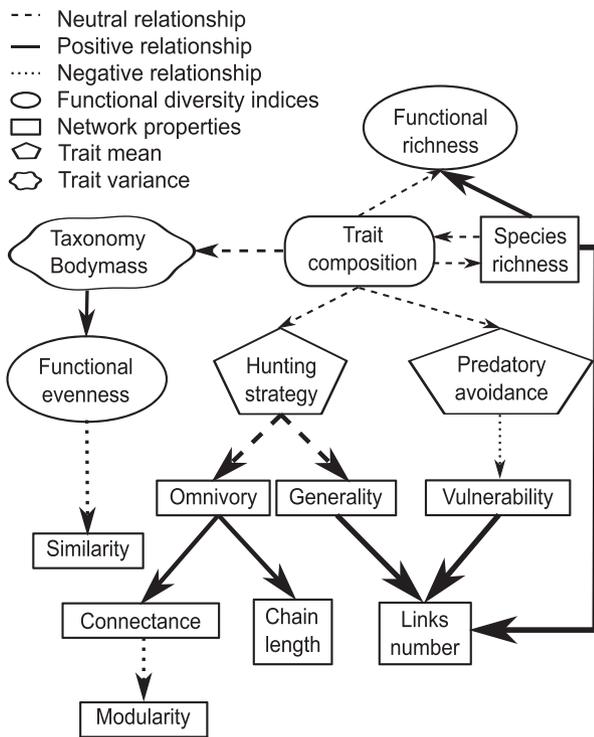
Figure 4. Summary diagram showing observed relationships between functional diversity properties, network properties and traits mean and variance. Functional properties are surrounded, network properties are framed, trait means are framed with a pentagon and trait variances are written inside a cloud. Relationships between variables are represented with arrows: solid = positive correlation; dotted = negative; dashed = neither positive nor negative. The width of the arrows represents the strength of the relationship.

as soil food webs is also related to the different aspects of the trait distribution. We found that traits such as body size, mobility and hunting mode affect the occurrence of trophic interactions in soil food webs, but other aspects of the community functional structure such as trait identity and diversity were also related to the summary characteristics of the network structure. These results show how the trait-based approach to food webs can efficiently and mechanistically position species within a community and explain its macroscopic properties.

From traits to interactions

We successfully used a trait-matching approach to determine soil food web interactions. We were able to predict 98% of interactions between animal resources and consumers using a small set of traits in addition to taxonomy as a proxy for unmeasured traits. Our method, inspired by the matching centrality formalism (Rohr et al. 2010), compared the accuracy of three statistical approaches (generalized linear models, gradient boosted model, and random forest models) to predict species interactions from their traits. Random forest (RF) was the most efficient method, indicating that

trait-matching constraints are not linear and possibly multivariate. The random forest algorithm is based on regression trees, allowing the classification of interactions with a set of conditions, a process similar to what ecologists would intuitively do when constructing an interaction network. Cross-validation results were satisfying, both when tested in different geographical areas (leaving one area out of the calibration data) and with different species (leaving a set of species out of the calibration data). Non-carnivorous interactions were better predicted in the independent interactions dataset, which contained more precise non-animal species identification. On the other hand, carnivorous interactions were better predicted in the food webs dataset, in which species identification was done only for animal species. We nonetheless have to remain careful about the ability of the model to predict interactions of species in new systems. Indeed, the random forest algorithm tends to over-fit the data, as we exemplified by the validation across datasets.

In agreement with previous studies conducted in other systems (Brose et al. 2006, Riede et al. 2010, Williams 2010, Gravel et al. 2013), body mass was found the trait that best predicts interactions, followed by taxonomy. Taxonomy can be used as a proxy, approximating traits that are hard to document (Mouquet et al. 2012), such as chemical defense or behavior, following the assumption that some traits are highly conserved within taxa. Taxonomy also helps capturing trait syndromes, such as the ant *Plectroctena* which uses nutcracker mandibles in combination with a hunting strategy specific to the predation of millipedes (Polidori 2011). Further, kingdom of the resource and diet of the consumer provides an approximate way to characterize the backbone of network structure, with the identification of interactions that are trivially prohibited also called forbidden interactions, such as herbivores not feeding on animals (Morales-Castilla et al. 2015). As more studies are published, it will be interesting to see whether the amount of variance explained by taxonomy decreases with increasing quality of trait information, or if some fundamental constraints cannot be disentangled only using trait data and remain best explained by evolutionary history.

From functional structure to network structure

Having verified the existence of a relationship between traits and the occurrence of interactions, we then explored relationships between functional and network structure. The strongest associations were observed between trait composition (i.e. CWM), network properties, and motifs frequency. While we did not had a priori expectations, the results we found were all coherent. For instance, it was not surprising to find that a higher number of carnivorous species was correlated to the average trophic level, but also connectance. We also found that the use of poison or webs were positively related to the number of links, and the frequency of the motif that represents omnivory. In contrast, vulnerability decreased with the number of species living below-ground,

presumably because it limits predation by spiders. These results suggest that hunting strategy is a representative trait of species generality, and in combination with species vulnerability, both determine the number of links.

The relationships we found between functional diversity metrics and network properties were more intriguing. First, relationships between the range of trait values and network properties were more interesting when looking at each trait individually, than when considering functional richness. Correlations between functional richness and network properties were likely driven by the strong dependence of this index to species richness (Villéger et al. 2008), which in turn increase the number of links per species (Martinez 1992), as well generality and vulnerability. We expected that functional richness, by increasing functional differences among species, would be negatively correlated to trophic similarity. However, trophic similarity was rather negatively correlated to variance in body mass and taxonomy. This observation is consistent with results of the first part of the study. While functional richness could provide insights about the amount of functions a community can provide, it should considered in conjunction with individual trait investigation to approximate network structure.

Functional evenness also had a negative relationship with trophic similarity, but not with modularity and connectance as we would have expected. Low functional evenness characterizes the distribution of species in functional groups within the community. Therefore, low functional evenness increased the availability of resource and the number of consumers (mainly determined by matching of their taxonomy and body mass), thus increasing trophic similarity. Similarly, we would have expected a decrease in modularity with functional evenness. Modularity can be seen as a measure of trophic complementarity between functional groups of species (Poisot et al. 2013, Montoya et al. 2015). For example, Olesen et al. (2007) showed that modules of species with convergent traits values are present in species rich pollinators–plant communities. Thus, it would have been reasonable to expect modules of similar consumers (i.e. functional groups) sharing the same resources. We suspect our results might stand from the high amount of generalist predators linking modules. We also only found a weak relationship between functional evenness and connectance. Then, functional evenness provides insight into species trophic similarity and thus competition and complementary among species.

Our analysis had some limitations, which should orient future research. Taxonomic resolution of basal trophic groups was very low, therefore influencing the number of species and links in each network. This limitation does not impact the analysis of trophic interactions among carnivorous species, but certainly does for detritus and plants. These nodes are key for major ecosystem processes such as nutrient cycling, therefore, more precise species identification and documentation of interactions are needed. The importance of taxonomy for the explanation of pairwise interactions also underlines the coarse resolution of trait information

we were able to gather from literature. Taxonomy does not inform us about the mechanisms involved in interactions, and limits our model to taxa present in the training data. Moreover, the decision tree computed by the random forest, for the two datasets together, required more than 10 000 nodes, for 250 972 interactions. The rules of the random forest are likely specific to the training data. If we want to correctly predict interaction for a new pair of species absent from the training dataset, their traits should be similar to one of the 20 000 groups of pairs of the training species. We found that traits mediate species interactions, but the complexity of statistical model made the interpretation difficult. Other traits could perform better at discriminating interactions by requiring less nodes, and could replace taxonomy. Traits potentially mediating species interactions (Moretti et al. 2016) were difficult to document precisely due to a lack of available information. We expect that the documentation of traits representing foraging strategy and predatory avoidance, as well as spatial and temporal co-occurrence should improve significantly our ability to predict predatory interactions. Non-animal interactions could also be predicted more accurately with traits such as chemical defense and stoichiometry. Trait measured directly on the individuals for which interactions were observed could also improve our understanding of trait-matching constraints (Bartomeus et al. 2016). That said, we emphasize the difficulty of finding traits that are relevant and could be measured across the range of organisms we considered. There is currently no standard that would account for the variability of characteristics that could be found from nematodes to spiders. Further, while additional traits might refine our ability to predict exceptional interactions, like highly specialized consumer–resource matches, each additional trait is susceptible to decrease the generality of the model. Future investigations of trait-matching constraints will therefore require a proper analysis of the optimal solution to the generality–accuracy tradeoff.

Conclusion

We offer a new approach to answer quantitatively questions which ecologists often addressed intuitively. Despite some limitations, we showed how trait composition, taxonomy and body mass in particular, determine the distribution of species interactions in a community. The growing interest for trait-based approaches to soil fauna will likely improve trait documentation, and thereby the precision of future analyses. As Schleuning et al. (2015) and Gravel et al. (2016) have recently pointed out, the next challenge in biodiversity and ecosystem functioning studies will be to relate network ecology to ecosystem processes. Answers to these questions have the potential to upend existing theories by demonstrating the changing effect and importance of functional diversity across trophic levels, as suggested by Lefcheck and Duffy (2015). The joint study of functional and network

structure may provide valuable insight by bridging important conceptual gaps.

References

Albouy, C. et al. 2014. From projected species distribution to food-web structure under climate change. – Global Change Biol. 20: 730–741.

Allouche, O. et al. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). – J. Appl. Ecol. 43: 1223–1232.

Andrassy, I. 1956. Die Rauminhalts- und Gewichtsbestimmung der Fadenwürmer (Nematoden). – Acta Zool. Acad. Sci. Hung. 2: 1–15.

Aubin, I. et al. 2013. Can a trait-based multi-taxa approach improve our assessment of forest management impact on biodiversity? – Biodiv. Conserv. 22: 2957–2975.

Bardgett, R. D. and Wardle, D. A. 2010. Aboveground–belowground linkages: biotic interactions, ecosystem processes and global change. – Oxford Univ. Press.

Bartomeus, I. et al. 2016. A common framework for identifying rules across different types of interactions. – Funct. Ecol. 10: 1–40.

Berg, M. P. and Bengtsson, J. 2007. Temporal and spatial variability in soil food web structure. – Oikos 116: 1789–1804.

Breiman, L. 2001. Random forests. – Machine Learning 45: 5–32.

Brose, U. and Scheu, S. 2014. Into darkness: unravelling the structure of soil food webs. – Oikos 123: 1153–1156.

Brose, U. et al. 2006. Allometric scaling enhances stability in complex food webs. – Ecol. Lett. 9: 1228–1236.

Cadotte, M. W. et al. 2009. Using phylogenetic, functional and trait diversity to understand patterns of plant community productivity. – PloS One 4: e5695.

Coulis, M. et al. 2015. Functional dissimilarity across trophic levels as a driver of soil processes in a Mediterranean decomposer system exposed to two moisture levels. – Oikos 124: 1304–1316.

de Vries, F. T. et al. 2013. Soil food web properties explain ecosystem services across European land use systems. – Proc. Natl Acad. Sci. USA 110: 14296–14301.

Dehling, D. M. et al. 2014. Functional relationships beyond species richness patterns: trait matching in plant–bird mutualisms across scales. – Global Ecol. Biogeogr. 23: 1085–1093.

Digel, C. et al. 2011. Body sizes, cumulative and allometric degree distributions across natural food webs. – Oikos 120: 503–509.

Digel, C. et al. 2014. Unravelling the complex structure of forest soil food webs: higher omnivory and more trophic levels. – Oikos 123: 1157–1172.

Dolédec, S. et al. 1996. Matching species traits to environmental variables: a new three-table ordination method. – Environ. Ecol. Stat. 3: 143–166.

Dray, S. and Dufour, A. 2007. The ade4 package: implementing the duality diagram for ecologists. – J. Stat. Software. 22: 1–20.

Eklöf, A. et al. 2012. Relevance of evolutionary history for food web structure. – Proc. R. Soc. B 279: 1588–1596.

Eklöf, A. et al. 2013. The dimensionality of ecological networks. – Ecol. Lett. 16: 577–583.

Frainer, A. et al. 2014. When does diversity matter? Species functional diversity and ecosystem functioning across habitats and seasons in a field experiment. – J. Anim. Ecol. 83: 460–469.

Freckman, D. 1982. Nematodes in soil ecosystems. – Univ. of Texas Press.

Gagic, V. et al. 2015. Functional identity and diversity of animals predict ecosystem functioning better than species-based indices. – Proc. R. Soc. B 282: 201442620.

Ganihar, S. R. 1997. Biomass estimates of terrestrial arthropods based on body length. – J. Biosci. 22: 219–224.

Garnier, E. et al. 2004. Plant functional markers capture ecosystem properties during secondary succession. – Ecology 85: 2630–2637.

Gravel, D. et al. 2013. Inferring food web structure from predator–prey body size relationships. – Methods Ecol. Evol. 30: 347–356.

Gravel, D. et al. 2016. The meaning of functional trait composition of food webs for ecosystem functioning. – Phil. Trans. R. Soc. B 371: 20150268.

Hedde, M. et al. 2010. Patterns and mechanisms responsible for the relationship between the diversity of litter macro-invertebrates and leaf degradation. – Basic Appl. Ecol. 11: 35–44.

Hooper, D. U. et al. 2002. Species diversity, functional diversity and ecosystem functioning. – In: Loreau, M. et al. (eds), Biodiversity and ecosystem functioning: a current synthesis. Oxford Univ. Press, pp. 195–208.

Ibanez, S. et al. 2013. Herbivory differentially alters litter dynamics of two functionally contrasted grasses. – Funct. Ecol. 27: 1064–1074.

Ings, T. C. et al. 2009. Ecological networks–beyond food webs. – J. Anim. Ecol. 78: 253–269.

Isbell, F. et al. 2011. High plant diversity is needed to maintain ecosystem services. – Nature 477: 199–202.

Jackson, D. A. 1995. PROTEST: a procrustean randomization test of community environment concordance. – Écoscience 2: 297–303.

Krasnov, B. R. et al. 2016. Trait-based and phylogenetic associations between parasites and their hosts: a case study with small mammals and fleas in the Palearctic. – Oikos 125: 29–38.

Kunstler, G. et al. 2016. Plant functional traits have globally consistent effects on competition. – Nature 529: 204–207.

Laigle, I. et al. 2017. Data from: Species traits as drivers of food web structur. – Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.1sg4j>.

Lavorel, S. et al. 2013. A novel framework for linking functional diversity of plants with other trophic levels for the quantification of ecosystem services. – J. Veg. Sci. 24: 942–948.

Lefcheck, J. S. and Duffy, E. J. 2015. Multitrophic functional diversity predicts ecosystem functioning in experimental assemblages of estuarine consumers. – Ecology 96: 2973–2983.

Liaw, A. and Wiener, M. 2002. Classification and regression by randomForest. – R News 2: 18–22.

Loreau, M. and Hector, A. 2001. Partitioning selection and complementarity in biodiversity experiments. – Nature 412: 72–76.

Martinez, N. D. 1992. Constant connectance in community food webs. – Am. Nat. 139: 1208–1218.

McGill, B. J. et al. 2006. Rebuilding community ecology from functional traits. – Trends Ecol. Evol. 21: 178–185.

McIntyre, S. et al. 1999. Disturbance response in vegetation – towards a global perspective on functional traits. – J. Veg. Sci. 10: 621–630.

Melguizo-Ruiz, N. et al. 2012. Potential drivers of spatial structure of leaf-litter food webs in southwestern European beech forests. – Pedobiologia 55: 311–319.

Minden, V. and Kleyer, M. 2011. Testing the effect–response framework: key response and effect traits determining above-ground biomass of salt marshes. – J. Veg. Sci. 22: 387–401.

Mokany, K. et al. 2008. Functional identity is more important than diversity in influencing ecosystem processes in a temperate native grassland. – J. Ecol. 96: 884–893.

Montoya, D. et al. 2015. Functional group diversity increases with modularity in complex food webs. – Nat. Comm. 6: 7379.

Morales-Castilla, I. et al. 2015. Inferring biotic interactions from proxies. – Trends Ecol. Evol. 30: 347–356.

Moretti, M. and Legg, C. J. 2009. Combining plant and animal traits to assess community functional responses to disturbance. – Ecography 32: 299–309.

Moretti, M. et al. 2016. Handbook of protocols for standardized measurement of terrestrial invertebrate functional traits. – Funct. Ecol. 31: 558–567.

Mouchet, M. A. et al. 2010. Functional diversity measures: an overview of their redundancy and their ability to discriminate community assembly rules. – Funct. Ecol. 24: 867–876.

Mouillot, D. et al. 2013. A functional approach reveals community responses to disturbances. – Trends Ecol. Evol. 28: 167–177.

Mouquet, N. et al. 2012. Ecophylogenetics: advances and perspectives. – Biol. Rev. Camb. Phil. Soc. 87: 769–785.

Moya-Laraño, J. et al. 2014. Eco-evolutionary spatial dynamics. Rapid evolution and isolation explain food web persistence. – Adv. Ecol. Res. 50: 75–144.

Newman, M. E. J. and Girvan, M. 2004. Finding and evaluating community structure in networks. – Phys. Rev. E 69: 1–15.

Olesen, J. M. et al. 2007. The modularity of pollination networks. – Proc. Natl Acad. Sci. USA 104: 19891–19896.

Paradis, E. et al. 2004. APE: Analyses of phylogenetics and evolution in R language. – Bioinformatics 20: 289–290.

Petchey, O. L. and Gaston, K. J. 2002. Functional diversity (FD), species richness and community composition. – Ecol. Lett. 5: 402–411.

Pey, B. et al. 2014. Current use of and future needs for soil invertebrate functional traits in community ecology. – Basic Appl. Ecol. 15: 194–206.

Podani, J. and Schmera, D. 2006. On dendrogram-based measures of functional diversity. – Oikos 115: 179–185.

Poisot, T. et al. 2013. Trophic complementarity drives the biodiversity–ecosystem functioning relationship in food webs. – Ecol. Lett. 16: 853–861.

Polidori, C. 2011. Predation in the hymenoptera: an evolutionary perspective. – Transworld Research Network.

Polis, G. and Holt, R. D. 1992. Intraguild predation: the dynamics of complex trophic interactions. – Trends Ecol. Evol. 7: 151–154.

Polis, G. et al. 1997. Toward an integration of landscape and food web ecology: the dynamics of spatially subsidized food webs. – Annu. Rev. Ecol. Syst. 28: 289–316.

Rao, C. R. 1982. Diversity: its measurement, decomposition, apportionment and analysis. – Sankhya Indian J. Stat. A 44: 1–22.

Ridgeway, G. 2015. Gbm: Generalized boosted regression models. – R package ver 1(3), 55.

Riede, J. O. et al. 2010. Scaling of food-web properties with diversity and complexity across ecosystems. – In: Woodward, G. (ed), Ecological networks. Advances in ecological research. Academic Press, pp. 139–170.

Rohr, R. P. et al. 2010. Modeling food webs: exploring unexplained structure using latent traits. – Am. Nat. 176: 170–177.

Schleuning, M. et al. 2015. Predicting ecosystem functions from biodiversity and mutualistic networks: an extension of trait-based concepts to plant–animal interactions. – Ecography 38: 380–392.

Sonnier, G. et al. 2010. Quantifying relationships between traits and explicitly measured gradients of stress and disturbance in early successional plant communities. – J. Veg. Sci. 21: 1014–1024.

Spitz, J. et al. 2014. Let's go beyond taxonomy in diet description: testing a trait-based approach to prey–predator relationships. – J. Anim. Ecol. 83: 1137–1148.

Stouffer, D. B. 2010. Scaling from individuals to networks in food webs. – Funct. Ecol. 24: 44–51.

Thébault, E. and Loreau, M. 2006. The relationship between biodiversity and ecosystem functioning in food webs. – Ecol. Res. 21: 17–25.

Vermaat, J. E. et al. 2009. Major dimensions in food-web structure properties. – Ecology 90: 278–282.

Villéger, S. et al. 2008. New multidimensional functional diversity indices for a multifaceted framework in functional ecology. – Ecology 89: 2290–2301.

Violle, C. et al. 2007. Let the concept of trait be functional! – Oikos 116: 882–892.

Wardle, D. A. 2006. The influence of biotic interactions on soil biodiversity. – Ecol. Lett. 9: 870–886.

Williams, R. J. 2010. Simple MaxEnt models explain food web degree distributions. – Theor. Ecol. 3: 45–52.

Williams, R. J. and Martinez, N. D. 2000. Simple rules yield complex food webs. – Nature 404: 180–183.

Supplementary material (available online as Appendix oik-04712 at <www.oikosjournal.org/appendix/oik-04712>). Appendix 1–3.