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

Running headline: Multiscale network structure and 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.

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

All authors conceived the ideas and designed methodology; B.P performed research. B.P, H.M and B.P.L led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Data availability

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

Abstract

1. 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., pollination) interaction types.

2. 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.

3. 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. We also find structural uniqueness linked to specific ecologies (e.g., pollination, parasitism) beyond the classical mutualism/antagonism dichotomy.

4. 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

1. Introduction

Species in ecosystems engage in multiple types of antagonistic or mutualistic interactions, such as predation, parasitism, pollination, seed dispersal, or mycorrhizal symbioses (Bronstein, 2015; Thomas et al., 2005). Such interactions between two groups of potentially taxonomically distant species are often represented by bipartite networks (Bascompte et al., 2003; Bascompte & Jordano, 2007; Fontaine et al., 2011). Studying the structure of these interaction networks is fundamental for unraveling the different processes behind their ecological organization (Bastolla et al., 2009; Fontaine et al., 2011; Suweis et al., 2013; Thompson, 2005), understanding their evolutionary dynamics, and prioritizing conservation efforts (Bastolla et al., 2009; Olesen et al., 2007; Rezende, Lavabre, et al., 2007).

Early studies on bipartite ecological networks, focused mainly on pollination and herbivory networks, suggested that mutualistic and antagonistic networks have different structures (Bascompte et al., 2003; Thébault & Fontaine, 2008, 2010). 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 (Krause et al., 2003; Newman, 2006b). Conversely, mutualistic networks tend to have a nested structure, where specialist species mainly interact with generalists, and generalists form a core of interactions (Jordano et al., 2002; Rohr et al., 2014; Bascompte et al., 2003). Such structural differences have been observed in both terrestrial (Fontaine et al., 2011) and marine networks (Guimarães et al., 2007). However, several studies have challenged the generality of the modular/nested dichotomy (Chagnon, 2016; Fontoura et al., 2020; Michalska-Smith & Allesina, 2019; Pilosof et al., 2014). 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 (Fortuna et al., 2010). 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 (Fontaine et al., 2011; Guimarães et al., 2007; Pires & Guimarães, 2013). 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 (Michalska-Smith & Allesina, 2019). 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 (Song & Saavedra, 2020).

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 (Simmons, Cirtwill, et al., 2019). 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 (Milo et al., 2002). Motifs have been informative in various contexts, like the assessment of network stability or resistance to biological invasions (Losapio et al., 2021; Stouffer & Bascompte, 2010; Vitali et al., 2022). 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 (Lewitus & Morlon, 2016) or brain networks (de Lange et al., 2014), but only marginally to study the structure of ecological networks (Michalska-Smith & Allesina, 2019). 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 (Chagnon et al., 2016; Selosse et al., 2018). More generally, networks involving microorganisms, such as bacteria in soils (Freilich et al., 2010) and plankton in the oceans (Chaffron et al., 2021; Lima-Mendez et al., 2015), have often been constructed using presence/absence data rather than direct observations, and the type of a large proportion of these interactions remains undetermined (Bjorbækmo et al., 2020; Lima-Mendez et al., 2015).

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 accuracy of the classifiers. We compare the predictive power of the different structural metrics and classification methods. Finally, we discuss the implications of our findings for the understanding of natural communities.

2. 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 (Michalska-Smith & Allesina, 2019), among which some come from the Web of Life database (web-of-life.es (Fortuna et al., 2014)). We kept networks containing at least 10 species in each guild, as the structure of small networks is difficult to characterize (Olesen et al., 2007), 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 (Jacquemyn et al., 2011, 2015; Martos et al., 2012; Montesinos-Navarro et al., 2012; Öpik et al., 2009; Xing et al., 2019) and 3 from the database generated by (Põlme et al., 2018). 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 modularity. We quantified nestedness using the NODFc index that corrects for network size ("NODFc" function from the R-package maxnodf, version 1.0.0 (Hoeppke & Simmons, 2021)). We quantified modularity using the function *computeModules* from the R-package bipartite (version 2.16) (Dormann et al., 2008) with the "Beckett" method.

Second, we characterized the mesoscale structure of the networks by performing bipartite motif analyses that depict the patterns of interactions between subsets of species (called motifs; Fig. 1) (Simmons, Sweering, et al., 2019). For each network, we computed the frequencies of the 44 possible motifs involving between 2 and 6 nodes using the function *mcount* from the bmotif R-package (version 2.0.2) (Simmons et al., 2018) with the normalization procedure "normalise_sum" (*i.e.* which computes the frequency of motifs among the total number of motifs). The function *mcount* does not allow counting motifs with more than 6 nodes, and limiting the size to 6 seemed reasonable given the size of our smaller networks (10 species): the frequencies of larger motifs is likely to be more important in larger networks, independently of the type of interactions.

171 Third, we used a multi-scale approach stemming from graph theory that integrates the
172 whole structural information of a network, from micro- to macro-scales: the spectral density
173 of the Laplacian graph (Newman, 2006a). We computed the normalized Laplacian graph of
174 each network (Newman, 2006a), defined as $L_{norm} = D^{-1/2} L D^{-1/2}$, where D is the degree matrix (a
175 diagonal matrix counting the number of interactions for each species) and L is the Laplacian
176 graph, defined as the difference between the adjacency matrix (a square binary matrix whose
177 elements represent the interactions between species of the network) and the degree matrix
178 (Fig. S2A). The normalized Laplacian graph is a symmetric and positive-definite matrix whose
179 eigenvalues are constrained between 0 and 2 (Banerjee & Jost, 2008), which makes them
180 comparable between networks of different sizes, and symmetrical around 1. There are as many
181 eigenvalues equal to 0 (or 2) as there are connected components (*i.e.* perfect modules of species
182 that are not connected with any species from other modules), and we can thus expect a
183 modular network to have more eigenvalues close to 0 (and 2) than a nested network. To be
184 able to efficiently compare Laplacian graphs across networks of different sizes, we converted
185 the eigenvalues into a spectral density using a Gaussian convolution (Lewitus & Morlon,
186 2016)(Fig. S2A). To do so, we adapted the *spectR* function from the RPANDA R-package
187 (version 1.9) (Morlon et al., 2016) for bipartite networks and used a bandwidth of 0.025 for the
188 Gaussian convolution. We tested the influence of the bandwidth value (from 0.001 to 0.091) on
189 the resulting spectral densities. We characterized each network by extracting its spectral
190 density values at 100 points uniformly distributed between 0 and 1.

191 The type of ecology (*e.g.,* pollination, host-parasite) can impact network structure
192 (Chagnon et al., 2016; Guimarães et al., 2007). To assess whether there are structural differences
193 between mutualistic and antagonistic networks beyond ecology, we controlled for the

194 confounding effect of ecology by fitting linear mixed models (function *lmer* from the R-package
195 lme4 (version 1.1-27)) with the type of interaction (*i.e.* mutualistic or antagonistic) as a fixed
196 effect, the type of ecology as a random effect, and one structural metric as the response variable
197 (Type II Wald test). In addition, for each structural metric, we fitted linear models with only
198 the type of ecology or the type of interaction as fixed effects and evaluated the best-fitting
199 model by comparing their Akaike information criterion (AIC).

200

201 **Unsupervised classification**

202 To visualize structural differences across antagonistic *versus* mutualistic networks, we
203 used principal component analyses (PCA), with networks characterized either in terms of their
204 global metrics, motif frequencies, or Laplacian spectral densities. PCAs were performed with
205 the functions *PCA* and *fviz_pca_ind* from the FactoMineR and factoextra R-packages (versions
206 2.4 and 1.0.7 respectively) (Lê et al., 2008).

207 For each type of characterization, we then used unsupervised K-means classification
208 on the two main principal components projections. K-means were performed with the function
209 *kmeans* with 2 clusters, and we performed chi-squared tests to investigate whether the two
210 clusters tend to separate mutualistic from antagonistic networks. To quantify the ability to
211 separate antagonistic and mutualistic networks, we used a measure of classification accuracy
212 computed as the F-score with the function *F1_score* from the R-package MLmetrics. The F-
213 score is the harmonic mean of the precision (*e.g.*, the frequency of true mutualistic among
214 predicted mutualistic networks) and the recall (*e.g.*, the frequency of predicted mutualistic
215 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. We measured the accuracy of the approach using both the percentages of mutualistic and antagonistic networks from the test set that are correctly classified by the classifier and the F-scores of the classification of both types of networks.

We first investigated the accuracy 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) (Friedman et al., 2010). We found that training the classifiers with 80% of the interaction networks and testing it on the remaining 20% provided a good trade-off between the size of the test set and the accuracy of the classification.

Second, we used artificial neural networks (ANN) implemented in the *neuralnet* function from the R-package neuralnet (version 1.44.2) (Günther & Fritsch, 2010). We optimized their configuration (number of layers and number of neurons in each layer) according to the input variables. We selected a 3 layers network, constituted of the input and

239 output layers and an intermediate layer with 5, 20, and 60 neurons for global metrics, motif
240 frequencies, and spectral density inputs respectively. Preliminary analyses showed that these
241 choices provided a good compromise between classification accuracy and overfitting (results
242 not shown). We also selected a non-linear logistic transformation, and the maximal number of
243 steps was fixed at 10^8.

244 We independently trained and tested the classifiers on empirical networks using the
245 different sets of structural input variables. For each configuration, the ANN trainings were
246 replicated 50 times with different training and test sets, and we reported the average
247 percentages of correct classifications among antagonistic and mutualistic networks and their
248 respective F-scores.

249

250 **Sensitivity analyses on the supervised classification using the motif frequencies**

251 We selected the most accurate classification approach (supervised classification using
252 ANN on motif frequencies; see Results) and carried out a series of additional sensitivity
253 analyses: we (i) developed a measure of classification confidence, (ii) assessed the impact of
254 overfitting and the over-representation of mutualistic *versus* antagonistic networks and of
255 specific types of ecology (e.g. pollination) in the training set, (iii) tested the specificity of our
256 approach, *sensu* (Michalska-Smith & Allesina, 2019) (iv) investigated the effect of the types of
257 ecology on our classification, and finally (iv) analyzed which motif complexity is required to
258 accurately classify networks.

259 First, because the training sets were limited in size (we only had a total of 343 empirical
260 networks), we introduced a measure of the classification confidence for each empirical

261 network. We trained and tested the ANN classifier on motif frequencies for 10,000
262 independent training and test sets of empirical networks: each empirical network was
263 therefore classified on average 2,000 times with different training sets. We then computed the
264 percentage of classifications that fell in each category, which gives an idea of classification
265 confidence: we considered that a given network was classified as mutualistic (resp.
266 antagonistic) with "high confidence" if it was classified as mutualistic (resp. antagonistic) in
267 more than 75% of the classifications. Conversely, a given network was classified as mutualistic
268 (resp. antagonistic) with "low confidence" if it was classified as mutualistic (resp. antagonistic)
269 in more than 50% of the classifications but in less than 75% of them.

270 Second, we measured overfitting by comparing the percentages of correct
271 classifications obtained on the training *versus* test sets as a function of the number of neurons
272 in the hidden layer. In addition, we assessed the impact on our results of the over-
273 representation of mutualistic *versus* antagonistic networks in the training set (more than 60%
274 of the networks were mutualistic): we trained the classifier with the same number of
275 antagonistic and mutualistic networks (from 5 to 125 networks for each type of interaction)
276 and compared the percentages of correct classifications with the ones obtained with the same-
277 size training sets composed of networks randomly chosen in the two interaction types. In each
278 condition, we replicated the classification 50 times with different training sets. Similarly, we
279 assessed the impact on our results of the over or under-representation of some types of ecology
280 in the training set (e.g. the pollination networks represent ~50% of the complete database). To
281 do so, we trained the classifier with the same number of networks per type of ecology (from 5
282 to 20 networks for each type of ecology; if a type of ecology had fewer networks than the
283 number per type of ecology, all networks were included in the training set) and compared the

284 percentages of correct classifications with the ones obtained with same-size training sets
285 composed of networks chosen randomly across ecology types.

286 Third, we tested the specificity of our approach, *sensu* (Michalska-Smith & Allesina,
287 2019): we applied perturbations to the empirical networks in order to investigate whether the
288 classification accuracies we obtained were not driven only by differences in network size,
289 connectance, and/or degree distribution between antagonistic and mutualistic networks (Fig.
290 2A). We used three different randomization strategies, all of which maintained network size
291 constant, and tested our classification approach on each randomized dataset. First, we
292 randomized the networks while keeping their connectance fixed using the *shuffle.web* function
293 of the **bipartite** R-package. Second, we randomized the networks while keeping their degree
294 distributions fixed using the r2table algorithm implemented in the *null.model* function of the
295 **bipartite** R-package**.** Third, we randomized the networks while keeping both the connectance
296 and their degree distributions fixed using the *permatswap* function with quasiswap method
297 from the **vegan** R-package (v2.5-7). For each of these 3 randomization strategies, we
298 randomized 200 times each network, and for each randomization, we classified them using
299 ANN on motif frequencies with 50 different training and testing sets.

300 Fourth, we investigated if the classifiers were strictly learning to separate antagonistic
301 from mutualistic networks or if they also learned the potential structural singularities of each
302 type of ecology from our database (*i.e.*, data-linkage). We performed two different analyses to
303 control for the signal from the type of ecology. First, we removed one by one each type of
304 ecology from the training test (*e.g.* pollination networks), we trained the classifiers on 50
305 independent training sets constituted by 80% of the remaining networks, and we tested the
306 classification on the networks from the removed type of ecology (*e.g.* are pollination networks

classified among mutualistic networks when no pollination network has been used in the training?). If the classifiers learn to separate antagonistic from mutualistic networks regardless of ecology, the classification should perform as well as the original classification. If they learn to classify ecologies rather than the type of interaction, then it should perform very badly. Second, we investigated whether the structural differences between mutualistic and antagonistic networks learned by our ANN classifier were explained only by structural differences specific to each type of ecology (*e.g.,* pollination *versus* host-parasite). To do so, we shuffled the "mutualistic" or "antagonistic" nature of each type of ecology while keeping 5 (*resp.* 4) different types of ecology as "mutualistic" (*resp.* "antagonistic"). In such shuffled datasets, "antagonistic" networks could for example correspond to all pollination, host-parasitoid, mycorrhizal and anemone-fish networks, whereas "mutualistic" networks could correspond to all bacteria-phage, herbivory, host-parasite, ant-plant, and seed dispersal networks. If the structural differences from the type of ecology alone drive the ANN classifications, we expect similar accuracies with these shuffled datasets compared with the original dataset. We performed 300 of such shuffling, and for each shuffled dataset, we classified "mutualistic" *versus* "antagonistic" networks using ANN on motif frequencies with 50 different training and testing sets. In this analysis, we noticed cases when most networks were classified in one type of interaction (the most represented interaction in terms of number of networks); in this case, the percentage of correct classifications will appear very high for this interaction type, and very low for the other. Therefore, in addition to the mean accuracy of antagonistic and mutualistic networks, we measured the difference in accuracy. A large difference reflects a loss of the ANN ability to classify properly.

329 Finally, we investigated which level of motif complexity (size) is required to
330 distinguish mutualistic from antagonistic networks. To do so, we measured the classification
331 accuracy obtained when training the ANN on two subsets of all motifs, corresponding to the
332 removal of the 6-nodes motifs, and then the removal of both 5 and 6-nodes motifs.

333

334 **Simulations of bipartite interaction networks**

335 The accuracy of supervised classifiers often depends on the size of the available training
336 set (i.e., data for which the classification is known), which in the case of bipartite interaction
337 networks is rather small (n=343 here). A possible approach to increase the training set is to use
338 simulated data. We explored this possibility with the recently-developed BipartiteEvol model
339 (Maliet et al., 2020), an eco-evolutionary model that simulates the evolution of interacting
340 individuals within two clades (as detailed in Supplementary Note 1).

341

342 # 3. Results

343 **Structure of antagonistic versus mutualistic networks**

344 We found differences in macro-scale structures between antagonistic and mutualistic
345 networks.. For instance, nestedness was significantly higher in mutualistic networks (mean:
346 2.16 ± s.d. 0.71) than in antagonistic ones (1.94 ± 1.09; Wilcoxon-Mann-Whitney test: W = 8673,
347 p-value < 0.001), whereas antagonistic networks were significantly more connected (0.23 ±
348 0.17) than mutualistic ones (0.15 ± 0.09; Wilcoxon-Mann-Whitney test: W = 17709, p-value <
349 0.001). The difference in terms of modularity was not significant (Wilcoxon-Mann-Whitney

350 test: W = 12559, p-value = 0.19) (Fig. 2A). In addition, we found that mutualistic networks were
351 composed of more species compared to antagonistic ones (76.1± 92.3 for antagonistic networks
352 and 82.7 ± 72.7 for mutualistic ones; Wilcoxon-Mann-Whitney test: W = 11124, p-value = 0.003).
353 These differences were however better explained by specific ecologies rather than the overall
354 interaction type. Specifically, when controlling for the different types of ecology (*e.g.*,
355 pollination, herbivory, seed dispersal...), antagonistic and mutualistic networks were not
356 significantly different in terms of their nestedness (Wald test: $\chi 2$ = 0.060, p-value = 0.81),
357 connectance (Wald test: $\chi 2$ = 0. 41, p-value = 0.52), modularity (Wald test: $\chi 2$ = 0.42, p-value =
358 0. 52), the total number of species (Wald test: $\chi 2$ = 0. 29, p-value = 0.59), and absolute difference
359 in the number of species per guild (Wald test: $\chi 2$ = 0. 95, p-value = 0.33). Model comparisons
360 confirmed that differences in terms of global metrics were better explained by the different
361 types of ecology than by the antagonistic/mutualistic type of interactions (Table S1). For
362 instance, compared to other types of ecology, bacteria-phage networks were more connected
363 but generally less nested, while herbivory networks were more nested (Fig. S1, Table S2 A-C).

364 Next, using the spectral density of the Laplacian graph (Fig. S2A) on perfectly modular
365 or nested networks revealed that modular networks have rather smooth densities with high
366 peaks whereas nested networks have noisier densities with several small peaks (Fig. S2B-E).
367 We found that mutualistic networks were enriched in eigenvalues near 1, while antagonistic
368 networks had significantly more spectral eigenvalues near 0 (Fig. 2B), which was mainly
369 driven by herbivory but not host-parasite networks (Fig. S3).

370 Finally, using motif frequencies, we found, following the motif denomination of
371 (Simmons, Sweering, et al., 2019), that antagonistic networks were significantly enriched in
372 densely connected motifs (such as motif 12 and motifs 31 to 37), whereas mutualistic networks

were enriched in motifs with asymmetrical specialization, wherein specialist species tend to interact with the most generalist species in the motif (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, 4 motifs were significantly different in frequency between antagonistic and mutualistic networks (motifs 17, 25, 28, and 44). 31 out of 44 motifs differed significantly in frequency across the different types of ecology (Fig. S4C): for example, seed-dispersal and host-parasite networks had relatively high frequencies of 4- to 6-nodes motifs, in particular motifs 29 to 43, which may be linked to their high connectance (Figs. S1, S4., Table S2A).

Unsupervised classifications of empirical networks

A principal component analysis (PCA) on the global metrics did not identify distinct clusters for mutualistic and antagonistic networks (Fig. 3A). Consequently, an unsupervised K-means classification on the principal component axes 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 (Fig. S5) and did not systematically fell in one of the two clusters of the K-means classification (Table S3A).

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

395 the high number of pollination networks in the database (Figs. S6-7). Unsupervised K-means
396 classifications on the principal component axes were not sufficient to confidently classify
397 mutualistic and antagonistic networks (Fig. 3B-C). Networks with different types of ecology
398 were largely mixed between the two clusters from the K-means classification (Tables S3 B-C),
399 suggesting that unsupervised classifications do not perform well at classifying ecological
400 networks, regardless of the metric used to characterize their structure.

401

402 **Supervised classifications of empirical networks**

403 We found that the two linear classifiers (Logit and Lasso regressions) mainly classified
404 empirical networks as mutualistic, regardless of their interaction type (Table S4). Artificial
405 neural networks (ANN) performed much better, in particular when using motif frequencies.
406 Motif frequencies led to better classifications than the spectral density or the global metrics,
407 and adding other variables to motif frequencies did not improve the classifications (Table S4).
408 The ANN based on motif frequencies classified an average of 65% of the antagonistic networks
409 (resp. 84% of the mutualistic networks) classified as antagonistic (resp. mutualistic), resulting
410 in an F-score of 0.68 and 0.82 for antagonistic and mutualistic networks respectively (Table S4).
411 Irrespectively of the bandwidth used to smooth the Laplacian spectrum, using the ANN with
412 the spectral density led to lower accuracies (Tables S4-5). We therefore used ANN on motif
413 frequencies for the remaining analyses.

414

415 **Supervised classification of the empirical networks using motif frequencies**

416 Using our measure of classification accuracy with ANN on motif frequencies, we found
417 that 69% (52% and 17% with high and low confidence respectively; F-score=0.74) of the
418 antagonistic networks and 89% (76% and 13% with high and low confidence respectively; F-
419 score=0.86) of the mutualistic networks were correctly classified (Table 1). In total, 81% (67%
420 and 14% with high and low confidence respectively) of the networks were correctly classified
421 (Table 1). The percentage of correct classifications varied according to the type of ecology.
422 Indeed, most bacteria-phage, seed dispersal, and pollination networks were correctly
423 classified by this approach (*e.g.* 23/29 of the seed dispersal networks were correctly classified,
424 among which ~85% with high confidence). Conversely, high intimacy mutualistic networks
425 (*i.e.* ant-plant and mycorrhizal networks) presented frequent misclassification (only 7 out of 12
426 networks (~60%) were classified as mutualistic), herbivory networks (which are low intimacy
427 antagonistic networks) were mainly classified as mutualistic (13/23, *i.e.* 57%), and 50% (5/10)
428 of the host-parasitoid networks were also misclassified (Table 1).

429

430 **Robustness of the classification using motif frequencies**

431 Overfitting had a limiting impact on the result just presented (Fig. S8). In addition,
432 controlling for the overrepresentation of mutualistic network in the database resulted in
433 similar mean percentage of correct classifications (Fig. S9A). Similarly, equalizing the number
434 of networks per type of ecology in the training set did not significantly affect the overall
435 percentages of correct classifications (Fig. S9B). Finally, we found lower classification
436 accuracies when randomizing the networks compared to the original classifications (Fig. S16).
437 Yet, classifications were not completely random: for instance, 56% (mean F-score = 0.57) and
438 77% (mean F-score = 0.76) of the antagonistic and mutualistic networks respectively were

439 correctly classified using randomized networks with fixed network size and connectance (Fig.
440 S16), suggesting that differences in network size, connectance, and/or degree distributions
441 between mutualistic and antagonistic networks partially contribute to our ability to tell
442 mutualistic and antagonistic ecological networks apart from their structure but fail to explain
443 the high classification accuracy we report. Besides, we found that removing a given type of
444 ecology during the ANN training decreased the percentage of correct classifications but
445 remained quite good (Table S7). For instance, 65% of the seed dispersal and 63% of the ant-
446 plant networks were correctly classified when none of these networks were used during
447 training, compared to 79% and 66% respectively when they were included (Table S7; Table 1).
448 Also, artificially shuffling the mutualistic or antagonistic nature of each type of ecology
449 decreased the mean percentages of correct classifications by about 10%, and the F-scores by
450 about 0.08, while increasing the difference in percentages of correct classifications and F-scores
451 between antagonistic and mutualistic networks by 30% and 0.3 respectively, which reflects a
452 major loss of the classification ability (Fig. S17). Hence, although the ANN classifiers partially
453 learned the singularities of the networks linked to ecology (*i.e.* data linkage), they also learned
454 the antagonistic/mutualistic dichotomy beyond ecology.

455 Finally, we found that the classification accuracy increased when including more
456 complex motifs (Table S9). The difference was most notable between analyses carried with
457 motifs of size at most 4 and of size at most 5 or 6 (Table S9), and rather insignificant between
458 analyses carried with motifs of size at most 5 and at most 6. This suggests that complex motifs
459 (with at least 5 nodes) are needed to distinguish between mutualistic and antagonistic
460 networks, but that adding even larger motifs would not significantly increase classification
461 accuracy.

462

463 **Ability of simulated networks to improve the accuracy of classifiers**

464 Using simulated networks did not improved our ability to distinguish apart
465 mutualistic from antagonistic networks both when using K-means classifications (results not
466 shown) and ANN (Table S8B). Mutualistic networks, in particular, were often wrongly
467 classified as antagonistic. When classifying simulated BipartiteEvol networks from the test
468 sets, 70% (F-score =0.71) of antagonistic networks and 58% (F-score = 0.57) of mutualistic ones
469 were accurately classified when using motif frequencies and ANN (Table S8A), suggesting
470 difficulties in separating antagonistic *versus* mutualistic networks simulated with
471 BipartiteEvol using ANN. When comparing the structure of simulated and empirical
472 networks, we found differences in global metrics, motif composition, and density of eigen-
473 values that may explain why simulations failed to improve the accuracy of classification
474 (Supplementary Note 2).

475

4. 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 81% 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. Beyond the classical dichotomy between mutualism/antagonism, we emphasize that the different types of ecology have a important structural imprint on the structure of ecological networks.

Using motif frequencies and ANN, we were able to correctly classify around 81% 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 (Michalska-Smith & Allesina, 2019), and achieved without accounting for environmental conditions (as in Song & Saavedra, 2020). We also demonstrated the relative specificity of our approach (*sensu* Michalska-Smith & Allesina, 2019): when randomizing the network structures, we obtained lower classification accuracies, suggesting that motif frequencies, which depict patterns of species interactions at the meso-scale, capture structural differences between mutualistic and antagonistic networks that are not only due to differences in network size, connectance, or degree distributions. This therefore indicates that there are significant structural differences at the network meso-scale, which cannot be fully captured when looking at macro- or -micro-scales alone using global metrics and/or Laplacian spectral densities.

Our analyses confirm that global metrics, including the modular/nested dichotomy, are not sufficient to discriminate between mutualistic and antagonistic network structures (Fontoura et al., 2020; Michalska-Smith & Allesina, 2019). 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 (Bascompte et al., 2003; Thébault & Fontaine, 2010), 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 (Pilosof et al., 2014). Within our database, some pollination and mycorrhizal networks are known to be modular (Chagnon et al., 2015; Jacquemyn et al., 2015; Martos et al., 2012; Olesen et al., 2007; Perez-Lamarque et al., 2022; Xing et al., 2019), while they assemble mainly mutualistic interactions.

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.

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 (Simmons, Cirtwill, et al., 2019). Densely connected motifs (*e.g.* motifs 12 and 31 to 37) are significantly more frequent in antagonistic networks, while mutualistic networks are

522 characterized by motifs with asymmetrical specialization (*e.g.,* motifs 10, 19, 20, 25, 28). In
523 mutualistic networks, the emergence of motifs with asymmetrical specialization, which often
524 parallels the emergence of nestedness, has been related to three different types of explanations.
525 First, a pattern of asymmetric specialization may be partly explained from trait
526 complementarity and convergence. While generalist species can establish interactions with
527 broad range of mutualistic partners, trait complementarity generates specialization that
528 constrains the possible interactions of specialist species (Rezende, Jordano, et al., 2007;
529 Thompson, 2005). Second, such a pattern can emerge from differences in species abundances
530 (neutrality hypothesis): if interactions between individuals occur at random, rare specialist
531 species will be more likely to interact with abundant generalist species than with another rare
532 specialist (Vázquez et al., 2009). It has been proposed that a combination of trait
533 complementarity and abundance mechanisms best explains the observed structure of
534 mutualistic networks (Santamaría & Rodríguez-Gironés, 2007; Vázquez et al., 2009). Finally, it
535 has also been suggested that asymmetric specialization reduces interspecific competition
536 (Bastolla et al., 2009; Thompson, 2005), therefore enhancing the number of coexisting species
537 (Bascompte et al., 2006; Bastolla et al., 2009) and widening the conditions over which
538 mutualists coexist (Rohr et al., 2014), which in the end increases the persistence and resilience
539 of mutualistic networks to perturbations (Okuyama & Holland, 2008; Thébault & Fontaine,
540 2010). On the contrary, we expect antagonistic partners to engage in eco-evolutionary arm
541 races, which may result in modules formed by highly specialized within-module interactions
542 (i.e. densely connected motifs, which tend to be more frequent in antagonistic networks)
543 (Maliet et al., 2020; Thompson, 2005). Such modular structures would increase the resilience
544 of antagonistic communities (Thébault & Fontaine, 2010). Differences in motif frequencies can
545 also be influenced by sampling biases, including missing links or preferential detection of

some species. Future works will be required to identify the ecological processes and/or sampling effects that generate such structural differences between mutualistic and antagonistic networks (Brimacombe et al., 2023).

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 partially learns to identify each type of ecology. In particular, we showed that a better classification of networks of a particular type of ecology is achieved by using networks from the same ecology during the training of the classifiers. This is consistent with the result from linear mixed models that, after controlling for the type of ecology, the only significant difference between mutualistic and antagonistic networks is the frequency of a few motifs (motifs 17, 25, 28 and 44). These results on the importance of the type of ecology corroborate previous studies that found, for instance, significant structural differences between pollination and seed dispersal networks (Mello et al., 2011) or between mutualistic plant-fungus and plant-animal networks (Toju et al., 2015). This impact of the type of ecology may also explain the low classification accuracy of some networks, like the herbivory networks, which share similar structural properties with the pollination networks in both spectral density and motifs frequency. Nevertheless, we also demonstrated by shuffling the dataset that the high classification accuracy we obtained is achieved by learning structural differences between antagonistic and mutualistic networks beyond the type of ecology. Beyond the singularities of the type of ecology, there are structural differences

568 between mutualistic and antagonistic networks that can be captured using our meso-scale
569 approach and ANNs.

570 The accuracy of supervised classification approaches strongly depends on the size and
571 quality of the dataset used for training. We used one of the largest publicly available databases
572 of empirical networks, but it contained only 343 networks in total. In addition, it had a strong
573 unbalanced representation of the different types of ecology, with 80% of the mutualistic
574 networks represented by pollination networks and 60% of the antagonistic networks
575 represented by host-parasite networks. There was also a strong imbalance in the
576 representation of interaction intimacy, which is known to impact network structure (Fontaine
577 et al., 2011; Guimarães et al., 2007; Hembry et al., 2018; Pires & Guimarães, 2013): mutualistic
578 networks were mainly low intimacy networks, except ant-plants (n=3) or mycorrhizal
579 networks (n=9), and antagonistic networks were mainly high intimacy networks, except
580 herbivory networks (n=23). As a consequence, only ~60% of the high-intimacy mutualistic
581 networks and ~40% of the low-intimacy antagonistic networks were correctly classified. These
582 results confirm that interaction intimacy may confer a particular structure to ecological
583 networks (Fontaine et al., 2011; Guimarães et al., 2007). Increasing the number of empirical
584 networks, especially mutualistic networks with high intimacy (*e.g.* symbiotic networks) and
585 antagonistic networks with low intimacy (*e.g.* trophic networks) would improve further our
586 ability to classify ecological networks by better representing the large structural heterogeneity
587 of ecological networks (Michalska-Smith & Allesina, 2019).

588 The accuracy of supervised classification approaches can also strongly depend on the
589 type, architecture, and tuning of the classifier. For example, the overfitting we noticed in the
590 ANN (which did not seem to negatively impact our results) could potentially be avoided by

using other regularization technics, such as dropout (Srivastava et al., 2014). 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 (Wu et al., 2021).

One approach to better train the ANNs would be to increase the size of the training data thanks to synthetic data, obtained through either simulations or data augmentation. We investigated the possibility to train from simulated data with BipartiteEvol, an individual-based model that mimics the eco-evolutionary emergence of bipartite interaction networks (Maliet et al., 2020). 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. (Maliet et al., 2020) 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 frequencies of complex motifs and a lower connectance compared to empirical networks. Consequently, their projection only covers a small proportion of the space occupied by empirical networks. One possible explanation is the continuum between mutualism and antagonism in many interactions found in natural communities (Montesinos-Navarro et al., 2017). For instance, it has been shown that parrot-plant networks can combine the two types of interactions (Montesinos-Navarro et al., 2017). Similarly, pollination or mycorrhizal networks often contain cheating species that use antagonistic strategies and induce structural

614 changes within the interaction networks (Genini et al., 2010; Klironomos, 2003; Perez-
615 Lamarque et al., 2020). Interactions can also shift from mutualism to antagonism on short
616 timescales, depending on some external conditions (Canestrari et al., 2014; Sachs et al., 2011).
617 By contrast, BipartiteEvol models the emergence of a network in strict antagonistic or
618 mutualistic communities. This calls for future efforts to improve the BipartiteEvol model or to
619 use simulations under other generative models (*e.g.,* (Cai et al., 2020; Thebault & Fontaine,
620 2010). Alternatively, data augmentation methods could also be tested to generate synthetic
621 networks by randomly subsampling, mixing and rewiring empirical networks; but whether
622 such approaches produce realistic interaction networks would need to be tested first.

623

624 Ultimately, supervised classifying approaches could be particularly useful to predict
625 the type of undetermined interactions. For instance, many plant-fungus interaction networks
626 are assembled from high-throughput sequencing data, but we currently ignore the mutualistic
627 *versus* the antagonistic type of many of these endophytic interactions (Chagnon et al., 2016;
628 Selosse et al., 2018). Similarly, an increasing number of networks are built from co-occurrence
629 data generated with high-throughput sequencings, such as the ones between microscopic
630 planktonic species sampled by the *Tara Oceans* expeditions (Chaffron et al., 2021; Lima-
631 Mendez et al., 2015). The ecology of the majority of species in these networks is currently
632 undetermined (Bjorbækmo et al., 2020; Lima-Mendez et al., 2015; Strydom et al., 2021).
633 However, a preliminary analysis of bipartite networks built from co-occurrence data in the
634 *Tara Oceans* data revealed that these networks were structurally very different from the
635 empirical and simulated networks we used here (results not shown). Classifying these

636 networks would thus require using a specific training dataset composed of microbial networks
637 with known interaction types (as discussed in Supplementary Note 3).

638

639 **Conclusion**

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

651

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) network with high confidence (highlighted in grey) if it was classified as a mutualistic (resp. antagonistic) network in more than 75% of the classifications. It was predicted as a mutualistic (resp. antagonistic) network with low confidence if it was classified as a mutualistic (resp. antagonistic) network in more than 50% but less than 75% 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 (and the percentage) 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. Rows are ordered based on the % of networks classified as mutualistic within each type of ecology. In the end, 88 antagonistic and 191 mutualistic networks are correctly classified, i.e., 279 of the 343 empirical networks (81%).

Network ecology	Number of networks	Prediction			
		Antagonistic network (high confidence)	**Antagonistic network (low confidence)**	**Mutualistic network (high confidence)**	**Mutualistic network (low confidence)**
Anemone-Fish (mutualistic)	1	0 (0%)	0 (0%)	1 (100%)	0 (0%)
Pollination (mutualistic)	174	7 (4%)	7 (4%)	139 (80%)	21 (12%)
Seed dispersal (mutualistic)	29	6 (21%)	0 (0%)	20 (69%)	3 (10%)
Ant-Plant (mutualistic)	3	0 (0%)	1 (33%)	1 (33%)	1 (33%)
Herbivory (antagonistic)	23	4 (17.5%)	6 (26%)	9 (39%)	4 (17.5%)
Mycorrhiza (mutualistic)	9	3 (33.5%)	1 (11%)	3 (33.5%)	2 (22%)
Host-Parasitoid (antagonistic)	10	4 (40%)	1 (10%)	4 (40%)	1 (10%)
Bacteria-Phage (antagonistic)	17	12 (70%)	1 (6%)	3 (18%)	1 (6%)
Host-Parasitism (antagonistic)	77	46 (60%)	14 (18%)	10 (13%)	7 (9%)
Antagonistic networks	127	**66 (52%)**	**22 (17%)**	**26 (21%)**	**13 (10%)**
Mutualistic networks	216	**16 (7%)**	**9 (4%)**	**164 (76%)**	**27 (13%)**
Correct classifications (%) F-score		**88 (69%) F-score = 0.74**		**191 (88%) F-score = 0.86**	

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.

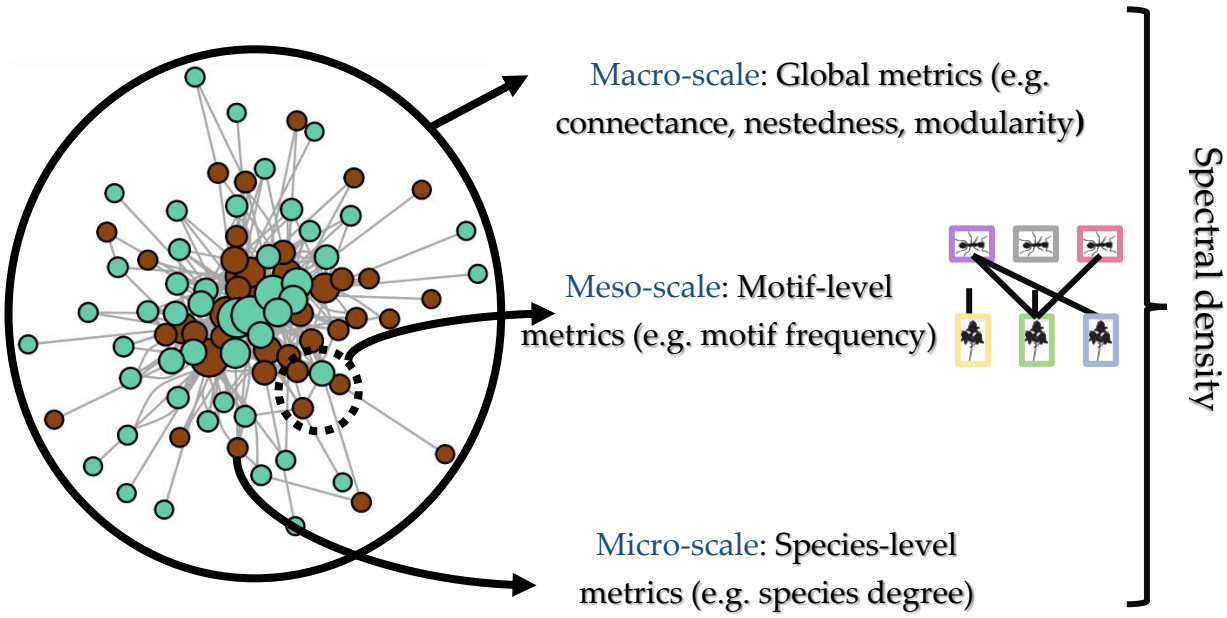


Figure 2. Differences between antagonistic and mutualistic networks across scales of network structure.

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

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

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

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

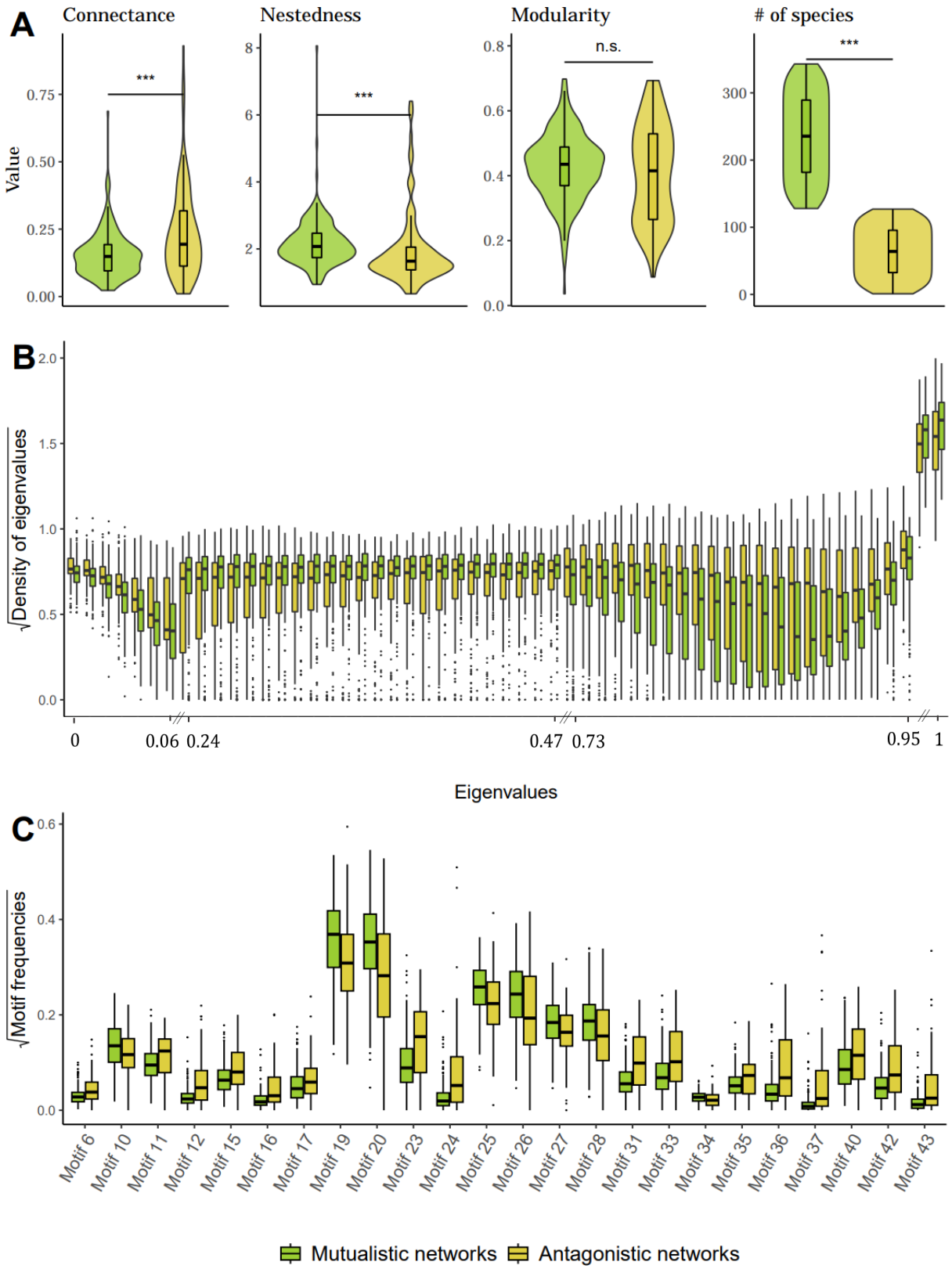
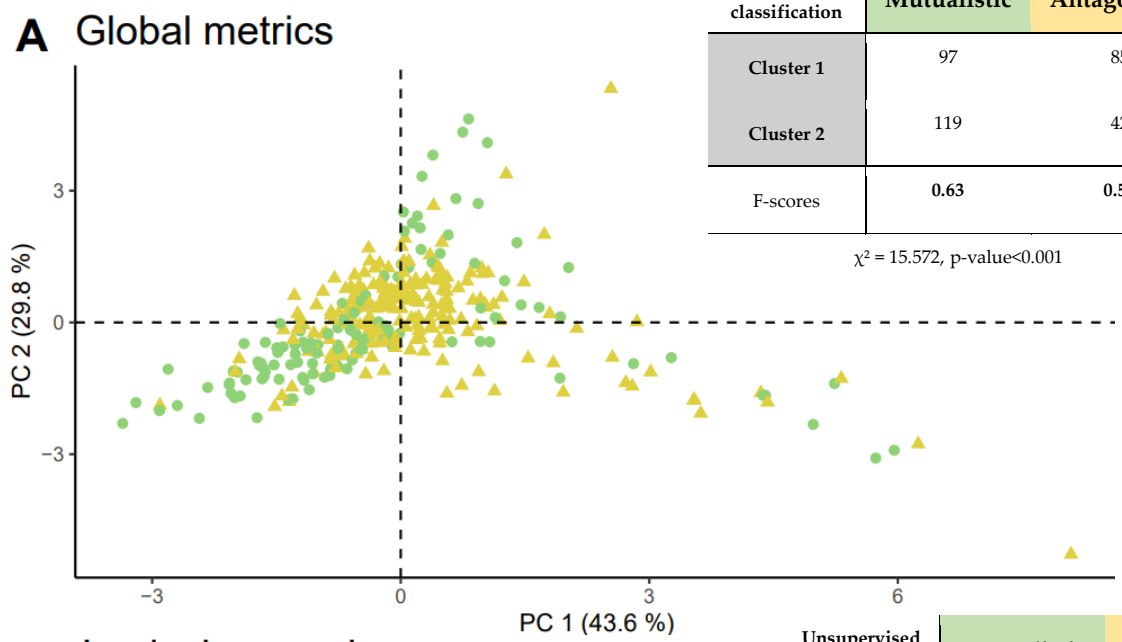

A
Connectance
Nestedness
Modularity
# of species
Value
***
n.s.
B
√Density of eigenvalues
0
0.06
0.24
0.47
0.73
0.95
1
Eigenvalues
C
√Motif frequencies
Motif 6
Motif 10
Motif 11
Motif 12
Motif 15
Motif 16
Motif 17
Motif 19
Motif 20
Motif 23
Motif 24
Motif 25
Motif 26
Motif 27
Motif 28
Motif 31
Motif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

Figure 3. Mesoscale network structure as a scale to discriminate mutualistic *from* antagonistic networks.

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.

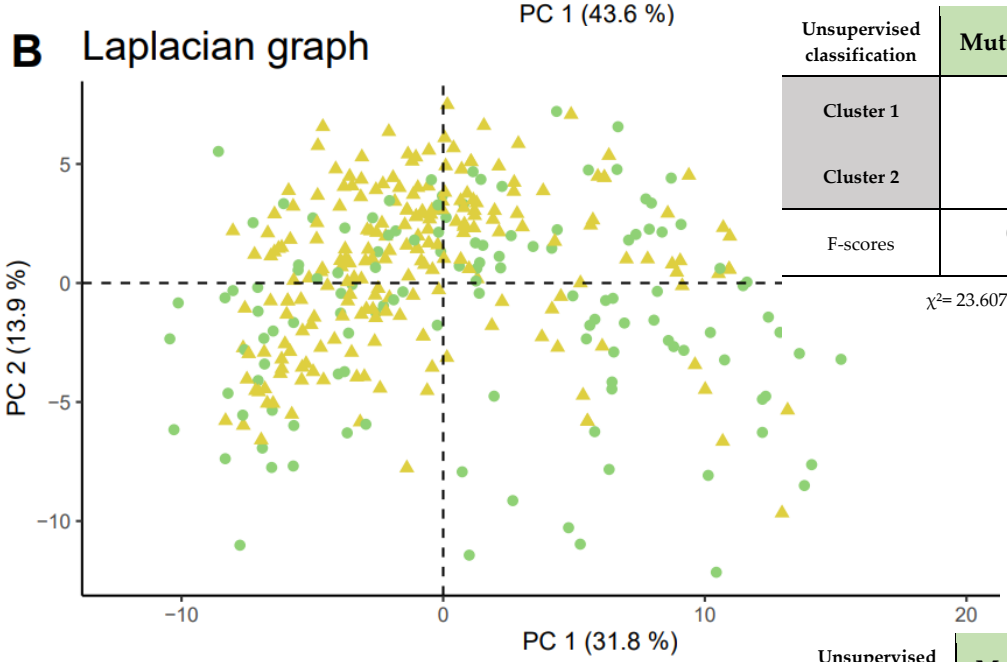
Each associated table corresponds to the contingency table of unsupervised K-means classification based on the global metrics (A), spectral density of the Laplacian graph (B), and motifs frequencies (C). Chi-squared test associated with the K-means classification is indicated.

752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775



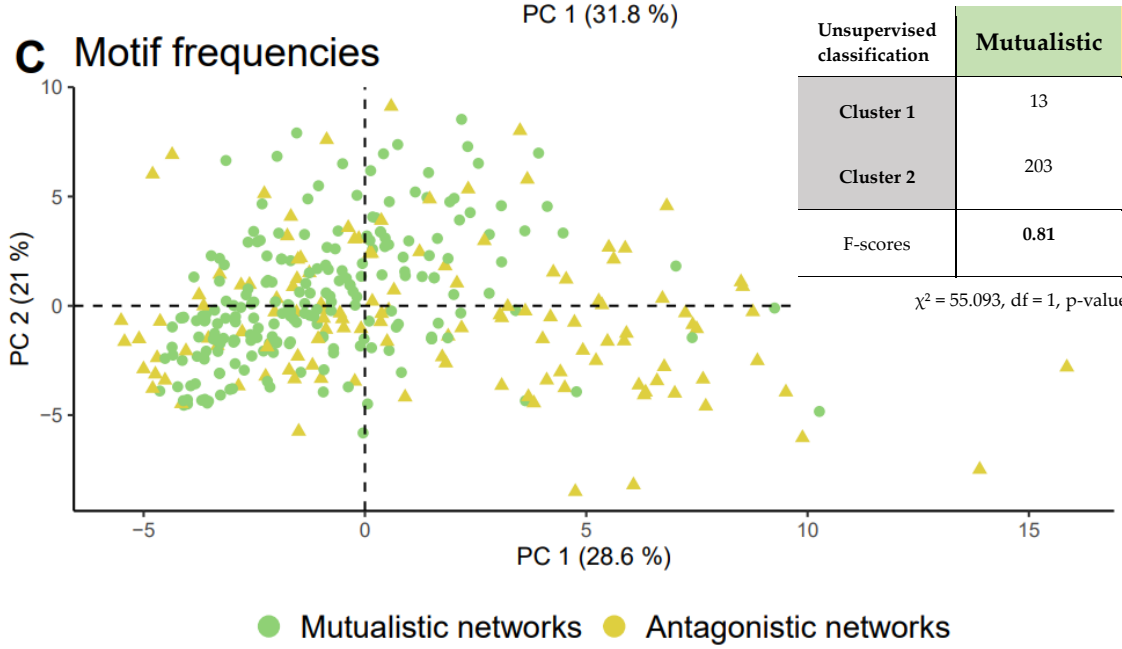
Unsupervised classification	Mutualistic	Antagonistic
Cluster 1	97	85
Cluster 2	119	42
F-scores	**0.63**	**0.55**

$\chi^2 = 15.572$, 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

Banerjee, A., & Jost, J. (2008). On the spectrum of the normalized graph Laplacian. *Linear Algebra and Its Applications*, *428*(11–12), 3015–3022. https://doi.org/10.1016/j.laa.2008.01.029

Bascompte, J., & Jordano, P. (2007). Plant-animal mutualistic networks: The architecture of biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, *38*(1), 567–593. https://doi.org/10.1146/annurev.ecolsys.38.091206.095818

Bascompte, J., Jordano, P., Melian, C. J., & Olesen, J. M. (2003). The nested assembly of plant-animal mutualistic networks. *Proceedings of the National Academy of Sciences*, *100*(16), 9383–9387. https://doi.org/10.1073/pnas.1633576100

Bascompte, J., Jordano, P., & Olesen, J. M. (2006). Asymmetric coevolutionary networks facilitate biodiversity maintenance. *Science*, *312*(5772), 431–433. https://doi.org/10.1126/science.1123412

Bastolla, U., Fortuna, M. A., Pascual-García, A., Ferrera, A., Luque, B., & Bascompte, J. (2009). The architecture of mutualistic networks minimizes competition and increases biodiversity. *Nature*, *458*(7241), 1018–1020. https://doi.org/10.1038/nature07950

Bjorbækmo, M. F. M., Evenstad, A., Røsæg, L. L., Krabberød, A. K., & Logares, R. (2020). The planktonic protist interactome: Where do we stand after a century of research? *The ISME Journal*, *14*(2), 544–559. https://doi.org/10.1038/s41396-019-0542-5

Blanchet, F. G., Cazelles, K., & Gravel, D. (2020). Co-occurrence is not evidence of ecological interactions. *Ecology Letters*, *23*(7), 1050–1063. https://doi.org/10.1111/ele.13525

Brimacombe, C., Bodner, K., Michalska-Smith, M., Poisot, T., & Fortin, M.-J. (2023). Shortcomings of reusing species interaction networks created by different sets of researchers. *PLOS Biology*, *21*(4), e3002068. https://doi.org/10.1371/journal.pbio.3002068

Bronstein, J. L. (Ed.). (2015). *Mutualism*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199675654.001.0001

Cai, W., Snyder, J., Hastings, A., & D'Souza, R. M. (2020). Mutualistic networks emerging from adaptive niche-based interactions. *Nature Communications*, *11*(1), 5470. https://doi.org/10.1038/s41467-020-19154-5

Canestrari, D., Bolopo, D., Turlings, T. C. J., Roder, G., Marcos, J. M., & Baglione, V. (2014). From parasitism to mutualism: Unexpected interactions etween a cuckoo and Its host. *Science*, *343*(6177), 1350–1352. https://doi.org/10.1126/science.1249008

Chaffron, S., Delage, E., Budinich, M., Vintache, D., Henry, N., Nef, C., Ardyna, M., Zayed, A. A., Junger, P. C., Galand, P. E., Lovejoy, C., Murray, A. E., Sarmento, H., null, null, Acinas, S. G., Babin, M., Iudicone, D., Jaillon, O., Karsenti, E., … Eveillard, D. (2021). Environmental vulnerability of the global ocean epipelagic plankton community interactome. *Science Advances*, *7*(35), eabg1921. https://doi.org/10.1126/sciadv.abg1921

Chagnon, P.-L. (2016). Seeing networks for what they are in mycorrhizal ecology. *Fungal Ecology*, *24*, 148–154. https://doi.org/10.1016/j.funeco.2016.05.004

Chagnon, P.-L., Bradley, R. L., & Klironomos, J. N. (2015). Trait-based partner selection drives mycorrhizal network assembly. *Oikos*, *124*(12), 1609–1616. https://doi.org/10.1111/oik.01987

Chagnon, P.-L., U'Ren, J. M., Miadlikowska, J., Lutzoni, F., & Arnold, A. E. (2016). Interaction type influences ecological network structure more than local abiotic conditions: Evidence from endophytic and endolichenic fungi at a continental scale. *Oecologia*, *180*(1), 181–191. https://doi.org/10.1007/s00442-015-3457-5

de Lange, S. C., de Reus, M. A., & van den Heuvel, M. P. (2014). The Laplacian spectrum of neural networks. *Frontiers in Computational Neuroscience*, *7*. https://doi.org/10.3389/fncom.2013.00189

Dormann, C. F., Gruber, B., & Fründ, J. (2008). *Introducing the bipartite Package: Analysing Ecological Networks*. *8*, 5.

Fontaine, C., Guimarães, P. R., Kéfi, S., Loeuille, N., Memmott, J., van der Putten, W. H., van Veen, F. J. F., & Thébault, E. (2011). The ecological and evolutionary implications of merging different types of networks: Merging networks with different interaction types. *Ecology Letters*, *14*(11), 1170–1181. https://doi.org/10.1111/j.1461-0248.2011.01688.x

Fontoura, L., Cantor, M., Longo, G. O., Bender, M. G., Bonaldo, R. M., & Floeter, S. R. (2020). The macroecology of reef fish agonistic behaviour. *Ecography*, ecog.05079. https://doi.org/10.1111/ecog.05079

Fortuna, M. A., Ortega, R., & Bascompte, J. (2014). The Web of Life. *arXiv:1403.2575 [q-Bio]*. http://arxiv.org/abs/1403.2575

Fortuna, M. A., Stouffer, D. B., Olesen, J. M., Jordano, P., Mouillot, D., Krasnov, B. R., Poulin, R., & Bascompte, J. (2010). Nestedness versus modularity in ecological networks: Two sides of the same coin? *Journal of Animal Ecology*. https://doi.org/10.1111/j.1365-2656.2010.01688.x

Freilich, S., Kreimer, A., Meilijson, I., Gophna, U., Sharan, R., & Ruppin, E. (2010). The large-scale organization of the bacterial network of ecological co-occurrence interactions. *Nucleic Acids Research*, *38*(12), 3857–3868. https://doi.org/10.1093/nar/gkq118

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, *33*(1). https://doi.org/10.18637/jss.v033.i01

Genini, J., Morellato, L. P. C., Guimarães, P. R., & Olesen, J. M. (2010). Cheaters in mutualism networks. *Biology Letters*, *6*(4), 494–497. https://doi.org/10.1098/rsbl.2009.1021

Guimarães, P. R., Rico-Gray, V., Oliveira, P. S., Izzo, T. J., dos Reis, S. F., & Thompson, J. N. (2007). Interaction intimacy affects structure and coevolutionary dynamics in mutualistic networks. *Current Biology*, *17*(20), 1797–1803. https://doi.org/10.1016/j.cub.2007.09.059

Günther, F., & Fritsch, S. (2010). neuralnet: Training of Neural Networks. *The R Journal*, *2*(1), 30. https://doi.org/10.32614/RJ-2010-006

Hembry, D. H., Raimundo, R. L. G., Newman, E. A., Atkinson, L., Guo, C., Guimarães, P. R., & Gillespie, R. G. (2018). Does biological intimacy shape ecological network structure? A test using a brood pollination mutualism on continental and oceanic islands. *Journal of Animal Ecology*, *87*(4), 1160–1171. https://doi.org/10.1111/1365-2656.12841

Hoeppke, C., & Simmons, B. (2021). Maxnodf: An R package for fair and fast comparisons of nestedness between networks. *Methods in Ecology and Evolution*, *12*. https://doi.org/10.1111/2041-210x.13545

Jacquemyn, H., Brys, R., Waud, M., Busschaert, P., & Lievens, B. (2015). Mycorrhizal networks and coexistence in species-rich orchid communities. *New Phytologist*, *206*(3), 1127–1134. https://doi.org/10.1111/nph.13281

Jacquemyn, H., Merckx, V., Brys, R., Tyteca, D., Cammue, B. P. A., Honnay, O., & Lievens, B. (2011). Analysis of network architecture reveals phylogenetic constraints on mycorrhizal specificity in the genus Orchis (Orchidaceae). *New Phytologist*, *192*(2), 518–528. https://doi.org/10.1111/j.1469-8137.2011.03796.x

Jordano, P., Bascompte, J., & Olesen, J. M. (2002). Invariant properties in coevolutionary networks of plant-animal interactions: Invariant properties in coevolutionary networks. *Ecology Letters*, *6*(1), 69–81. https://doi.org/10.1046/j.1461-0248.2003.00403.x

Klironomos, J. N. (2003). Variation in plant response to native and exotic arbuscular mycorrhizal fungi. *Ecology*, *84*(9), 2292–2301. https://doi.org/10.1890/02-0413

Krause, A. E., Frank, K. A., Mason, D. M., Ulanowicz, R. E., & Taylor, W. W. (2003). Compartments revealed in food-web structure. *Nature*, *426*(6964), 282–285. https://doi.org/10.1038/nature02115

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR : an R package for multivariate analysis. *Journal of Statistical Software*, *25*(1). https://doi.org/10.18637/jss.v025.i01

Lewitus, E., & Morlon, H. (2016). Characterizing and comparing phylogenies from their Laplacian spectrum. *Systematic Biology*, *65*(3), 495–507. https://doi.org/10.1093/sysbio/syv116

Lima-Mendez, G., Faust, K., Henry, N., Decelle, J., Colin, S., Carcillo, F., Chaffron, S., Ignacio-Espinosa, J. C., Roux, S., Vincent, F., Bittner, L., Darzi, Y., Wang, J., Audic, S., Berline, L., Bontempi, G., Cabello, A. M., Coppola, L., Cornejo-Castillo, F. M., … Raes, J. (2015). Determinants of community structure in the global plankton interactome. *Science*, *348*(6237), 1262073–1262073. https://doi.org/10.1126/science.1262073

Losapio, G., Schöb, C., Staniczenko, P. P. A., Carrara, F., Palamara, G. M., De Moraes, C. M., Mescher, M. C., Brooker, R. W., Butterfield, B. J., Callaway, R. M., Cavieres, L. A., Kikvidze, Z., Lortie, C. J., Michalet, R., Pugnaire, F. I., & Bascompte, J. (2021). Network motifs involving both competition and facilitation predict biodiversity in alpine plant communities. *Proceedings of the National Academy of Sciences*, *118*(6), e2005759118. https://doi.org/10.1073/pnas.2005759118

Maliet, O., Loeuille, N., & Morlon, H. (2020). An individual-based model for the eco-evolutionary emergence of bipartite interaction networks. *Ecology Letters*, ele.13592. https://doi.org/10.1111/ele.13592

Martos, F., Munoz, F., Pailler, T., Kottke, I., Gonneau, C., & Selosse, M.-A. (2012). The role of epiphytism in architecture and evolutionary constraint within mycorrhizal networks of tropical orchids. *Molecular Ecology*, *21*(20), 5098–5109. https://doi.org/10.1111/j.1365-294X.2012.05692.x

Mello, M. A. R., Marquitti, F. M. D., Guimarães, P. R., Kalko, E. K. V., Jordano, P., & de Aguiar, M. A. M. (2011). The missing part of seed dispersal networks: Structure and robustness of bat-fruit interactions. *PLoS ONE*, *6*(2), e17395. https://doi.org/10.1371/journal.pone.0017395

Michalska-Smith, M. J., & Allesina, S. (2019). Telling ecological networks apart by their structure: A computational challenge. *PLOS Computational Biology*, *15*(6), e1007076. https://doi.org/10.1371/journal.pcbi.1007076

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., & Alon, U. (2002). Network motifs: Simple building blocks of complex networks. *Science*, *298*(5594), 824–827.

Montesinos-Navarro, A., Hiraldo, F., Tella, J. L., & Blanco, G. (2017). Network structure embracing mutualism–antagonism continuums increases community robustness. *Nature Ecology & Evolution*, *1*(11), 1661–1669. https://doi.org/10.1038/s41559-017-0320-6

Montesinos-Navarro, A., Segarra-Moragues, J. G., Valiente-Banuet, A., & Verdú, M. (2012). Plant facilitation occurs between species differing in their associated arbuscular mycorrhizal fungi. *New Phytologist*, *196*(3), 835–844. https://doi.org/10.1111/j.1469-8137.2012.04290.x

Morlon, H., Lewitus, E., Condamine, F. L., Manceau, M., Clavel, J., & Drury, J. (2016). RPANDA : an R package for macroevolutionary analyses on phylogenetic trees. *Methods in Ecology and Evolution*, *7*(5), 589–597. https://doi.org/10.1111/2041-210X.12526

Newman, M. E. J. (2006a). Finding community structure in networks using the eigenvectors of matrices. *Physical Review E*, *74*(3), 036104. https://doi.org/10.1103/PhysRevE.74.036104

Newman, M. E. J. (2006b). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, *103*(23), 8577–8582. https://doi.org/10.1073/pnas.0601602103

Okuyama, T., & Holland, J. N. (2008). Network structural properties mediate the stability of mutualistic communities. *Ecology Letters*, *11*(3), 208–216. https://doi.org/10.1111/j.1461-0248.2007.01137.x

Olesen, J. M., Bascompte, J., Dupont, Y. L., & Jordano, P. (2007). The modularity of pollination networks. *Proceedings of the National Academy of Sciences*, *104*(50), 19891–19896. https://doi.org/10.1073/pnas.0706375104

Öpik, M., Metsis, M., Daniell, T. J., Zobel, M., & Moora, M. (2009). Large-scale parallel 454 sequencing reveals host ecological group specificity of arbuscular mycorrhizal fungi in a boreonemoral forest. *New Phytologist*, *184*(2), 424–437. https://doi.org/10.1111/j.1469-8137.2009.02920.x

Perez-Lamarque, B., Petrolli, R., Strullu-Derrien, C., Strasberg, D., Morlon, H., Selosse, M.-A., & Martos, F. (2022). Structure and specialization of mycorrhizal networks in phylogenetically diverse

tropical communities. *Environmental Microbiome*, *17*(1), 38. https://doi.org/10.1186/s40793-022-00434-0

Perez-Lamarque, B., Selosse, M., Öpik, M., Morlon, H., & Martos, F. (2020). Cheating in arbuscular mycorrhizal mutualism: A network and phylogenetic analysis of mycoheterotrophy. *New Phytologist*, *226*(6), 1822–1835. https://doi.org/10.1111/nph.16474

Pilosof, S., Fortuna, M. A., Cosson, J.-F., Galan, M., Kittipong, C., Ribas, A., Segal, E., Krasnov, B. R., Morand, S., & Bascompte, J. (2014). Host–parasite network structure is associated with community-level immunogenetic diversity. *Nature Communications*, *5*(1), 5172. https://doi.org/10.1038/ncomms6172

Pires, M. M., & Guimarães, P. R. (2013). Interaction intimacy organizes networks of antagonistic interactions in different ways. *Journal of The Royal Society Interface*, *10*(78), 20120649. https://doi.org/10.1098/rsif.2012.0649

Põlme, S., Bahram, M., Jacquemyn, H., Kennedy, P., Kohout, P., Moora, M., Oja, J., Öpik, M., Pecoraro, L., & Tedersoo, L. (2018). Host preference and network properties in biotrophic plant-fungal associations. *New Phytologist*, *217*(3), 1230–1239. https://doi.org/10.1111/nph.14895

Rezende, E. L., Jordano, P., & Bascompte, J. (2007). Effects of phenotypic complementarity and phylogeny on the nested structure of mutualistic networks. *Oikos*, *116*(11), 1919–1929. https://doi.org/10.1111/j.0030-1299.2007.16029.x

Rezende, E. L., Lavabre, J. E., Guimarães, P. R., Jordano, P., & Bascompte, J. (2007). Non-random coextinctions in phylogenetically structured mutualistic networks. *Nature*, *448*(7156), 925–928. https://doi.org/10.1038/nature05956

Rohr, R. P., Saavedra, S., & Bascompte, J. (2014). On the structural stability of mutualistic systems. *Science*, *345*(6195), 1253497–1253497. https://doi.org/10.1126/science.1253497

Sachs, J. L., Skophammer, R. G., & Regus, J. U. (2011). Evolutionary transitions in bacterial symbiosis. *Proceedings of the National Academy of Sciences*, *108*(Supplement_2), 10800–10807. https://doi.org/10.1073/pnas.1100304108

Santamaría, L., & Rodríguez-Gironés, M. A. (2007). Linkage Rules for Plant–Pollinator Networks: Trait Complementarity or Exploitation Barriers? *PLoS Biology*, *5*(2), e31. https://doi.org/10.1371/journal.pbio.0050031

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

Simmons, B. I., Cirtwill, A. R., Baker, N. J., Wauchope, H. S., Dicks, L. V., Stouffer, D. B., & Sutherland, W. J. (2019). Motifs in bipartite ecological networks: Uncovering indirect interactions. *Oikos*, *128*(2), 154–170. https://doi.org/10.1111/oik.05670

Simmons, B. I., Sweering, M. J. M., Schillinger, M., Dicks, L. V., Sutherland, W. J., & Clemente, R. D. (2018). *bmotif: A package for motif analyses of bipartite networks* [Preprint]. Ecology. https://doi.org/10.1101/302356

Simmons, B. I., Sweering, M. J. M., Schillinger, M., Dicks, L. V., Sutherland, W. J., & Di Clemente, R. (2019). bmotif: A package for motif analyses of bipartite networks. *Methods in Ecology and Evolution*, *10*(5), 695–701. https://doi.org/10.1111/2041-210X.13149

Song, C., & Saavedra, S. (2020). Telling ecological networks apart by their structure: An environment-dependent approach. *PLOS Computational Biology*, *16*(4), e1007787. https://doi.org/10.1371/journal.pcbi.1007787

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, *15*(56), 1929–1958.

Stouffer, D. B., & Bascompte, J. (2010). Understanding food-web persistence from local to global scales. *Ecology Letters*, *13*(2), 154–161. https://doi.org/10.1111/j.1461-0248.2009.01407.x

Strydom, T., Catchen, M. D., Banville, F., Caron, D., Dansereau, G., Desjardins-Proulx, P., Forero-Muñoz, N. R., Higino, G., Mercier, B., Gonzalez, A., Gravel, D., Pollock, L., & Poisot, T. (2021). A roadmap towards predicting species interaction networks (across space and time). *Philosophical*

Transactions of the Royal Society B: Biological Sciences, *376*(1837), 20210063. https://doi.org/10.1098/rstb.2021.0063

Suweis, S., Simini, F., Banavar, J. R., & Maritan, A. (2013). Emergence of structural and dynamical properties of ecological mutualistic networks. *Nature*, *500*(7463), 449–452. https://doi.org/10.1038/nature12438

Thébault, E., & Fontaine, C. (2008). Does asymmetric specialization differ between mutualistic and trophic networks? *Oikos*, *117*(4), 555–563. https://doi.org/10.1111/j.0030-1299.2008.16485.x

Thébault, E., & Fontaine, C. (2010). Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science*, *329*(5993), 853–856. https://doi.org/10.1126/science.1188321

Thebault, E., & Fontaine, C. (2010). Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science*, *329*(5993), 853–856. https://doi.org/10.1126/science.1188321

Thomas, F., Renaud, F., & Guegan, J.-F. (Eds.). (2005). *Parasitism and Ecosystems*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198529873.001.0001

Thompson, J. N. (2005). *The Geographic Mosaic of Coevolution*. University of Chicago Press. https://doi.org/10.7208/9780226118697

Toju, H., Guimarães, P. R., Olesen, J. M., & Thompson, J. N. (2015). Below-ground plant–fungus network topology is not congruent with above-ground plant–animal network topology. *Science Advances*, *1*(9), e1500291. https://doi.org/10.1126/sciadv.1500291

Vázquez, D. P., Chacoff, N. P., & Cagnolo, L. (2009). Evaluating multiple determinants of the structure of plant–animal mutualistic networks. *Ecology*, *90*(8), 2039–2046. https://doi.org/10.1890/08-1837.1

Vitali, A., Sasal, Y., Vázquez, D. P., Miguel, M. F., & Rodríguez-Cabal, M. A. (2022). The disruption of a keystone interaction erodes pollination and seed dispersal networks. *Ecology*, *103*(1). https://doi.org/10.1002/ecy.3547

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2021). A comprehensive survey on graph neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, *32*(1), 4–24. https://doi.org/10.1109/TNNLS.2020.2978386

Xing, X., Jacquemyn, H., Gai, X., Gao, Y., Liu, Q., Zhao, Z., & Guo, S. (2019). The impact of life form on the architecture of orchid mycorrhizal networks in tropical forest. *Oikos*, *128*(9), 1254–1264. https://doi.org/10.1111/oik.06363