Telling mutualistic and antagonistic ecological networks apart by learning their multiscale structure

Short title: Linking multiscale network structure with interaction type

Benoît Pichon[1,2,*] (ORCID : 0009-0007-6206-5482)

Rémy Le Goff [2] (ORCID : 0009-0002-7827-7096)

Hélène Morlon[2, #] (ORCID : 0000-0002-3195-7521

Benoît Perez-Lamarque[2,3, #] (ORCID : 0000-0001-7112-7197)

[1] ISEM, Univ Montpellier, CNRS, EPHE, IRD, F-34095 Montpellier, France

[2] Institut de biologie de l'École normale supérieure (IBENS), École normale supérieure, CNRS, INSERM, Université PSL, 46 rue d'Ulm, 75 005 Paris, France

[3] Institut de Systématique, Évolution, Biodiversité (ISYEB), Muséum national d'histoire naturelle, CNRS, Sorbonne Université, EPHE, UA, CP39, 57 rue Cuvier 75 005 Paris, France

[*] *Corresponding author: benoit.pichon0@gmail.com*

[#] B.P.L and H.M. jointly supervised this work.

Abstract

Characterizing and understanding the processes that shape the structure of ecological networks, which represent who interacts with whom in a community, has many implications in ecology, evolutionary biology, and conservation. A highly debated question is whether and how the structure of a bipartite ecological network differs between antagonistic (e.g., herbivory) and mutualistic (e.g., pollinators) interaction types. Here, we tackle this question by using a multiscale characterization of network structure, machine learning tools, and a large database of empirical and simulated bipartite networks. Contrary to previous studies focusing on global structural metrics, such as nestedness and modularity, which concluded that antagonistic and mutualistic networks cannot be told apart from only their structure, we find that they can be told apart with a multiscale characterization of their structure. Mesoscale structures, such as motif frequencies, appear particularly informative, with an over-representation of densely connected motifs in antagonistic networks, and of motifs with asymmetrical specialization in mutualistic networks. These characteristics can be used to predict interaction types with relatively good confidence. Our results clarify structural differences between antagonistic and mutualistic networks and highlight machine learning as a promising approach for characterizing interaction types in systems where it is not directly observable.

Keywords: ecological interactions, network structure, motif frequency, interaction classification, machine learning

Introduction

Species in ecosystems engage in multiple types of antagonistic or mutualistic interactions, such as predation, parasitism, pollination, seed dispersal, or mycorrhizal symbioses [1,2]. Such interactions between two groups of potentially taxonomically distant species are often represented by bipartite networks [3–5]. Studying the structure of these interaction networks is fundamental for unraveling the different processes behind their ecological organization [5–8], understanding their evolutionary dynamics, and prioritizing conservation efforts [7,9,10].

Early studies on bipartite ecological networks, focused mainly on pollination and herbivory networks, suggested that mutualistic and antagonistic networks have different structures [3,11,12]. Antagonistic networks tend to have a modular structure, where species within modules preferentially interact with each other and rarely interact with species from other modules [13,14]. Conversely, mutualistic networks tend to have a nested structure, where specialist species mainly interact with generalists, and generalists form a core of interactions [15,16,3]. Such structural differences have been observed in both terrestrial [5] and marine networks [17]. However, several studies have challenged the generality of the modular/nested dichotomy [18–21]. The dichotomy seems to apply only to highly connected networks (*i.e.* with a high proportion of realized interactions), whereas networks with lower connectance are often simultaneously modular and nested [22]. In addition, networks tend to be more modular when their interactions are more intimate, *i.e.* when the degree of biological integration between interacting individuals is high, regardless of interaction type [5,17,23]. Recently, using a large and diverse dataset of empirical networks with different types of ecology (*e.g.* ant-plant, bacteria-phage, host-parasitoid,…), Michalska-Smith and Allesina (2019) showed that the structure of ecological networks is extremely heterogeneous and that global measures of network structure including nestedness and modularity are not sufficient to discriminate antagonistic and mutualistic networks [20]. Song and Saavedra (2020) suggested that

the confounding effect of environmental conditions indeed prevents telling antagonistic and mutualistic networks apart from such global metrics unless environmental differences are accounted for [24].

Macro-scale metrics of network structure, like nestedness and modularity, are not the only way to characterize network structure. The nestedness/modularity paradigm, as the main method to characterize bipartite network structure, has recently been questioned [25]. The authors highlighted the limits of macro-scale metrics for capturing structural differences between bipartite networks and showed the usefulness of motifs (a small subset of species interactions exhibiting varying topologies), which capture the mesoscale structure. Motifs have been informative in various contexts, like the assessments of network stability or resistance to biological invasions [26–28]. Another way to characterize the structure of interaction networks is through the spectral density of the Laplacian graph. Laplacian graphs have been used in other fields, for instance, to compare structural properties of species phylogenies [29] or brain networks [30], but only marginally to study the structure of ecological networks [20]. If such approaches can be informative to identify structural components that differ between antagonistic and mutualistic networks, this will both improve our understanding of the processes involved in the evolution and maintenance of these networks and offer the possibility to predict interaction type in systems where interactions cannot be directly observed. For instance, many widespread fungal endophytes have ecologies that are either unknown or controversial [31,32]. More generally, networks involving microorganisms, such as bacteria in soils [33] and plankton in the oceans [34,35], have often been constructed using presence/absence data rather than direct observations, and the type of a large proportion of these interactions remains undetermined [35,36].

Here, we use a variety of network structural metrics combined with different machine learning tools to (i) investigate if there are structural components of networks,

and which ones, that are distinct in mutualistic *versus* antagonistic networks, (ii) assess whether unsupervised and supervised classification approaches can robustly predict interspecific interaction type from network structure, and (iii) evaluate whether currently available simulation tools for bipartite networks can help improve the performance of the classifiers. We compare the predictive power of the different structural metrics and classification methods.

Results

Structure of antagonistic versus mutualistic networks

We characterized the structure of 343 bipartite ecological networks with well-known interaction types (antagonistic *versus* mutualistic) and ecologies (*e.g.,* pollination, herbivory, seed dispersal) using three different approaches that focus on different network scales (see Methods, Fig. 1; [25]). We first investigated the macro-scale structure of the networks by using global metrics: connectance, nestedness, modularity, network size (*i.e.,* the total number of species), and absolute difference in guild size (*i.e.,* the number of species on each side of the network). We found differences in macro-scale structures between antagonistic and mutualistic networks; however, these were explained by specific ecologies rather than the overall interaction type. Specifically, nestedness was significantly higher in mutualistic networks (mean: 2.16 ± s.d. 0.71) than in antagonistic ones (1.94 ± 1.09; Wilcoxon-Mann-Whitney test: W = 8673, p-value < 0.001), whereas antagonistic networks were significantly more connected (0.23 ± 0.17) than mutualistic ones (0.15 ± 0.09; Wilcoxon-Mann-Whitney test: W = 17709, p-value < 0.001). The difference in terms of modularity was not significant (0.41± 0.15 for antagonistic networks and 0.43 ± 0.10 for mutualistic ones; Wilcoxon-Mann-Whitney test: W = 12559, p-value = 0.19) (Fig. 2A). In addition, we found that mutualistic networks were composed of more species compared to antagonistic ones (76.1± 92.3 for antagonistic networks and 82.7 ± 72.7 for mutualistic ones; Wilcoxon-Mann-Whitney test: W = 11124, p-value = 0.003) and that absolute differences between guild sizes were also more important in mutualistic networks (28.5 ± 46.7 for antagonistic networks and 35.4 ± 50.1 for mutualistic ones; Wilcoxon-Mann-Whitney test: W = 10768, p-value < 0.001). When controlling for the different types of ecology (*e.g.,* pollination, herbivory, seed dispersal...) using mixed models, antagonistic and mutualistic networks were not significantly different in terms of their nestedness (Wald test: $\chi2 = 0.060$, p-value = 0.81), connectance (Wald test: $\chi2 = 0.41$, p-

value = 0.52), modularity (Wald test: $\chi 2$ = 0.42, p-value = 0. 52), the total number of species (Wald test: $\chi 2$ = 0. 29, p-value = 0.59), and absolute difference in the number of species per guild (Wald test: $\chi 2$ = 0. 95, p-value = 0.33). Model comparisons confirmed that differences in terms of global metrics were better explained by the different types of ecology than by the antagonistic/mutualistic type of the interactions (Supplementary Table 1). For instance, compared to other types of ecology, bacteria-phage networks were more connected but generally less nested, while herbivory networks were more nested (Supplementary Fig. 1, Supplementary Table 2 A-C).

Next, we investigated the micro- to the macro-scale structure of the networks using the spectral density of the Laplacian graph, a graph-theoretical multi-scale approach that integrates the whole structural information of a network (see Methods, Supplementary Fig. 2A). Visual inspections of the spectral densities of perfectly modular or perfectly nested networks revealed that modular networks have rather smooth densities with high peaks whereas nested networks have noisier densities with several small peaks (Supplementary Fig. 2B-E). We found that mutualistic networks were enriched in eigenvalues near 1, while antagonistic networks had significantly more spectral eigenvalues near 0 (Fig. 2B), which was mainly driven by herbivory but not host-parasite networks (Supplementary Fig. 3).

Finally, we analyzed patterns of interactions between subsets of species, referred to as motifs (Fig. 1), which represent the mesoscale structure of the networks (see Methods). We found, following the motif denomination of [37], that antagonistic networks were significantly enriched in densely connected motifs (such as motif 12 and motifs 31 to 37), whereas mutualistic networks were enriched in motifs with asymmetrical specialization (such as motifs 10, 19, 20, 25, 28 Fig. 2C). In total, 24 out of the 44 motifs differed significantly in frequency between antagonistic and mutualistic networks ($p<0.01$). When controlling for the type of ecology using mixed models, 4 motifs remained significantly different in frequency between antagonistic and mutualistic networks. These 4 motifs were all asymmetrically specialized; we

observed in particular that motifs 17 and 44 were significantly more frequent in antagonistic networks (Supplementary Fig. 4B). 31 out of 44 motifs differed significantly in frequency across the different types of ecology (Supplementary Fig. 4C): for example, seed-dispersal and host-parasite networks had relatively high frequencies of 4- to 6-species motifs, in particular motifs 30 to 43, which may be linked to their high connectance (Supplementary Fig. 1 & 4, Supplementary Table 2A).

Unsupervised classifications of empirical networks

We tested unsupervised classifications of the mutualistic *versus* antagonistic empirical networks based on the different structural metrics.

A principal component analysis (PCA) on the global metrics did not identify distinct clusters for mutualistic and antagonistic networks (Fig. 3A). An unsupervised K-means classification on the principal component axes detected significant differences, but that was not sufficient to confidently classify the networks (F-scores of 0.63 and 0.55 for mutualistic and antagonistic networks respectively) (Fig. 3A). Similarly, networks with different types of ecology (*e.g.* seed dispersal, herbivory, ant-plant…) were largely mixed on the PCA projection (Supplementary Fig. 5) and did not systematically fell in one of the two clusters of the K-means classification (Supplementary Table 3A).

PCAs performed on spectral density values (Fig. 3B) or motif frequencies (Fig. 3C) were better at separating antagonistic *versus* mutualistic empirical networks than when using the global metrics, but only slightly so (Fig. 3A). Antagonistic networks appeared to be spread on the projection, while mutualistic networks were more clustered, which is potentially linked to the high number of pollination networks in the database (Supplementary Fig. 6-7). Unsupervised K-means classifications on the principal component axes were not sufficient to confidently classify mutualistic and antagonistic networks: mutualistic networks were rather well classified (F-scores=0.76

for the spectral density and 0.81 for the motifs; Fig. 3B-C), but not antagonistic ones (F-scores = 0.55 for spectral density and 0.48 for motifs). Networks with different types of ecology were largely mixed between the two clusters from the K-means classification (Supplementary Tables 3 B-C). Altogether, these results suggest that unsupervised classifications do not perform well at classifying ecological networks, regardless of the metric used to characterize their structure.

Supervised classifications of empirical networks

We then tested different supervised approaches (Logit regressions, Lasso regressions, and artificial neural networks) to classify empirical networks based on their structure. In short, empirical networks were separated between a training set and a test set, and the different classifiers were trained to separate antagonistic and mutualistic networks using global metrics, the Laplacian spectral density, motif frequencies, or any combination of these three different sets of input variables (see Methods).

We found that the two linear classifiers (Logit and Lasso regressions) mainly classified empirical networks as mutualistic, regardless of their interaction type (Supplementary Table 4). Artificial neural networks (ANN) performed much better, in particular when using motif frequencies. Motif frequencies led to better classifications than the spectral density or the global metrics, and adding other variables to motif frequencies did not improve the classifications (Supplementary Table 4). The ANN based on motif frequencies classified an average of 65% of the antagonistic networks (resp. 84% of the mutualistic networks) as antagonistic (resp. mutualistic), resulting in an F-score of 0.68 and 0.82 for antagonistic and mutualistic networks respectively. The other classification methods had lower accuracies; for example, the ANN based on the Laplacian spectral density provided F-scores of 0.59 and 0.77 for antagonistic and mutualistic networks respectively (Supplementary Table 4). This was true irrespective

214 of the bandwidth value used to smooth the Laplacian spectrum into a spectral density
215 (Supplementary Table 5). As the best classification was reached using ANN on motif
216 frequencies, we used this approach for the remaining analyses.

217 We tested the potential effect of overfitting and found that it had a limiting
218 impact on the accuracy of the classifications (Supplementary Fig. 8). The classification
219 was affected by the over-representation of mutualistic networks in our database, but
220 only slightly so: when equalizing the number of networks per interaction type in the
221 training set, antagonistic networks were better classified (from 58 to 68% correct
222 classifications on average) while there were more misclassifications for mutualistic
223 networks (from 78 to 70% correct classifications on average; Supplementary Fig. 9A).

224 Finally, we explored whether the ANN learned singularities of the types of
225 ecology rather than learning the mutualistic/antagonistic interaction type. When we
226 excluded a given type of ecology during the ANN training, the percentage of correct
227 classifications of these networks decreased significantly, but remained quite good
228 (Supplementary Table 7B): for instance, 65% of the seed dispersal networks and 63%
229 of the plant-ant networks were accurately classified when none of these networks were
230 used during training, compared to 79% and 66% respectively when they were included
231 (Supplementary Table 7B; Table 1). Thus, the ANN classifiers not only learned from
232 the antagonistic/mutualistic dichotomy but also partially "learned" the singularities of
233 the networks (*i.e.* data linkage). Accordingly, removing one type of ecology (e.g. seed
234 dispersal networks) from the training set did not significantly decrease the accuracy of
235 our ANN classifier, except when removing pollination or host-parasite networks that
236 represent the majority of the mutualistic and antagonistic networks in our database,
237 respectively (Supplementary Table 7A). Similarly, equalizing the number of networks
238 per type of ecology in the training set did not significantly affect classification accuracy
239 (Supplementary Fig. 9B).

240

Confidence in the supervised classifications of the empirical networks

To overcome the limited size of our database of empirical networks, we introduced a measure of the classification confidence for each empirical network: we replicated 10,000 ANN classifications with different empirical training and test sets (see Methods). When looking at the confidence in the classification of each empirical network, we found that 69% (52% and 17% with high and low confidence respectively; F-score =0.74) of the antagonistic networks and 89% (76% and 13% with high and low confidence respectively; F-score =0.86) of the mutualistic ones were accurately classified in the majority of the ANN classification replicates (Table 1). The percentage of accurately classified networks varied according to the type of ecology. Indeed, most bacteria-phage, seed dispersal, and pollination networks were accurately classified by this approach (*e.g.* 23/29 of the seed dispersal networks were accurately classified, among which ~85% with high confidence). Conversely, high intimacy mutualistic networks (*i.e.* ant-plant and mycorrhizal networks) presented frequent misclassification (only 7 out of 12 networks (~60%) were classified as mutualistic), herbivory networks (which are low intimacy antagonistic networks) were mainly classified as mutualistic (13/23, *i.e.* 57%), and 50% (5/10) of the host-parasitoid networks were also misclassified (Table 1).

Ability of simulated networks to improve the performance of classifiers

The rather small size of our empirical database limits the performance of classifiers. We tested whether simulated networks can help improve these performances. We used simulations of the BipartiteEvol model, a recently developed individual-based eco-evolutionary model that mimics the emergence of mutualistic or antagonistic bipartite interaction networks ([38]; see Methods). Unfortunately, both unsupervised K-means classifications (results not shown) and ANN (Supplementary Table 8B) trained on BipartiteEvol simulated networks were not able to properly

classify the mutualistic *versus* antagonistic empirical networks, no matter the metric used (results not shown). Mutualistic networks, in particular, were often wrongly classified as antagonistic. When classifying simulated BipartiteEvol networks from the test sets, 70% (F-score =0.71) of antagonistic networks and 58% (F-score = 0.57) of mutualistic ones were accurately classified when using motif frequencies and ANN (Supplementary Table 8A), suggesting difficulties in separating antagonistic *versus* mutualistic networks simulated with BipartiteEvol using ANN.

When comparing one-by-one the global metrics of the simulated and the empirical networks, we found networks simulated using BipartiteEvol to be on average significantly more nested and less modular than empirical ones (Supplementary Fig. 10). They also remained slightly less connected (0.11) than empirical networks (0.19) despite the strategy we used to increase their connectance (see Methods). Principal component analyses (PCA) using either global metrics, motifs frequencies, or spectral density values revealed that BipartiteEvol networks occupied a limited space within the space occupied by the empirical networks. This indicates that our simulations are realistic but mimic only a fraction of what empirical networks can look like (Supplementary Figs. 11-13). In addition, BipartiteEvol networks often had significantly higher frequencies of complex motifs (involving 6 species and highly connected) compared with empirical networks that were often enriched with simpler motifs (Supplementary Fig. 14). In line with these results, the Laplacian density of BipartiteEvol networks was characterized by larger eigenvalues than those of empirical networks (Supplementary Fig. 15). Altogether, networks simulated using BipartiteEvol are not as structurally diverse as empirical networks and slightly distinct in their structure, explaining at least partly why ANN trained on BipartiteEvol networks performed badly at classifying empirical networks (Supplementary Table 8B).

Discussion

In this study, we showed that discriminating antagonistic *versus* mutualistic ecological networks based on their structure is overall possible. Supervised machine learning classifiers such as artificial neural networks trained on motif frequencies reached an average of 82% correct classifications (F-score of 0.74 and 0.86 for antagonistic and mutualistic networks respectively). Structural differences between antagonistic and mutualistic networks can be captured at the network mesoscale through differences in some motif frequencies, whereas the macro-scale modular/nested dichotomy does not seem sufficient. Simulations under a recently developed eco-evolutionary model of bipartite interaction networks did not improve the classification of empirical networks, revealing a diversity of structures in the empirical networks that is not exhaustively represented by the model.

Unsupervised classifiers and supervised linear classifiers (Logit and Lasso) failed at accurately classifying antagonistic *versus* mutualistic networks. Supervised artificial neural networks (ANN) performed much better. Using motif frequencies and ANN, we were able to properly classify around 82% of the empirical interaction networks. This percentage of successful classifications is higher than the success rate achieved by the only previous attempt we are aware of to classify networks with ANN based exclusively on their structure [20], and achieved without accounting for environmental conditions (as in 24). Our results suggest that motif frequencies, which depict patterns of species interactions at the mesoscale, are better at capturing structural differences between antagonistic and mutualistic networks compared to global metrics or the Laplacian spectral density, and that these differences are robust to different environmental conditions.

Our analyses confirm that global metrics, including the modular/nested dichotomy, are not sufficient to discriminate between mutualistic and antagonistic

network structures [20,21]. We did not find a significant difference in modularity values between mutualistic and antagonistic networks. We did find that empirical mutualistic networks tend to be more nested and less connected than antagonistic networks, as previously suggested by [3,12], but this dichotomy is not consistent enough (*i.e.,* there are too many exceptions) to provide a robust classification criterion. There are several empirical examples of exceptions to the nested/modular dichotomy in the literature, such as a nested structure in host-parasite antagonistic networks [18]. Within our database, some pollination and mycorrhizal networks are known to be modular [9,39–43], while they assemble mainly mutualistic interactions.

We found that structural differences between mutualistic and antagonistic networks are much clearer when looking at motif frequencies. Densely connected motifs (*e.g.* motifs 31 to 37) are significantly more frequent in antagonistic networks, while mutualistic networks are characterized by motifs with asymmetrical specialization (*e.g.,* motifs 10, 19, 20). Motifs focus on precise structural patterns of interspecific interactions that can be missed when summarizing all the information in a single global value at the network macro-scale [25]. Why motifs performed better than the spectral density of the Laplacian graph is unclear, as this latter representation is supposed to integrate all scales. Using a constant bandwidth value to convert the Laplacian graph spectrum into a density may result in information loss when comparing networks of different sizes. Also, transforming the discrete Laplacian spectrum into a continuous density may be problematic for small networks with only a few eigenvalues.

Mutualistic and antagonistic networks cover a wide range of ecologies (*e.g.* pollination, parasitism, herbivory) and our analyses showed that the type of ecology matters in the supervised classification, a phenomenon referred to as data linkage: networks with different ecologies tend to have their own structural signatures beyond the mutualistic/antagonistic dichotomy, and the ANN classifier at least partially learns to identify each type of ecology rather than to discriminate mutualistic *versus*

351 antagonistic networks. In fact, after controlling for the type of ecology, the only
352 significant difference between mutualistic and antagonistic networks was the
353 frequency of a few motifs with asymmetrical specialization (*e.g.*, motifs 17 and 44),
354 suggesting that a significant part of the structural singularities comes from the type of
355 ecology rather than the type of interaction (antagonistic *versus* mutualistic). These
356 results on the importance of the type of ecology corroborate previous studies that
357 found, for instance, significant structural differences between pollination and seed
358 dispersal networks [44] or between mutualistic plant-fungus and plant-animal
359 networks [45]. Interestingly, we found herbivory and pollination networks to have
360 similar structural properties in both spectral density and motifs frequency, which may
361 explain the low classification rate of herbivory networks. In terms of the ability to
362 classify mutualistic and antagonistic networks, our results on ecology suggest that a
363 better classification of networks of a particular type of ecology will be achieved by
364 using networks from the same ecology during the training of the classifiers.

365 The efficiency of supervised classification approaches strongly depends on the
366 size and quality of the dataset used for training. We used one of the largest publicly
367 available databases of empirical networks, but it contained only 343 networks in total.
368 In addition, it had a strong unbalanced representation of the different types of ecology,
369 with 80% of the mutualistic networks represented by pollination networks and 60% of
370 the antagonistic networks represented by host-parasite networks. There was also a
371 strong imbalance in the representation of interaction intimacy, which is known to
372 impact network structure [5,17,23,46]: mutualistic networks were mainly low intimacy
373 networks, except ant-plants (n=3) or mycorrhizal networks (n=9), and antagonistic
374 networks were mainly high intimacy networks, except herbivory networks (n=23). As
375 a consequence, only ~60% of the high-intimacy mutualistic networks and ~40% of the
376 low-intimacy antagonistic networks were correctly classified. These results confirm
377 that interaction intimacy may confer a particular structure to ecological networks
378 [5,17]. Increasing the number of empirical networks, especially mutualistic networks

with high intimacy (*e.g.* symbiotic networks) and antagonistic networks with low intimacy (*e.g.* trophic networks) would improve further our ability to classify ecological networks by better representing the large structural heterogeneity of ecological networks [20]. Another way to improve the classification would be to consider interaction weights (the abundance of interacting individuals in a pair of interacting species). Such information can emphasize structural differences that are not contained in binary networks [47,48]. Unfortunately, the database of weighted empirical networks is even smaller than that of binary networks.

The efficiency of supervised classification approaches can also strongly depend on the type, architecture, and tuning of the classifier. For example, the overfitting we noticed in the ANN (which did not seem to negatively impact our results) could potentially be avoided by using other regularization technics, such as dropout [49]. We chose to use simple, few-layers neural networks implemented in R for their availability and ease of usage to the large community of R users, as well as because our dataset for training was rather small. However, several other machine learning tools exist and could be tested, although they may require larger training sets; graph neural networks in particular have been specifically designed to learn information from graphs such as ecological networks [50].

One approach to better train the ANNs would be to increase the size of the training data thanks to simulated data. We investigated this possibility with BipartiteEvol, an individual-based model that mimics the eco-evolutionary emergence of bipartite interaction networks [38]. However, ANN classifiers trained on BipartiteEvol simulations were not able to discriminate mutualistic and antagonistic empirical networks, suggesting that the model somehow fails at capturing important structural differences between antagonistic and mutualistic networks. [38] showed that the model produces realistic networks, which we confirm here, and is able to reproduce major global differences observed between mutualistic and antagonistic networks. However, we find that BipartiteEvol simulations tend to have higher

407 frequencies of complex motifs and a lower connectance compared to empirical
408 networks. Consequently, their projection only covers a small proportion of the space
409 occupied by empirical networks. One possible explanation is the continuum between
410 mutualism and antagonism in many interactions found in natural communities [51].
411 For instance, it has been shown that parrot-plant networks can combine the two types
412 of interactions [51]. Similarly, pollination or mycorrhizal networks often contain
413 cheating species that use antagonistic strategies and induce structural changes within
414 the interaction networks [52–54]. Interactions can also shift from mutualism to
415 antagonism on short timescales, depending on some external conditions [55,56]. By
416 contrast, BipartiteEvol models the emergence of a network in strict antagonistic or
417 mutualistic communities. This calls for future efforts to improve the BipartiteEvol
418 model. Simulations under other generative models (*e.g.*, 57,58) or data augmentation
419 methods could also be tested.

420

421 Ultimately, supervised classifying approaches could be particularly useful to
422 predict the type of undetermined interactions. For instance, many plant-fungus
423 interaction networks are assembled from high-throughput sequencing data, but we
424 currently ignore the mutualistic *versus* the antagonistic type of many of these
425 endophytic interactions [31,32]. Similarly, an increasing number of networks are built
426 from co-occurrence data generated with high-throughput sequencings, such as the
427 ones between microscopic planktonic species sampled by the *Tara Oceans* expeditions
428 [34,35]. The ecology of the majority of species in these networks is currently
429 undetermined [35,36,59]. However, a preliminary analysis of bipartite networks built
430 from co-occurrence data in the *Tara Oceans* data revealed that these networks were
431 structurally very different from the empirical and simulated networks we used here
432 (results not shown). Classifying these networks would thus require using a different
433 training dataset. A potential source of difference might come from species delineation
434 in microbial lineages in such databases, which rely on the clustering into operational

taxonomic units (OTUs) of short marker gene sequences (e.g. the small subunit rRNA gene or the ITS marker). This difference in species delimitation in networks involving microbial groups compared to networks involving macro-organisms (*e.g.* pollination and herbivory) [19] may affect network structure. This may explain why the classification of mycorrhizal networks was not very accurate in our study. One approach to tackle this would be to incorporate more microbial networks with known interaction types in the training of the supervised classifier, which would be highly recommended anyway given the importance of the type of ecology in the classification accuracy. Another source of differences comes from the use of co-occurrence as a proxy for interaction, a topic that is currently intensively debated [60]. Classifiers would need to be trained with co-occurrence networks with known interaction types or simulated co-occurrence networks.

Conclusion

Mutualistic and antagonistic networks have structural differences that can be particularly well captured by motif frequencies. Antagonistic networks are enriched in densely connected motifs, while mutualistic networks are enriched in motifs with asymmetrical specialization. Besides this dichotomy, there are structural differences across networks with distinct ecologies. These structural differences can be used to classify bipartite interaction networks with supervised machine learning with pretty good accuracy. These performances could be further improved by increasing the number of empirical networks used for training, diversifying the type of ecology sampled, and improving currently available models to simulate networks. While more work is needed to predict interaction types for co-occurrence networks, our analyses provide promising results on the ability to predict interaction types from network structure.

Materials and methods

Empirical interaction networks

We analyzed 343 local or regional scales bipartite interaction networks with well-known interaction types. We obtained these networks by downloading the networks from [20], among which some come from the Web of Life database (web-of-life.es [61]). We kept networks containing at least 10 species in each guild, as the structure of small networks is difficult to characterize [9], in particular in terms of motif composition (preliminary results not shown). In addition, we gathered 9 mycorrhizal networks that also fit this size criterion: 6 were collected from individual papers [39,41,42,62–64] and 3 from the database generated by [65]. Among the 343 networks, 216 were mutualistic and 127 antagonistic. They comprised 9 different types of ecology: bacteria-phage (antagonistic; n=17), seed dispersal (mutualistic; n=29), host-parasite (antagonistic; n=77), host-parasitoid (antagonistic; n=10), plant-mycorrhiza (mutualistic; n=9), ant-plant (mutualistic; n=3), herbivory (antagonistic; n=23), anemone-fish (mutualistic; n=1), and pollination (mutualistic; n=174). Each network was converted into a binary matrix, taking the value 1 if species *i* interacts with species *j*, and 0 otherwise, *i.e.* we only focused on the presence/absence of interactions, not on their frequencies. The guild with the largest (resp. lowest) number of species was in rows (resp. columns).

Comparing the structure of mutualistic versus antagonistic networks

We characterized the structure of ecological networks in three ways focusing on different network scales (Fig. 1).

First, we characterized the macro-scale structure of the networks using 5 global metrics: the total size (number of species in both guilds), the absolute difference in the number of species in the two guilds, the connectance, the nestedness, and the

488 modularity. We quantified nestedness using the NODFc index that corrects for
489 network size ("NODFc" function from the R-package maxnodf, version 1.0.0 [66]). We
490 quantified modularity using the function *computeModules* from the R-package
491 bipartite (version 2.16) [67] with the "Beckett" method.

492 Second, we characterized the mesoscale structure of the networks by performing
493 bipartite motif analyses that depict the patterns of interactions between subsets of
494 species (called motifs; Fig. 1) [37]. For each network, we computed the frequencies of
495 the 44 possible motifs involving between 2 and 6 species using the function *mcount*
496 from the bmotif R-package (version 2.0.2) [68] with the normalization procedure
497 "normalise_sum" (*i.e.* which computes the frequency of motifs among the total
498 number of motifs).

499 Third, we used a multi-scale approach stemming from graph theory that integrates
500 the whole structural information of a network, from micro- to macro-scales: the
501 spectral density of the Laplacian graph [69]. We computed the normalized Laplacian
502 graph of each network [69], defined as $L_{norm} = D^{-1/2} L D^{-1/2}$, where D is the degree matrix
503 (a diagonal matrix counting the number of interactions for each species) and L is the
504 Laplacian graph, defined as the difference between the adjacency matrix (a square
505 binary matrix whose elements represent the interactions between species of the
506 network) and the degree matrix (Supplementary Fig. 2A). The normalized Laplacian
507 graph is a symmetric and positive-definite matrix whose eigenvalues are constrained
508 between 0 and 2 [70], which makes them comparable between networks of different
509 sizes, and symmetrical around 1. There are as many eigenvalues equal to 0 (or 2) as
510 there are connected components (*i.e.* perfect modules of species that are not connected
511 with any species from other modules), and we can thus expect a modular network to
512 have more eigenvalues close to 0 (and 2) than a nested network. To be able to efficiently
513 compare Laplacian graphs across networks of different sizes, we converted the
514 eigenvalues into a spectral density using a Gaussian convolution [29](Supplementary
515 Fig. 2A). To do so, we adapted the *spectR* function from the RPANDA R-package

516 (version 1.9) [71] for bipartite networks and used a bandwidth of 0.025 for the Gaussian
517 convolution. We tested the influence of the bandwidth value (from 0.001 to 0.091) on
518 the resulting spectral densities. We characterized each network by extracting its
519 spectral density values at 100 points uniformly distributed between 0 and 1.

520 In sum, we characterized the structure of each network with 149 variables: 5 global
521 metrics, 44 motif frequencies, and 100 values extracted from its Laplacian spectral
522 density. For each metric of network structure, we used a separate Wilcoxon-Mann-
523 Whitney test to assess whether antagonistic *versus* mutualistic networks were
524 significantly different according to that metric. In addition, we accounted for the
525 different types of ecology by performing mixed-effects models (function *lmer* from the
526 R-package lme4 (version 1.1-27)) with the type of interaction (*i.e.* mutualistic or
527 antagonistic) as a fixed effect and the type of ecology as a random effect (Type II Wald
528 test). We also investigated which metric significantly differed between each type of
529 ecology by performing an ANOVA for each metric (Fisher test). Finally, we analyzed
530 whether the differences in structural metrics were better explained by the type of
531 ecology or the type of interaction by comparing the Akaike information criterion (AIC)
532 of the different models (see Supplementary Table 1).

533

534 **Unsupervised classification**

535 To visualize structural differences across networks, we used principal
536 component analyses (PCA), with networks characterized either in terms of their global
537 metrics, motif frequencies, or Laplacian spectral densities. PCAs were performed with
538 the functions *PCA* and *fviz_pca_ind* from the FactoMineR and factoextra R-packages
539 (versions 2.4 and 1.0.7 respectively) [72].

540 For each type of characterization, we then used unsupervised K-means
541 classification on the two main principal components projections (see Fig. 3,
542 Supplementary Tables 3). K-means were performed with the function *kmeans* with 2

clusters. We investigated whether the two clusters tend to separate mutualistic from antagonistic networks, and networks with different types of ecology, by performing chi-squared tests on the K-means classifications. To quantify the ability to separate antagonistic and mutualistic networks, we used a measure of classification accuracy computed as the F-score with the function *F1_score* from the R-package MLmetrics. The F-score is the harmonic mean of the precision (*e.g.* the frequency of true mutualistic among predicted mutualistic networks) and the recall (*e.g.* the frequency of predicted mutualistic among true mutualistic networks).

Supervised classification

Next, we used supervised classifiers to classify antagonistic *versus* mutualistic networks based on their structure. As input for the supervised classifier, for each interaction network, we used a set of structural variables, either its global metrics, its motif frequencies, its Laplacian spectral density, or any combination of these sets of variables. As an output, the supervised classifier provides an interaction type: "mutualistic" or "antagonistic". Supervised classifiers are based on supervised learning: the classifiers have first to be trained with a random subset of the interaction networks (the "training set") in order to optimize its internal parameters; once trained, it can classify the other networks (the "test set") from the input variables. The percentage of networks from the test set that are accurately classified by the classifier indicates the performance of the approach.

We first investigated the performance of linear classifiers using logit and Lasso regressions. In Lasso regression, a constraint on the estimated model coefficients can make the regression with a high number of variables more reliable. Logit regressions were computed with the *glm* function, whereas the Lasso regressions were performed with the function *cv.glmnet* from the glmnet R-package (version 4.1-2) [73]. We

569 included 80% of the networks in the training set and replicated the classifications 50
570 times (with different training sets).

571 Second, we used artificial neural networks (ANN) implemented in the *neuralnet*
572 function from the R-package neuralnet (version 1.44.2) [74]. We optimized their
573 configuration (number of layers and number of neurons in each layer) according to the
574 input variables. We selected a 3 layers network, constituted of the input and output
575 layers and an intermediate layer with 5, 20, and 60 neurons for global metrics, motif
576 frequencies, and spectral density inputs respectively. Preliminary analyses showed
577 that these choices provided a good compromise between classification accuracy and
578 overfitting (results not shown). We also selected a non-linear logistic transformation,
579 and the maximal number of steps was fixed at 10^8. In addition, we found that training
580 the ANN (*i.e.* optimizing neuron weights to reduce classification bias) with 80% of the
581 interaction networks and testing it on the remaining 20% provided a good trade-off
582 between the size of the test set and the accuracy of the classification.

583 We independently trained and tested the classifiers on empirical networks
584 using the different sets of structural input variables. For each configuration, the ANN
585 trainings were replicated 50 times with different training and test sets, and we reported
586 the average percentages of correct classifications among antagonistic and mutualistic
587 networks and their standard deviations. By comparing the performance of each
588 classifier (percentage of correct classification and classification accuracy), we were
589 therefore able to select the best set of input structural variables and the best
590 classification approach to discriminate antagonistic and mutualistic networks.

591 Moreover, because the training sets were limited in size (we only had a total of
592 343 empirical networks), we introduced a measure of the classification confidence for
593 each empirical network. We selected the most efficient classification approach (ANN
594 on motif frequencies; see Results) and we trained and tested it for 10,000 independent
595 training and test sets of empirical networks: each empirical network was therefore
596 classified on average 2,000 times with different training sets. We then computed the

percentage of classifications that fell in each category, which gives an idea of classification confidence: we considered that a given network was classified as mutualistic (resp. antagonistic) with "high confidence" if it was classified as mutualistic (resp. antagonistic) in more than 75% of the classifications. Conversely, a given network was classified as mutualistic (resp. antagonistic) with "low confidence" if it was classified as mutualistic (resp. antagonistic) in more than 50% of the classifications but in less than 75% of them.

Next, we verified that our results were not impacted by the over-representation of mutualistic *versus* antagonistic networks in the training set (more than 60% of the networks were mutualistic). To do so, we trained the classifier with the same number of antagonistic and mutualistic networks (from 5 to 95 networks for each type of interaction) and compared the classification accuracy with the one obtained with the same-size training sets composed of randomly chosen networks. In each condition, we replicated the classification 50 times with different training sets.

Finally, we investigated if the classifiers were strictly learning to separate antagonistic from mutualistic networks or if they also learned the potential structural singularities of each type of ecology from our database. To do so, we removed one by one each type of ecology from the training test (*e.g.* pollination networks), we trained the classifiers on 50 independent training sets constituted by 80% of the remaining networks, and we (i) tested the classification on the 20% remaining networks (*e.g.* are other networks still as well classified as when pollination networks were not used in the training?) and (ii) tested the classification on the networks from the removed type of ecology (*e.g.* are pollination networks classified among mutualistic networks when no pollination network have been used in the training?). If the classifiers learn to separate antagonistic from mutualistic networks regardless of ecology, the classifications of tests (i) and (ii) should perform as well as the original classification. If they learn to classify ecologies rather than the type of interaction, then test (ii) should perform very badly, and test (i) may be less accurate. Lastly, we also verified whether

our results were impacted by the over or under-representation of some types of ecology in the training set (e.g. the pollination networks represent ~50% of the complete database). To do so, we trained the classifier with the same number of networks per type of ecology (from 5 to 20 networks for each type of ecology; if a type of ecology had fewer networks than the number per type of ecology, all networks were included in the training set) and compared the classification accuracy with the one obtained with same-size training sets composed of randomly chosen networks (see Supplementary Fig. 8B). If the classifiers learn to separate antagonistic from mutualistic networks regardless of ecology, then an over- or under-representation of some ecologies should not impact the results.

Simulations of bipartite interaction networks

The accuracy of supervised classifiers often depends on the size of the available training set (i.e., data for which the classification is known), which in the case of bipartite interaction networks is rather small (e.g. we were able to obtain 343 networks here). A possible approach to increase the training set is to use simulated data. We explored this possibility with the recently-developed BipartiteEvol model [38], an eco-evolutionary model that simulates the evolution of interacting individuals within two clades. We simulated 1,200 weighted networks using a wide range of antagonistic and mutualistic parameters (see [75], for details). We noticed that connectance values for these networks were on average lower than those of the empirical networks; to approach those of the empirical networks, we retained 70% randomly sampled individual interactions in the simulated networks, which had a direct effect of deleting rare, specialized species and increasing network connectance. We also carried out analyses with the original networks (no sampling), which did not qualitatively affect the results (results not shown). Retaining networks that had at least 10 species in each guild, we obtained a total of 1,175 binary networks, including 705 mutualistic and 470 antagonistic networks, hereafter referred to as BipartiteEvol networks.

We performed unsupervised classification using PCA and K-means. We also performed supervised classification: we independently trained and tested the supervised classifiers (Logit or Lasso regressions, or ANN) on BipartiteEvol networks using the different sets of structural input variables. We then tested the classifier on the empirical networks. As the ANN trained on BipartiteEvol simulations performed poorly to classify empirical networks, we investigated differences between networks simulated with BipartiteEvol and empirical networks by comparing the distributions of their global metrics (in terms of network size, connectance, nestedness, and modularity), their motifs frequencies, and the spectral densities of their Laplacian graph.

Acknowledgments

We thank I. Overcast, J. Voznica, L. Aristide, S. Lambert, I. Quintero, C. Fruciano, J. Clavel, A. Silva, for comments on the first version of the manuscript. We also thank S. Chaffron, C. Bowler, C. Nef, and E. Thébault for fruitful discussions.

The authors declare no conflict of interest.

Author contributions

Conceptualisation: Benoît Pichon, Rémy Le Goff, Hélène Morlon, and Benoît Perez-Lamarque

Software & Investigation : Benoît Pichon

Supervision : Hélène Morlon et Benoît Perez-Lamarque

Visualisation: Benoît Pichon, Rémy Le Goff, Hélène Morlon, and Benoît Perez-Lamarque

Writing – original draft: Benoît Pichon, Hélène Morlon, and Benoît Perez-Lamarque

Writing – review & editing: Rémy Le Goff

Data availability

The code used for the analyses is available on Zenodo: https://doi.org/10.5281/zenodo.7786771

Tables:

Table 1. The type of interaction of the empirical ecological networks can be predicted based on their motif frequencies using artificial neural network classifiers.

The following results were obtained by training the artificial neural networks (ANN) with 10,000 different training and test sets, using motif frequencies as the input structural variables. From those 10,000 classifications, we calculated the percentage of antagonistic *versus* mutualistic classifications for each network: a network was predicted as a mutualistic (resp. antagonistic) if it was classified as a mutualistic (resp. antagonistic) network in more than 75% (high confidence) or less than 75% (low confidence) of the classifications. We therefore obtained a predicted type of interaction for each of the 343 empirical networks.

The second column indicates the number of interaction networks in our database gathered for each ecology of interaction. Then, we indicate the number of networks classified as either mutualistic or antagonistic for each ecology of interaction and at the bottom, we show the global contingency table for all mutualistic and antagonistic networks. We also indicate the F-score associated with the global contingency table for each antagonistic or mutualistic type of interaction.

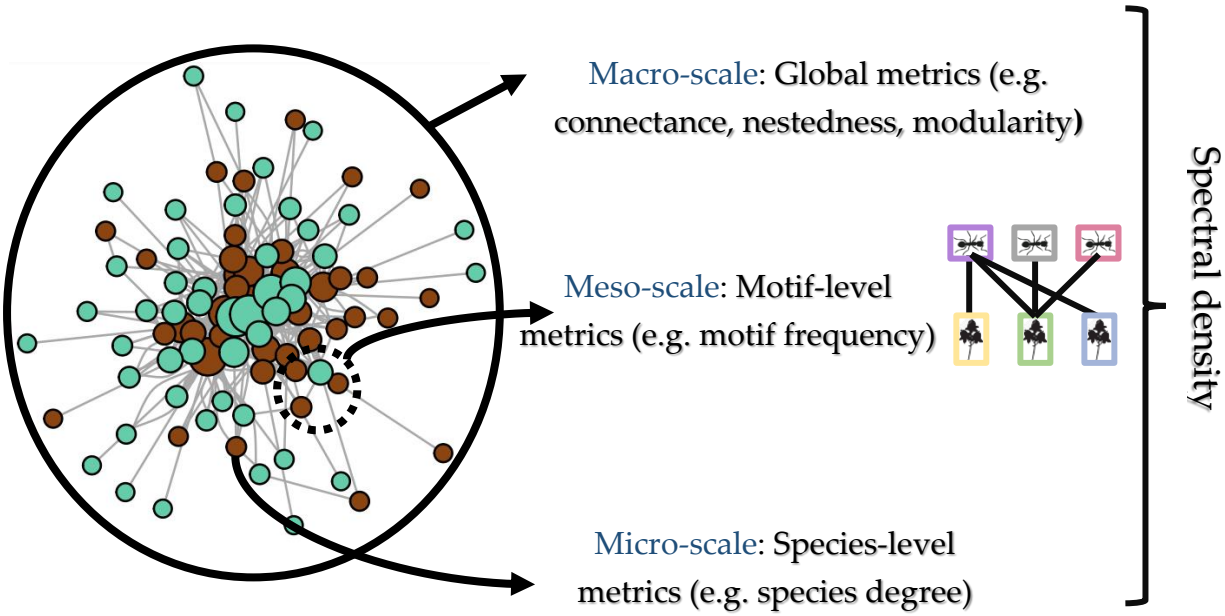
Network ecology	Number of networks	Prediction			
		Antagonistic network (high confidence)	Antagonistic network (low confidence)	Mutualistic network (high confidence)	Mutualistic network (low confidence)
Anemone-Fish	1	0	0	1	0
Bacteria-Phage	17	12	1	3	1
Herbivory	23	4	6	9	4
Host-Parasitism	77	46	14	10	7
Host-Parasitoid	10	4	1	4	1
Mycorrhiza	9	3	1	3	2
Plant-Ant	3	0	1	1	1
Pollination	174	7	7	139	21
Seed dispersal	29	6	0	20	3
Antagonistic network	127	**52%**	**17%**	**21%**	**10%**
Mutualistic network	216	**7%**	**4%**	**76%**	**13%**
F-scores		**0.74**		**0.86**	

701

Figures:

Figure 1: Multi-scale structure of bipartite networks

The structure of bipartite ecological networks, represented here by an ant-plant network (Blüthgen et al., 2004) can be studied at the macro-, meso-, and micro-scale. The Laplacian spectral density corresponds to a multi-scale approach from the macro to the micro-scale, whereas global metrics and motif frequencies focus on macro- and meso-scales, respectively. Blue nodes represent plant species, brown nodes ant species, and grey edges interspecific interactions.



723 **Figure 2. Differences between antagonistic and mutualistic networks across scales of**
724 **network structure.**

725 (A) Connectance, nestedness (NODFc), and modularity (Newman's index) of antagonistic and
726 mutualistic empirical networks. Stars indicate the statistical significance of the difference
727 between antagonistic and mutualistic networks, as measured by a Wilcoxon-Mann-Whitney
728 test (*** = p-value < 0.01, ns = non-significant).

729 (B) Densities of eigenvalues from the Laplacian spectral density that differ significantly
730 between antagonistic and mutualistic networks (Wilcoxon-Mann-Whitney test: p-value<0.01).
731 The discontinuities in the x-axis are due to the filtering of eigenvalues densities that differ
732 significantly between antagonistic and mutualistic networks.

733 (C) Frequencies of the motifs that differ significantly between antagonistic and mutualistic
734 networks (Wilcoxon-Mann-Whitney test: p-value<0.01). Each motif is illustrated at the top of
735 each boxplot (following the denomination of Simmons *et al.* 2018).

736 Mutualistic networks are shown in green and antagonistic ones in yellow. Boxplots present
737 the median surrounded by the first and third quartiles, and whiskers extend to the extreme
738 values but no further than 1.5 of the interquartile range.

739

740

741

742

743

744

745

746

747

748

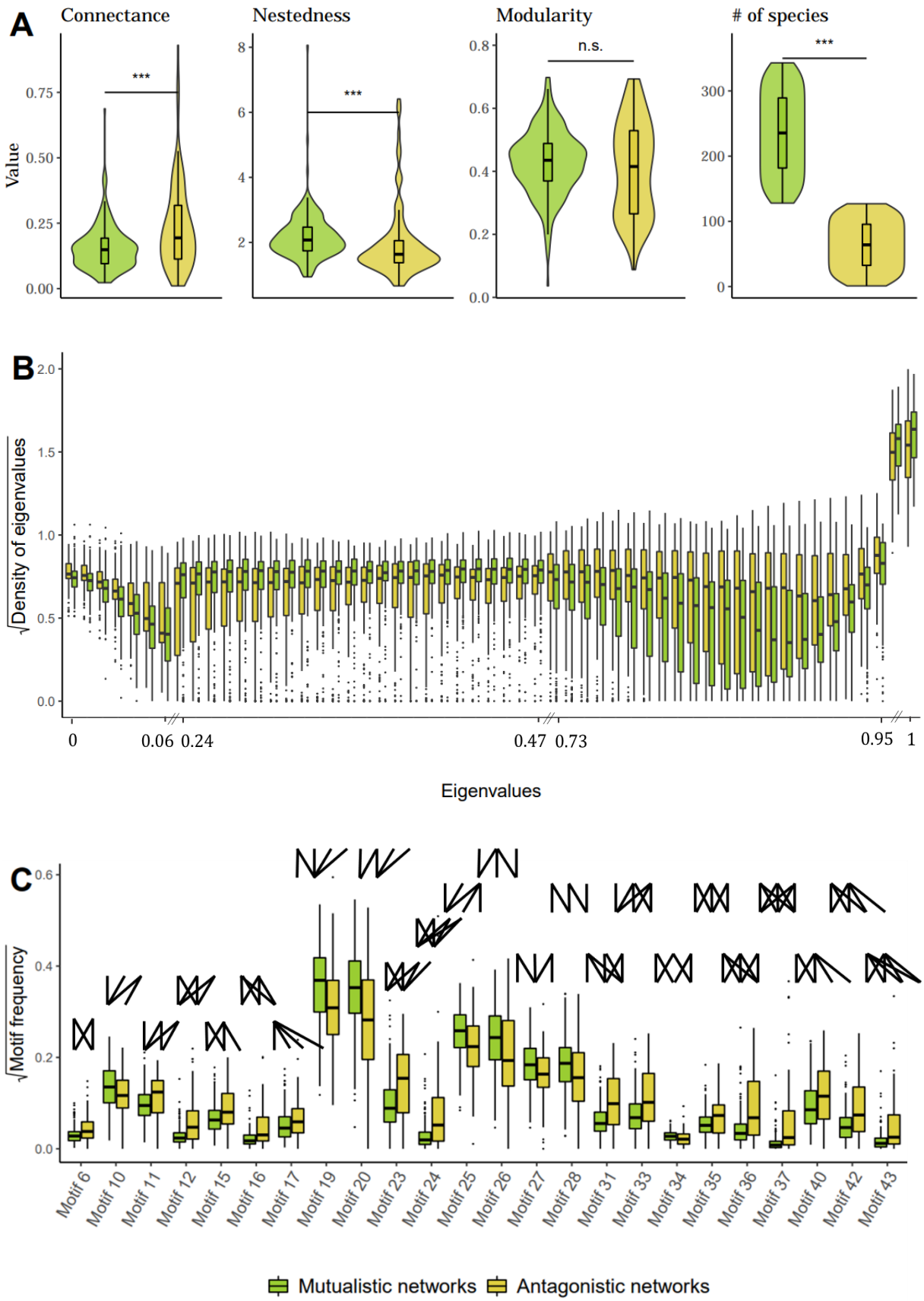
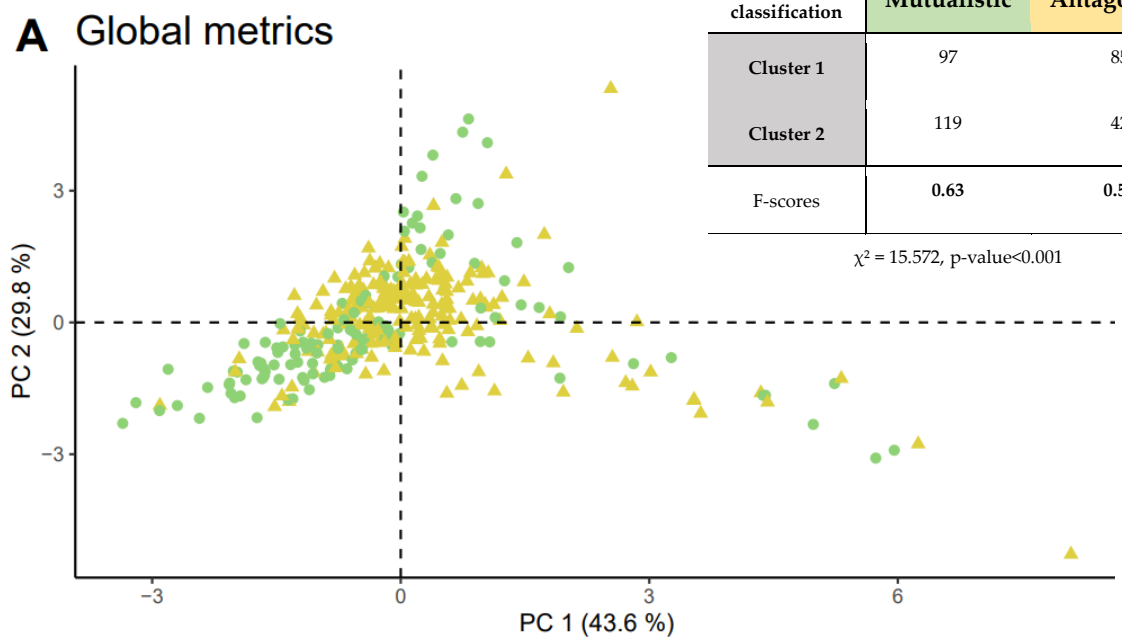

A
Connectance
Nestedness
Modularity
# of species
Value
***
n.s.
B
√Density of eigenvalues
Eigenvalues
0
0.06
0.24
0.47
0.73
0.95
1
C
√Motif frequency
Motif 6
Motif 10
Motif 11
Motif 12
Motif 15
Motif 16
Motif 17
Motif 19
Motif 20
Motif 23
Motif 24
Motif 25
Motif 26
Motif 27
Motif 28
Motif 31
Motif 33
Motif 34
Motif 35
Motif 36
Motif 37
Motif 40
Motif 42
Motif 43
Mutualistic networks
Antagonistic networks

Figure 3. Principal coordinate analysis performed on structural metrics at different scales tend to discriminate mutualistic *versus* antagonistic networks at the mesoscale.

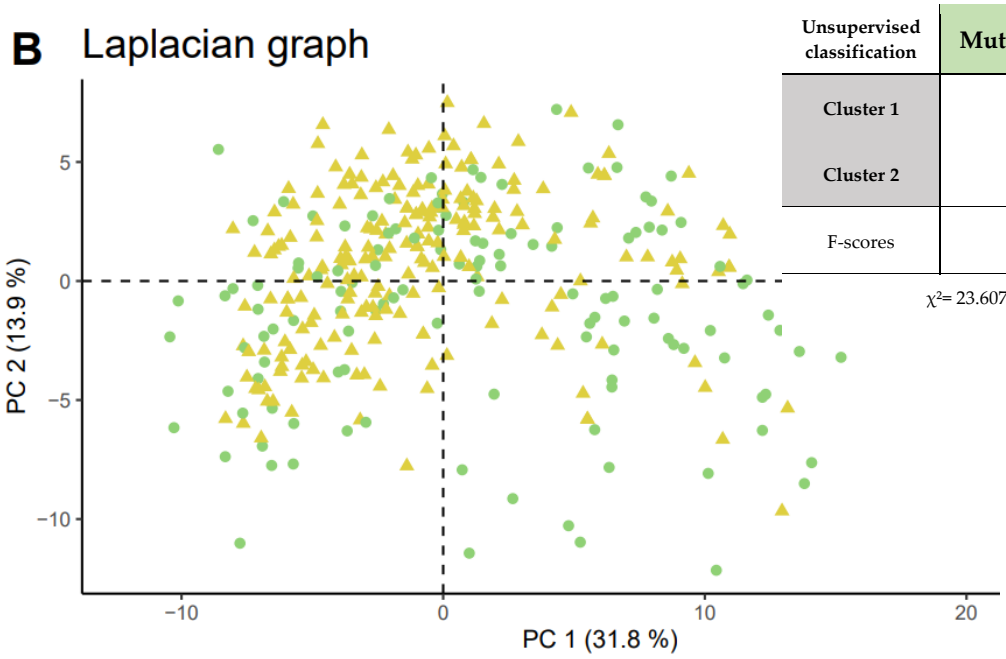
Projection of the 343 empirical networks on the two principal components (PC1 and PC2) obtained using principal coordinate analysis (PCA) on (A) global metrics (*i.e.* nestedness, modularity, connectance, and network size), (B) spectral density of the Laplacian graph and (C) motif frequencies. The percentage of explained variance is indicated on each axis.

Each associated table corresponds to the contingency table of unsupervised K-means classification based on the global metrics (A), motifs frequencies (B), and spectral density of the Laplacian graph (C). Chi-squared test associated with the K-means classification is indicated.



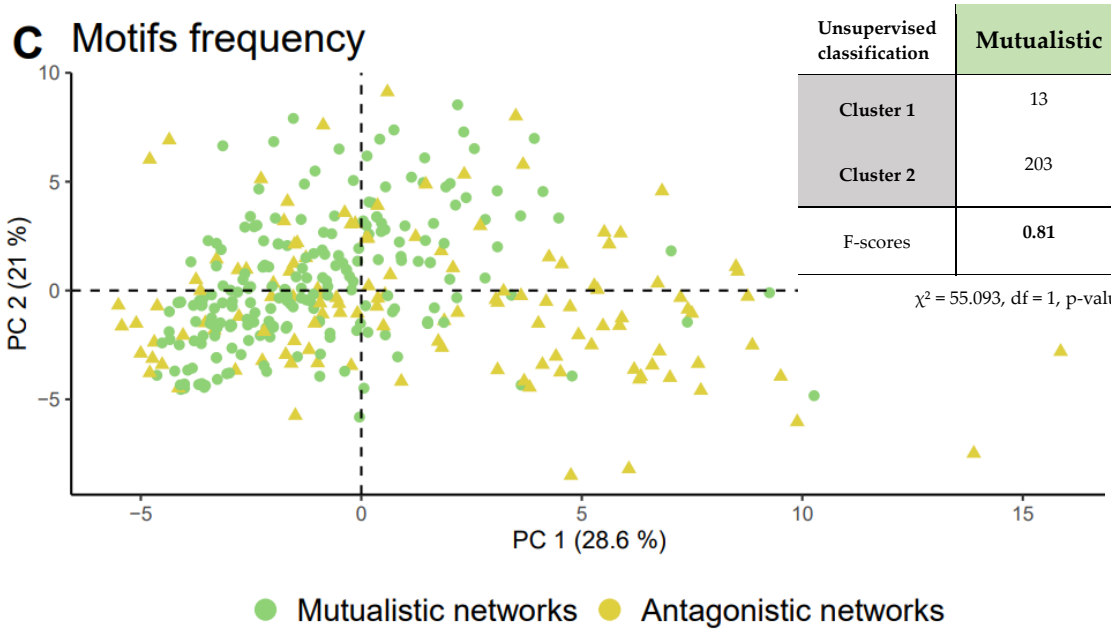
Unsupervised classification	Mutualistic	Antagonistic
Cluster 1	97	85
Cluster 2	119	42
F-scores	**0.63**	**0.55**

$\chi^2 = 15.572$, p-value<0.001



Unsupervised classification	Mutualistic	Antagonistic
Cluster 1	37	52
Cluster 2	179	75
F-scores	**0.76**	**0.48**

$\chi^2 = 23.607$, p-value < 0.001



Unsupervised classification	Mutualistic	Antagonistic
Cluster 1	13	49
Cluster 2	203	78
F-scores	**0.81**	**0.51**

$\chi^2 = 55.093$, df = 1, p-value <0.001

References

1. Thomas F, Renaud F, Guegan J-F, editors. Parasitism and Ecosystems. Oxford: Oxford University Press; 2005. doi:10.1093/acprof:oso/9780198529873.001.0001

2. Bronstein JL, editor. Mutualism. Oxford: Oxford University Press; 2015. doi:10.1093/acprof:oso/9780199675654.001.0001

3. Bascompte J, Jordano P, Melian CJ, Olesen JM. The nested assembly of plant-animal mutualistic networks. Proceedings of the National Academy of Sciences. 2003;100: 9383–9387. doi:10.1073/pnas.1633576100

4. Bascompte J, Jordano P. Plant-animal mutualistic networks: the architecture of biodiversity. Annu Rev Ecol Evol Syst. 2007;38: 567–593. doi:10.1146/annurev.ecolsys.38.091206.095818

5. Fontaine C, Guimarães PR, Kéfi S, Loeuille N, Memmott J, van der Putten WH, et al. The ecological and evolutionary implications of merging different types of networks: Merging networks with different interaction types. Ecology Letters. 2011;14: 1170–1181. doi:10.1111/j.1461-0248.2011.01688.x

6. Thompson JN. The Geographic Mosaic of Coevolution. University of Chicago Press; 2005. doi:10.7208/9780226118697

7. Bastolla U, Fortuna MA, Pascual-García A, Ferrera A, Luque B, Bascompte J. The architecture of mutualistic networks minimizes competition and increases biodiversity. Nature. 2009;458: 1018–1020. doi:10.1038/nature07950

8. Suweis S, Simini F, Banavar JR, Maritan A. Emergence of structural and dynamical properties of ecological mutualistic networks. Nature. 2013;500: 449–452. doi:10.1038/nature12438

9. Olesen JM, Bascompte J, Dupont YL, Jordano P. The modularity of pollination networks. Proceedings of the National Academy of Sciences. 2007;104: 19891–19896. doi:10.1073/pnas.0706375104

10. Rezende EL, Lavabre JE, Guimarães PR, Jordano P, Bascompte J. Non-random coextinctions in phylogenetically structured mutualistic networks. Nature. 2007;448: 925–928. doi:10.1038/nature05956

11. Thébault E, Fontaine C. Does asymmetric specialization differ between mutualistic and trophic networks? Oikos. 2008;117: 555–563. doi:10.1111/j.0030-1299.2008.16485.x

12. Thébault E, Fontaine C. Stability of ecological communities and the architecture of mutualistic and trophic networks. Science. 2010;329: 853–856. doi:10.1126/science.1188321

13. Krause AE, Frank KA, Mason DM, Ulanowicz RE, Taylor WW. Compartments revealed in food-web structure. Nature. 2003;426: 282–285. doi:10.1038/nature02115

14. Newman MEJ. Modularity and community structure in networks. Proceedings of the National Academy of Sciences. 2006;103: 8577–8582. doi:10.1073/pnas.0601602103

15. Jordano P, Bascompte J, Olesen JM. Invariant properties in coevolutionary networks of plant-animal interactions: invariant properties in coevolutionary networks. Ecology Letters. 2002;6: 69–81. doi:10.1046/j.1461-0248.2003.00403.x

16. Rohr RP, Saavedra S, Bascompte J. On the structural stability of mutualistic systems. Science. 2014;345: 1253497–1253497. doi:10.1126/science.1253497

17. Guimarães PR, Rico-Gray V, Oliveira PS, Izzo TJ, dos Reis SF, Thompson JN. Interaction intimacy affects structure and coevolutionary dynamics in mutualistic networks. Current Biology. 2007;17: 1797–1803. doi:10.1016/j.cub.2007.09.059

18. Pilosof S, Fortuna MA, Cosson J-F, Galan M, Kittipong C, Ribas A, et al. Host–parasite network structure is associated with community-level immunogenetic diversity. Nat Commun. 2014;5: 5172. doi:10.1038/ncomms6172

19. Chagnon P-L. Seeing networks for what they are in mycorrhizal ecology. Fungal Ecology. 2016;24: 148–154. doi:10.1016/j.funeco.2016.05.004

20. Michalska-Smith MJ, Allesina S. Telling ecological networks apart by their structure: A computational challenge. Bollenbach T, editor. PLoS Comput Biol. 2019;15: e1007076. doi:10.1371/journal.pcbi.1007076

21. Fontoura L, Cantor M, Longo GO, Bender MG, Bonaldo RM, Floeter SR. The macroecology of reef fish agonistic behaviour. Ecography. 2020; ecog.05079. doi:10.1111/ecog.05079

22. Fortuna M, Stouffer D, Olesen J, Jordano P, Mouillot D, Krasnov B, et al. Nestedness versus modularity in ecological networks: Two sides of the same coin? The Journal of animal ecology. 2010;79: 811–7. doi:10.1111/j.1365-2656.2010.01688.x

23. Pires MM, Guimarães PR. Interaction intimacy organizes networks of antagonistic interactions in different ways. J R Soc Interface. 2013;10: 20120649. doi:10.1098/rsif.2012.0649

24. Song C, Saavedra S. Telling ecological networks apart by their structure: An environment-dependent approach. O'Dwyer J, editor. PLoS Comput Biol. 2020;16: e1007787. doi:10.1371/journal.pcbi.1007787

25. Simmons BI, Cirtwill AR, Baker NJ, Wauchope HS, Dicks LV, Stouffer DB, et al. Motifs in bipartite ecological networks: uncovering indirect interactions. Oikos. 2019;128: 154–170. doi:10.1111/oik.05670

26. Stouffer DB, Bascompte J. Understanding food-web persistence from local to global scales. Ecology Letters. 2010;13: 154–161. doi:10.1111/j.1461-0248.2009.01407.x

27. Losapio G, Schöb C, Staniczenko PPA, Carrara F, Palamara GM, De Moraes CM, et al. Network motifs involving both competition and facilitation predict biodiversity in alpine plant communities. Proc Natl Acad Sci USA. 2021;118: e2005759118. doi:10.1073/pnas.2005759118

28. Vitali A, Sasal Y, Vázquez DP, Miguel MF, Rodríguez-Cabal MA. The disruption of a keystone interaction erodes pollination and seed dispersal networks. Ecology. 2022;103. doi:10.1002/ecy.3547

29. Lewitus E, Morlon H. Characterizing and comparing phylogenies from their Laplacian spectrum. Syst Biol. 2016;65: 495–507. doi:10.1093/sysbio/syv116

30. de Lange SC, de Reus MA, van den Heuvel MP. The Laplacian spectrum of neural networks. Front Comput Neurosci. 2014;7. doi:10.3389/fncom.2013.00189

31. Chagnon P-L, U'Ren JM, Miadlikowska J, Lutzoni F, Arnold AE. Interaction type influences ecological network structure more than local abiotic conditions: evidence from endophytic and endolichenic fungi at a continental scale. Oecologia. 2016;180: 181–191. doi:10.1007/s00442-015-3457-5

32. Selosse M, Schneider-Maunoury L, Martos F. Time to re-think fungal ecology? Fungal ecological niches are often prejudged. New Phytologist. 2018;217: 968–972. doi:10.1111/nph.14983

33. Freilich S, Kreimer A, Meilijson I, Gophna U, Sharan R, Ruppin E. The large-scale organization of the bacterial network of ecological co-occurrence interactions. Nucleic Acids Research. 2010;38: 3857–3868. doi:10.1093/nar/gkq118

34. Chaffron S, Delage E, Budinich M, Vintache D, Henry N, Nef C, et al. Environmental vulnerability of the global ocean epipelagic plankton community interactome. Science Advances. 2021;7: eabg1921. doi:10.1126/sciadv.abg1921

35. Lima-Mendez G, Faust K, Henry N, Decelle J, Colin S, Carcillo F, et al. Determinants of community structure in the global plankton interactome. Science. 2015;348: 1262073–1262073. doi:10.1126/science.1262073

36. Bjorbækmo MFM, Evenstad A, Røsæg LL, Krabberød AK, Logares R. The planktonic protist interactome: where do we stand after a century of research? ISME J. 2020;14: 544–559. doi:10.1038/s41396-019-0542-5

37. Simmons BI, Sweering MJM, Schillinger M, Dicks LV, Sutherland WJ, Di Clemente R. bmotif: A package for motif analyses of bipartite networks. Matthiopoulos J, editor. Methods Ecol Evol. 2019;10: 695–701. doi:10.1111/2041-210X.13149

38. Maliet O, Loeuille N, Morlon H. An individual-based model for the eco-evolutionary emergence of bipartite interaction networks. Poisot T, editor. Ecol Lett. 2020; ele.13592. doi:10.1111/ele.13592

39. Martos F, Munoz F, Pailler T, Kottke I, Gonneau C, Selosse M-A. The role of epiphytism in architecture and evolutionary constraint within mycorrhizal networks of tropical orchids. Molecular Ecology. 2012;21: 5098–5109. doi:10.1111/j.1365-294X.2012.05692.x

40. Chagnon P-L, Bradley RL, Klironomos JN. Trait-based partner selection drives mycorrhizal network assembly. Oikos. 2015;124: 1609–1616. doi:10.1111/oik.01987

41. Jacquemyn H, Brys R, Waud M, Busschaert P, Lievens B. Mycorrhizal networks and coexistence in species-rich orchid communities. New Phytol. 2015;206: 1127–1134. doi:10.1111/nph.13281

42. Xing X, Jacquemyn H, Gai X, Gao Y, Liu Q, Zhao Z, et al. The impact of life form on the architecture of orchid mycorrhizal networks in tropical forest. Oikos. 2019;128: 1254–1264. doi:10.1111/oik.06363

43. Perez-Lamarque B, Petrolli R, Strullu-Derrien C, Strasberg D, Morlon H, Selosse M-A, et al. Structure and specialization of mycorrhizal networks in phylogenetically diverse tropical communities. Environmental Microbiome. 2022;17: 38. doi:10.1186/s40793-022-00434-0

44. Mello MAR, Marquitti FMD, Guimarães PR, Kalko EKV, Jordano P, de Aguiar MAM. The missing part of seed dispersal networks: Structure and robustness of bat-fruit interactions. Traveset A, editor. PLoS ONE. 2011;6: e17395. doi:10.1371/journal.pone.0017395

45. Toju H, Guimarães PR, Olesen JM, Thompson JN. Below-ground plant–fungus network topology is not congruent with above-ground plant–animal network topology. Sci Adv. 2015;1: e1500291. doi:10.1126/sciadv.1500291

46. Hembry DH, Raimundo RLG, Newman EA, Atkinson L, Guo C, Guimarães PR, et al. Does biological intimacy shape ecological network structure? A test using a brood pollination mutualism on continental and oceanic islands. Ings T, editor. J Anim Ecol. 2018;87: 1160–1171. doi:10.1111/1365-2656.12841

47. Staniczenko PPA, Kopp JC, Allesina S. The ghost of nestedness in ecological networks. Nat Commun. 2013;4: 1391. doi:10.1038/ncomms2422

48. Miranda PN, Ribeiro JEL da S, Luna P, Brasil I, Delabie JHC, Dáttilo W. The dilemma of binary or weighted data in interaction networks. Ecological Complexity. 2019;38: 1–10. doi:10.1016/j.ecocom.2018.12.006

49. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A simple way to prevent neural networks from overfitting. Journal of Machine Learning Research. 2014;15: 1929–1958.

50. Wu Z, Pan S, Chen F, Long G, Zhang C, Yu PS. A comprehensive survey on graph neural networks. IEEE Trans Neural Netw Learn Syst. 2021;32: 4–24. doi:10.1109/TNNLS.2020.2978386

51. Montesinos-Navarro A, Hiraldo F, Tella JL, Blanco G. Network structure embracing mutualism–antagonism continuums increases community robustness. Nat Ecol Evol. 2017;1: 1661–1669. doi:10.1038/s41559-017-0320-6

52. Klironomos JN. Variation in plant response to native and exotic arbuscular mycorrhizal fungi. Ecology. 2003;84: 2292–2301. doi:10.1890/02-0413

53. Genini J, Morellato LPC, Guimarães PR, Olesen JM. Cheaters in mutualism networks. Biol Lett. 2010;6: 494–497. doi:10.1098/rsbl.2009.1021

54. Perez-Lamarque B, Selosse M, Öpik M, Morlon H, Martos F. Cheating in arbuscular mycorrhizal mutualism: a network and phylogenetic analysis of mycoheterotrophy. New Phytol. 2020;226: 1822–1835. doi:10.1111/nph.16474

55. Sachs JL, Skophammer RG, Regus JU. Evolutionary transitions in bacterial symbiosis. Proceedings of the National Academy of Sciences. 2011;108: 10800–10807. doi:10.1073/pnas.1100304108

56. Canestrari D, Bolopo D, Turlings TCJ, Roder G, Marcos JM, Baglione V. From parasitism to mutualism: Unexpected interactions etween a cuckoo and Its host. Science. 2014;343: 1350–1352. doi:10.1126/science.1249008

57. Thebault E, Fontaine C. Stability of ecological communities and the architecture of mutualistic and trophic networks. Science. 2010;329: 853–856. doi:10.1126/science.1188321

58. Cai W, Snyder J, Hastings A, D'Souza RM. Mutualistic networks emerging from adaptive niche-based interactions. Nat Commun. 2020;11: 5470. doi:10.1038/s41467-020-19154-5

59. Strydom T, Catchen MD, Banville F, Caron D, Dansereau G, Desjardins-Proulx P, et al. A roadmap towards predicting species interaction networks (across space and time). Phil Trans R Soc B. 2021;376: 20210063. doi:10.1098/rstb.2021.0063

60. Blanchet FG, Cazelles K, Gravel D. Co-occurrence is not evidence of ecological interactions. Jeffers E, editor. Ecol Lett. 2020;23: 1050–1063. doi:10.1111/ele.13525

61. Fortuna MA, Ortega R, Bascompte J. The Web of Life. arXiv:14032575 [q-bio]. 2014 [cited 10 Jun 2020]. Available: http://arxiv.org/abs/1403.2575

62. Öpik M, Metsis M, Daniell TJ, Zobel M, Moora M. Large-scale parallel 454 sequencing reveals host ecological group specificity of arbuscular mycorrhizal fungi in a boreonemoral forest. New Phytologist. 2009;184: 424–437. doi:10.1111/j.1469-8137.2009.02920.x

63. Jacquemyn H, Merckx V, Brys R, Tyteca D, Cammue BPA, Honnay O, et al. Analysis of network architecture reveals phylogenetic constraints on mycorrhizal specificity in the genus Orchis (Orchidaceae). New Phytologist. 2011;192: 518–528. doi:10.1111/j.1469-8137.2011.03796.x

64. Montesinos-Navarro A, Segarra-Moragues JG, Valiente-Banuet A, Verdú M. Plant facilitation occurs between species differing in their associated arbuscular mycorrhizal fungi. New Phytol. 2012;196: 835–844. doi:10.1111/j.1469-8137.2012.04290.x

65. Põlme S, Bahram M, Jacquemyn H, Kennedy P, Kohout P, Moora M, et al. Host preference and network properties in biotrophic plant-fungal associations. New Phytol. 2018;217: 1230–1239. doi:10.1111/nph.14895

66. Hoeppke C, Simmons B. Maxnodf: an R package for fair and fast comparisons of nestedness between networks. Methods in Ecology and Evolution. 2021;12. doi:10.1111/2041-210x.13545

67. Dormann CF, Gruber B, Fründ J. Introducing the bipartite Package: Analysing Ecological Networks. 2008;8: 5.

68. Simmons BI, Sweering MJM, Schillinger M, Dicks LV, Sutherland WJ, Clemente RD. bmotif: a package for motif analyses of bipartite networks. Ecology; 2018 Apr. doi:10.1101/302356

69. Newman MEJ. Finding community structure in networks using the eigenvectors of matrices. Phys Rev E. 2006;74: 036104. doi:10.1103/PhysRevE.74.036104

70. Banerjee A, Jost J. On the spectrum of the normalized graph Laplacian. Linear Algebra and its Applications. 2008;428: 3015–3022. doi:10.1016/j.laa.2008.01.029

71. Morlon H, Lewitus E, Condamine FL, Manceau M, Clavel J, Drury J. RPANDA : an R package for macroevolutionary analyses on phylogenetic trees. Fitzjohn R, editor. Methods Ecol Evol. 2016;7: 589–597. doi:10.1111/2041-210X.12526

72. Lê S, Josse J, Husson F. FactoMineR : an R package for multivariate analysis. J Stat Soft. 2008;25. doi:10.18637/jss.v025.i01

73. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Soft. 2010;33. doi:10.18637/jss.v033.i01

74. Günther F, Fritsch S. neuralnet: Training of Neural Networks. The R Journal. 2010;2: 30. doi:10.32614/RJ-2010-006

75. Perez-Lamarque B, Maliet O, Pichon B, Selosse M-A, Martos F, Morlon H. Do closely related species interact with similar partners? Testing for phylogenetic signal in bipartite interaction networks. Peer Community Journal. 2022;2: e59. doi:10.24072/pcjournal.179