

The non-random assembly of functional motifs in plant-pollinator networks worldwide

Abstract

Ecological processes leave distinct structural imprints on species interactions shaping the topology of mutualistic networks. Detecting those relationships is not trivial since they go beyond pairwise interactions, but may get blurred when considering full network descriptors. However, recent work has shown the network meso-scale can capture this important information. The meso-scale describes network subgraphs representing patterns of interactions between a small number of species (i.e. motifs) and those constitute the building blocks of the whole network. Here, we have compiled 60 networks from 18 different studies and show that some motifs are consistently under- and over-represented worldwide, suggesting that the building blocks of plant-pollinator networks are not random. Furthermore, we found that distinct motif positions describing species ecological roles (e.g., generalization, indirect interactions) are occupied by different functional groups on both trophic levels. For instance, bees appear consistently in generalist motif positions and plants with higher levels of outcrossing tend to have a large number of indirect interactions. Interestingly, the realized combinations of different functional groups within a motif can not be retrieved from their joint probability distributions, indicating that functional group combinations are not random either. Our result highlight the non-random structure of indirect interactions in plant-pollinator networks and opens the door to predict novel interactions beyond pairwise interactions within a community context.

Keywords: plant-pollinator networks, plant-pollinator interactions, network motifs, functional groups, plant traits, floral visitors

Introduction

Ecological communities are formed by a plethora of interacting species that form networks of interactions. Because of the tremendous complexity of these ecological networks, species interactions are generally condensed in metrics that summarize this information (Guimarães 2020). Plant-pollinator interactions are no exception, and they are often studied with indices that aggregate the information at the species (micro-scale) or network level (macro-scale). For example, plant-pollinator network level approaches have identified common invariant structural properties across networks, including a degree distribution that decays as a power law (Jordano 1987), nestedness (Bascompte et al. 2003), or modularity (Olesen et al. 2007). In addition, the local species position within the network can define its degree of specialization (Blüthgen et al. 2006) or its role connecting the rest of the community (Olesen et al. 2007). Despite the unquestionable progress with the use of these metrics at both species and network level, condensing the information into a single metric implies the loss of relevant ecological information that obscure the understanding of the ecological processes (Cirtwill et al.

2018, Simmons et al. 2019a).

Traditionally, plant-pollinator research has focused on direct interactions but overlooked indirect interactions (i.e., the mediated effect between two species by a third species), such as facilitative or competitive interactions between plants for pollinators (Moeller 2004, Sargent and Ackerly 2008). Despite the widespread nature of indirect interactions in ecological communities (Strauss 1991), plant-pollinator research often fails to finely capture those indirect interactions with the conventional analytical tools that condense the information either by species (e.g., interaction frequency) or in global topological indices (e.g., nestedness). Nonetheless, the emerging framework of network motifs (small sub-networks recorded within a given network) in plant-pollinator research allows to consider both direct and indirect interactions (Simmons et al. 2019a). Motifs are often referred to as the building blocks of a network that depict subsets of interactions (Milo et al. 2002). The analysis of network motifs in plant-pollinator networks have revealed that the different ecological processes that govern species interactions (e.g., species abundances versus trait-matching) can lead to different patterns of indirect interactions (Simmons et al. 2020). Yet, the global patterns of indirect interactions in real plant-pollinator networks are still unknown (e.g., over- and under- represented motifs).

Despite motifs are abstract representations often decoupled from species ecology, linking the structural properties of the meso-scale with the species ecology can help the understanding of ecological processes. For instance, different motifs can have different ecological meanings (Simmons et al. 2019a) and the position within a motif can determine the species functional role (Stouffer et al. 2012, Baker et al. 2015). However, it is unclear how the species ecology and life-history traits determines the species' functional role within the network of interactions (Coux et al. 2016). For example, large pollinators can forage larger distances (Greenleaf et al. 2007), deposit greater pollen quantities (Földesi et al. 2021) and handle complex zygomorphic flowers in comparison with small pollinators that are restricted to lower floral complexity (Gong and Huang 2009). How this different pollinator behavior translates into their interaction topology is unknown. Similarly, recent empirical findings indicate that the meso-scale is the best descriptor of plant reproductive success (Allen-Perkins et al. 2021), but little is known about how plants reproductive strategies shape their position within the network of interactions. Although some studies have evaluated plant reproductive strategies in plant-pollinator networks (Tur et al. 2013, Lázaro et al. 2020), this is an often overlooked aspect in a community context (Devaux et al. 2014) and rarely incorporated into plant-pollinator network studies. Hence, exploring how the main plant reproductive strategies integrate with the emergent motif framework can shed light on key aspects of ecosystem functioning.

Here, we used 60 plant-pollinator networks from 18 different studies and 14 countries, alongside functional groups of plants and taxonomic groups of floral visitors based on functional traits and taxonomic rank,

respectively. To obtain plant functional groups, we used a comprehensive dataset that included floral, reproductive and vegetative traits compiled in a larger set of networks. Floral visitor groups were divided by the main taxonomic groups that differed in life form and behaviour. Once we split the different plant-pollinator networks into their motif elements, we explored: (i) if there is a common invariant structural property in the overall motif networks (i.e. over- and under-represented motifs); (ii) which plant and pollinator groups are over- and under-represented in the different motif positions; and, (iii) if there are over- and under-represented plant and pollinator group combinations within a motif.

Methods

Plant-pollinator studies

We have compiled 60 plant-pollinator networks from 18 different studies (**Table S1**). All studies sampled plant-pollinator interactions in natural systems and were selected based on wide geographical coverage and presence of interaction frequency as a measure of interaction strength. In total, there were 503 plant species, 1,111 floral visitors species and 6,248 pairwise interactions registered. For ease of data manipulation, plant and floral visitor species names were standardized with the help of the R package *taxize* version *0.9.99* (Chamberlain et al. 2020). All analyses and data manipulation were conducted in R *version 4.0.5* (R Core Team 2021).

Plant and floral visitor groups

First, plant species were grouped through hierarchical clustering into the optimal number of functional groups that summarized the main plant reproductive strategies. For this, we used the trait dataset collated in Lanuza et al. (2021) that comprised a total of 1506 species including the 503 species considered in this study (**Table S2**). This dataset consisted on 8 floral, 4 reproductive and 3 vegetative traits excluding traits with high percentage of missing values (over 30%; **Table S3**). We opted to calculate the plant functional groups on this larger set of species because of the higher accuracy when delimiting functional groups with that many variables and species (Dolnicar et al. 2014). To feed the clustering analysis, we calculated the distance between the different qualitative and scaled quantitative variables with Gower distance (Gower 1971). For this, we used the function *gowdis* with method *ward.D2* from the package *FD* version *1.0-12* (Laliberté et al. 2014). Finally, we conducted hierarchical clustering with the function *hclust* from the *R* stats package version *4.0.5* and calculated the optimal number of clusters with the function *kgs* from the package *maptree* version *1.4-7* (White and Gramacy 2009).

Second, floral visitors were grouped into taxonomic groups based on taxonomic rank as done similarly in past

plant-pollinator studies (Fenster et al. 2004, Ollerton et al. 2009). We opted to divide floral visitors on the taxonomic rank level and not with functional traits because the main orders of floral visitors differed in form and behaviour and had lower superior taxonomical complexity (i.e., floral visitors had 6 orders versus plants that had 38). Thus, this allowed us to group floral visitors into groups that represented approximately the main life strategies of the possible pollinators: (i) bees (Hymenoptera-Anthophila), (ii) non-bee Hymenoptera (Hymenoptera-non-Anthophila), (iii) syrphids (Syrphidae-Diptera), (iv) non-syrphids-Diptera, (v) Lepidoptera and (vi) Coleoptera. However, a minor set of species belonged to other taxonomic groups that were considered in analyses but not discussed further because of their low representation in the full set of networks (3.55% of the total interactions recorded). These taxonomic groups were ‘lizards,’ ‘birds’ and ‘other insects.’ This last group was formed by a mix of uncommon insect taxa on the full set of networks.

Overall motif patterns

Following previous work (Simmons et al. 2019a, 2020), we broke down the plant-pollinator networks into their constituent motifs. Prior to analyses, we turned the quantitative networks into qualitative (or binary) ones, where interactions are present or absent. To test the effect of singletons, we ran all analysis to the same set of networks, but considering only interactions whose frequency was greater than one (64.98% of interactions). Results are qualitatively consistent and are not further discussed in the main text (**Figures S1 and S2**).

We calculated the frequency of all motifs up to five nodes (see **Figure 1**) for each empirical network, by using the *bmotif* package version 2.0.2 (Simmons et al. 2019b). To control for variation in network size, motif frequencies were normalised as a proportion of the total number of motifs within each motif class (i.e., the number of nodes a motif contains). Like Simmons et al. (2020), we classified the different motifs by their average path length into four different groups: (i) complete or strong ($\bar{x} = 1.38$), (ii) fan or medium-strong ($\bar{x} = 1.48$), (iii) asymmetric complete or medium-weak ($\bar{x} = 1.60$) and (iv) core-peripheral or weak ($\bar{x} = 1.85$). Moreover, we just used five-node motifs in our analyses for visualisation, interpretation and computational reasons. Finally, we also excluded two-node motifs (or links) from our analyses because their normalised frequencies would always equal one.

MA FORTUNA, what about degree?

To assess the significance of the observed frequencies, we created 1,000 simulated networks for each binary network using the *nullmodel* function and the *vaznull* model in the *bipartite* package version 2.14 (Dormann et al. 2009). Generated networks had the same number of plants and floral visitors, as well as the same connectance of their corresponding empirical networks. After extracting the motif frequencies from the simulated networks, for each motif type and empirical network, we calculated the percentage of simulated



Figure 1. Adapted figure of Simmons et al., 2019 with all the possible motifs from two to five species in bipartite networks. There is a total of 17 possible motifs with 46 different positions denoted within each node.

networks whose frequencies were smaller than the ones observed, that is, we estimated the percentile of the observed motif frequencies. Motifs whose percentile is close to 0 or 100 are under- or over-represented in the empirical networks, respectively, and thus they cannot be predicted by connectance and the number of species alone. To summarize general patterns across networks, we used an intercept-only linear mixed model (LMM) per motif with the help of the package *lmer* version *1.1-21* (Bates et al. 2015), where the response variable was the observed motif percentile per network. In these models, we used the study identifiers in **Table S1** as a random intercept. By doing so, we obtained estimates of the average motif frequency, but controlling the variation at the study level.

Over- and under-represented groups on motif positions

We calculated which plant and floral visitor groups were over- or under-represented in the different motif positions by comparing position frequencies of empirical networks with those of their corresponding simulated counterparts. To estimate the position frequencies of each group in a given network, we added the frequencies of those species that belong to the group, and then, we normalised the resulting frequencies by dividing the position measure for each group by the total number of times that a group appears in any position within the same motif size class. Then, we calculated the percentile of the observed position frequencies for each group and network, just like we did with motif frequencies. To outline the general patterns of position

frequencies across networks and groups, we fit a LMM per motif position, where the response variable was the observed position percentile per network. We used the group identifier as an explanatory variable and the study identifiers as a random intercept. By adding the group estimates to the model intercept, we assessed the average motif frequency, after controlling the variation at the study level. Finally, we visualized with the help of the package *ComplexHeatmap* version 2.6.2 (Gu et al. 2016) over- and under-representation of plant and floral visitor groups on the different motif positions.

To understand if the different plant and floral visitor groups are associated with motif positions that involve different network roles (e.g., specialization or generalism), we calculated the specificity of the different motif positions and evaluated its association through linear models with the average frequency of the different plant and floral visitor groups as response variable. For this, we previously calculated the specificity (s_i) of each motif position following Poisot et al. (2015) as $s_i = L - l_i/L - 1$, where L is the total number of groups and l_i the number of interaction partners of the group i in a given motif. Values of the vector s , range between 1, complete specialization (1 partner within the motif) and 0, complete generalism (all possible partners within the motif).

Over- and under-represented group combinations of motifs

We studied which combinations of plant and floral visitor groups tend to appear together within the same motifs (up to five nodes). That is which combinations are over- or under-represented. This analysis use 57 out of the 60 networks available due to computational limitations to identify all the nodes in the motifs of the three networks with the highest number of links. To do so, for each of the 53,250 possible motif combinations, we estimated the observed and the expected probability of finding those combination in empirical networks, respectively. Then, we determined whether the observed probabilities are likely to come from the expected probabilities or not. To calculate the observed probability (p_i^O) of the different groups within a motif i (e.g., motif 3, ‘bee’ + ‘bee’ + ‘selfing herbs’), we divided the number of times that i appears in our set of empirical networks (n_i^O) by the sum of the number of times that each possible combination appeared: $p_i^O = n_i^O / \sum_{k=1}^{53,250} n_k^O$. To estimate the expected probability of a given motif combination i , p_i^E , firstly, we calculated the probability of finding a given group x in the position α of i , $p_i(x, \alpha)$. Then, by assuming the independency of $p_i(x, \alpha)$, we computed the expected probability of the combination i as the product of the probability of its pairs (x, α) , that is, $p_i^E = \prod_{(x, \alpha) \in i} p_i(x, \alpha)$. To obtain $p_i(x, \alpha)$, we proceeded as follows. First, we calculated the number of times that the group x appears in the position α in each empirical network η , denoted as $n_i^\eta(x, \alpha)$. Then, by controlling the variation at the network level, we assessed the average value of the absolute frequency of group x appears in the position α , $n_i(x, \alpha) = E[n_i^\eta(x, \alpha)]$, and estimated $p_i(x, \alpha)$ as $p_i(x, \alpha) = n_i(x, \alpha) / \sum_k n_i(k, \alpha)$. To obtain $n_i(x, \alpha)$, we fitted a LMM per motif position, where

the response variable was the number of times that a given position was observed per network, the explanatory variable was the group identifier, and the random intercept was given by network identifiers nested within the study identifiers.

Once we obtained p_i^O and p_i^E , we used a simulation approach to determine whether the former is likely to come from the latter or not. This approach was preferred since the large number of possible motif combinations and the small probabilities for some of them advise against using an exact test of goodness-of-fit or a Chi-square one. Specifically, we created 1,000 random samples with repetition of possible motif combinations, where each sample contained 10 million elements and, for each combination, the probability of being selected was equal to its expected probability. From those random samples, we extracted the mean and the standard deviation of the expected probability of i , denoted as $E[p_i^E]$ and $\sigma[p_i^E]$, respectively, and calculated the z-scores of p_i^O as $z_i^O = (p_i^O - E[p_i^E]) / \sigma[p_i^E]$, for those motif combinations with $p_i^O > 0$. According to the usual interpretation of z-scores, combinations with $z_i^O > 1.96$ are over-represented, whereas those with $z_i^O < -1.96$ are under-represented, at the 95% confidence level. Notice that we focused on combinations with $p_i^O > 10^{-7}$ because we do not have enough numerical resolution to accurately detect whether combinations that took place only once are under-represented (due to the limited size of our random samples).

RESULTS

Plant and floral visitor groups

The hierarchical cluster analysis divided the dataset into five different groups with different and overlapping characteristics (**Figure S3** and **Figure S4**). The first group referred to as ‘selfing herbs’ consisted on herbs with hermaphroditic flowers with high levels of autonomous selfing. The second group named ‘small outcrossing perennials’ had small perennial species with a mixed of life forms (i.e., trees shrubs and herbs) with outcrossing hermaphroditic flowers. The third group referred to as ‘self-incompatible perennials with large flowers’ comprised perennial species with a mixed of life forms and large self-incompatible hermaphroditic flowers with high number of ovules. The fourth group named ‘tall plants with small unisexual flowers’ had the tallest species, highest proportion of shrub and tree life forms, dioecious and monoecious breeding systems, small flowers and the highest numbers of flowers and inflorescences per plant. Finally, the fifth group named ‘short-lived outcrossers with long zygomorphic flowers’ consisted on small perennial and short-lived herbs with long self-compatible zygomorphic flowers that were unable to self-pollinate.

In total, there were 1,126 species of floral visitors with 6,325 interactions recorded with plants. Most plants interacted with bees (2,256 interactions) and non-syrphid Diptera (1,768) followed by syrphids (845),

Lepidoptera (437), Coleoptera (432) and non-bee Hymenoptera (362).

Overall motif patterns

Most motifs were under- and over-represented (close to the 1st and 99th percentile, respectively) in the comparison between empirical and simulated networks (**Figure 2**). Motifs 3, 5, 9, 10 and 14 were under-represented in empirical networks, that is, all were close to the 1st percentile and under the 25th percentile. Interestingly, four out of five of these motifs belonged to the largest path length classification (i.e., core-peripheral). In addition, motifs 2, 6, 7, 16 and 17 were over-represented, all over the 75th percentile and thus the closest to the 99th percentile. In contrast to the under-represented motifs, over-represented motifs belonged to the two shortest path length groups (i.e., complete and fan). The remaining motifs (i.e., 4, 8, 11, 12, 13, and 15) were between the 25th and 75th percentile. Notably, the exclusion of singletons from this analysis did not change the observed results (**Figure S1**).

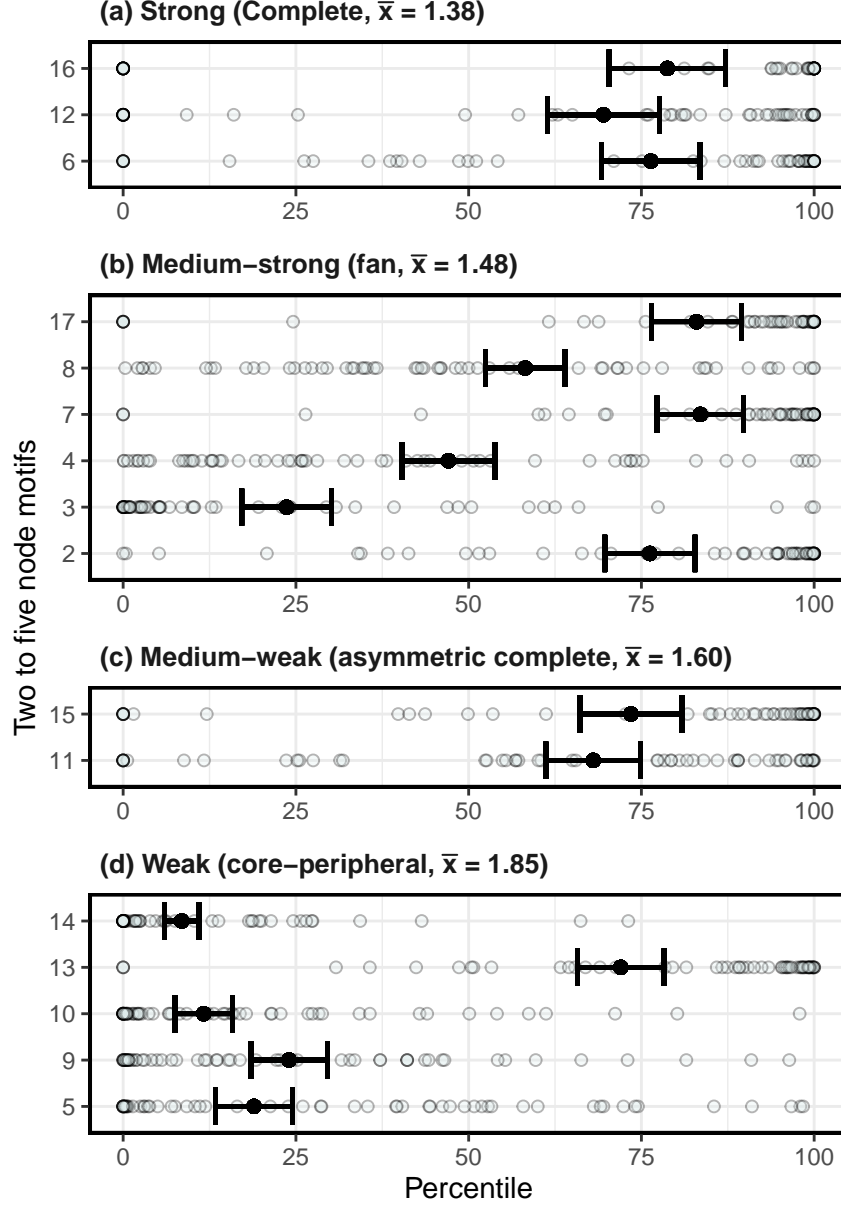


Figure 2. Comparison of motif frequencies between empirical and simulated networks grouped by average path length (plots a, b, c and d) as determined in Simmons et al. (2020). This is shown with the mean percentage of motif frequencies in empirical networks that were over the motif frequencies of the simulated ones (percentiles). This was done by network (light blue dots) and then averaged for all networks (black dots with error bars that correspond to the standard deviation).

Over- and under-represented groups on motif positions

The comparison of the plant and floral visitor group frequencies per motif position between empirical and simulated networks showed a 6.36% and 25.38% of over- ($>75^{th}$) and under-represented ($<25^{th}$) groups in the different positions, respectively (**Figure 3**). Floral visitors showed a total of 4.55% and 32.58% of over- and under-represented groups in the different positions and plants 8.18% and 18.18% of over- and

under-represented groups, respectively. Notably, the differences across groups were more marked for floral visitors than for plant groups (the differences between min and max percentiles per position were generally two-three times larger for floral visitors). From most over- to under-represented floral visitors groups on the different motif positions (indicated with the dendrogram order in **Figure 3**), we found: bees, non-syrphids Diptera, syrphids, Coleoptera, non-bee Hymenoptera and Lepidoptera. Although plant functional groups showed less differences between them, there were also more represented groups than others, thus from most over- to under-represented groups on the different motif positions we found: self incompatible perennials with large flowers, small outcrossing perennials, tall plants with unisexual flowers, selfing herbs and short lived outcrossers with long zygomorphic flowers.

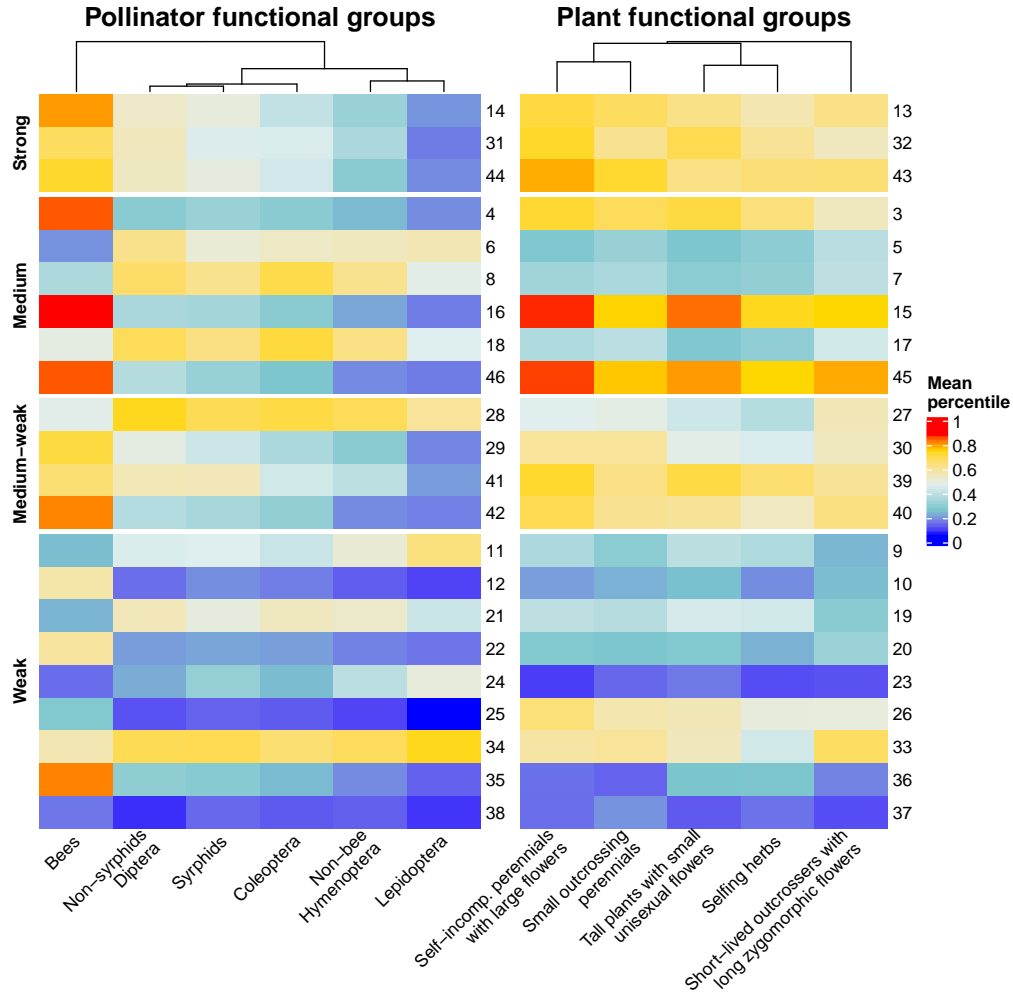


Figure 3. Heatmap indicating under- and over-representation of floral visitor and plant groups in the different motif positions. The different motif positions are divided by the average path length clasification determined by Simmons et al. (2020) and they are ordered from the most densely connected motifs (strong or complete) to the less connected ones (weak or core-peripheral). The superior dendrogram indicates the differences across groups with the more separated groups showing larger differences.

Over- and under-represented group combinations of motifs

The different observed frequencies on the different motif positions showed a hierarchical order of probabilities for floral visitors (**Figure 4**). That is, the taxonomic group of ‘bees’ was the most abundant one on all motif positions followed by non-syrphid Diptera which was after bees, the second most frequent group on all positions. Then, the group of syrphids was the most abundant in all positions but two (positions 28 and 31 from motifs 11 and 12, respectively). After syrphids, Coleoptera was the most frequent group on the different positions followed by non-bee Hymenoptera and finally, Lepidoptera which was the less frequent group on all positions but two (positions 12 and 22 from motifs 5 and 9, respectively). Remarkably, plant functional groups were more variable on the different motif positions. However, there were also predominant functional groups found across the different motif positions. The three most frequent groups were ‘tall plants with small unisexual flowers,’ ‘self-incompatible perennials with large flowers’ and ‘small outcrossing perennials.’ On the contrary, the groups of ‘selfing herbs’ and ‘short lived outcrossers with long zygomorphic flowers’ had the lowest probability to be present on the different motifs plant positions.

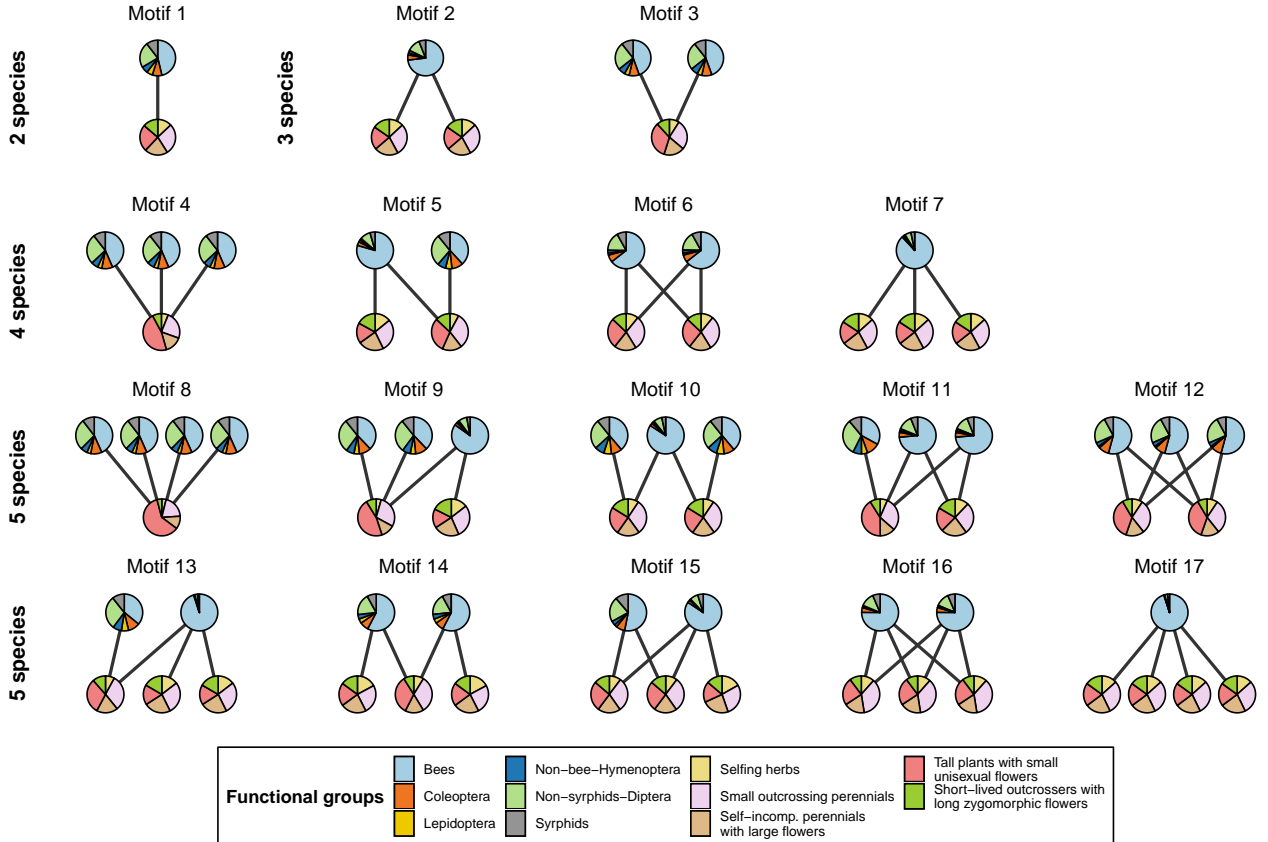


Figure 4. Graphical representation of the probability of finding a given plant and floral visitor group x in the position α of motif i , $p_i(x, \alpha)$, for all the possible motifs from two to five species in bipartite networks. The slices in the nodes for a given functional group x are proportional to the corresponding value of $p_i(x, \alpha)$.

The statistical comparison with z-scores between the observed and expected plant and floral visitor group combinations showed that we can not recover the observed combinations from its probability of occurrence. In fact, most combinations are under- or over- represented and follow a modified Gaussian distribution (Johnson's S_U) with 65% of under-represented group combinations, 20% of no statistical difference and 15% of under-represented ones (**Figure 5**). In addition, motifs with small node combinations (i.e., 2 and 3 nodes) only appeared as under-represented. Motifs with 4 and 5-node combinations appeared in the three statistical categories but 4 node motifs had a higher proportion in the under-represented category (10% in comparison with the 1% and the 3% of the no statistical difference and over-represented categories, respectively). Finally, 5-node motifs were in similar proportions in the no statistical difference and over-represented categories (99% and 97%, respectively) but they appeared in lower proportion in the under-represented category (89%). When, we explore the identity on the different positions of the most probable motif combinations we find as expected that the most common plant and floral visitor groups rule these positions on the most probable motif combinations (lower panel **Figure 5**).

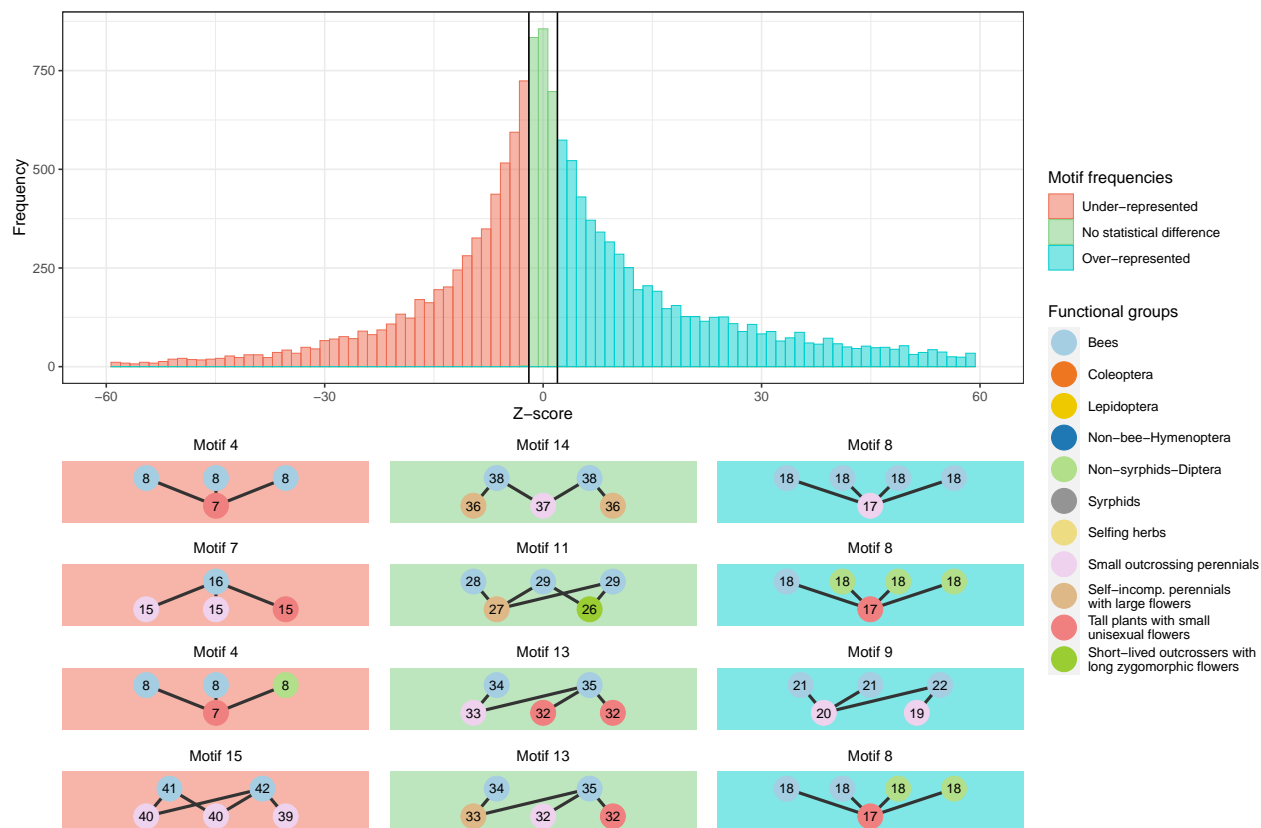


Figure 5. Z-score comparison (95% confidence level) of the observed and expected frequencies of motif combinations with four specific graphical examples of the most probable motifs of being observed for each category delimited by the z-score critical values: under-represented (red), no statistical difference (green) and over-represented (blue).

Discussion

Understanding the global patterns of network structure is key to unravel the processes that govern community dynamics (Bascompte and Jordano 2007, Bastolla et al. 2009, Guimarães 2020). However, common analytical approaches are unable to capture the full complexity of ecological systems (Simmons et al. 2019a, Bergamo et al. 2021). Because network motifs can capture both direct and indirect interactions, by analysing over- and under-representation of motifs we are able to explore in greater detail ecological interactions (Stouffer et al. 2007). We have explored for the first time the overall structure and representation of network motifs in real plant-pollinator networks and shown that the different motifs up to five nodes are consistently over- and under-represented in comparison with simulated networks indicating the presence of ecological processes shaping their structure. Despite the presence of multiple ecological processes likely operating at the same time (e.g., neutral effects, morphological matching or phenological overlap), we found a tendency for over-represented structures with more densely connected motifs (i.e., complete, fan and asymmetric complete) suggesting that neutral effects are likely the predominant processes governing the structure of these network motifs (Simmons et al. 2020). Thus, as indicated for other network properties like nestedness (Krishna et al. 2008, Suweis et al. 2013), species relative abundances (i.e., neutral effects) are possibly one of the main determinants of indirect interactions in real plant-pollinator networks.

From a pollinator perspective, we found that generalist network motifs were consistently over-represented (e.g. fan-type motifs; **Figure 2**). For instance, when we compare motifs with equal structure for both trophic levels, we found that floral visitors had always over-representation in the positions that implied the highest number of links. Despite this is in good agreement with the current view of the generalist tendency of pollinators (Waser et al. 1996, Olesen and Jordano 2002, Ollerton 2017), we found that this over-representation of the generalist network motifs was largely driven by the functional group of bees (e.g., positions 42, 46, 16, 4, 14 and 35 in motifs 15, 17, 7, 2, 6 and 13, respectively; see **Figure 3** for positions and **Figure 4** for motifs) while the rest of the functional groups tended to be under-represented in these generalist motifs and over-represented in motif positions with specialist roles (e.g., positions 11, 21, 28 and 34 in motifs 5, 9, 11 and 13, respectively; see **Figure 3** for positions and **Figure 4** for motifs). This large effect of bees on the overall patterns is no surprise, as bees were the functional group with largest number of interaction records in the whole set of networks as also recorded for other broad scale plant-pollinator network studies (Carvalheiro et al. 2014, Ollerton 2017). Unlike floral visitors, plant functional groups showed little differences on its appearance in different motif positions, and thus similar probability of appearing on all of them. If anything, the groups that involved higher levels of outcrossing showed a tendency to be more over-represented and probable on different motif positions. For instance, the group of self-incompatible

perennials with large flowers, one of the groups that is expected to have higher levels of outcrossing rate (Goodwillie et al. 2010), showed the highest over-representation of all plant functional groups (**Figure 3**); and the group of tall plants with unisexual flowers, the one that is expected to have the highest dependence of pollinators on reproduction (Rodger et al. 2021), was the most probable on the generalist motif for plant positions (i.e., plant nodes with 3 to 4 links; **Figure 4**).

The probabilities of appearance of each functional group on the different motif positions were insufficient to predict most of the observed network position combinations (**Figure 5**). This implies that not just the representation of network motifs types but also the identity of the elements of their structure is influenced by ecological processes that determine under- and over-representation of the different position combinations. In addition, we found that just 30% of the possible 53,250 motif combinations are realized. Thus, similarly to the presence of forbidden links at the network level (Olesen et al. 2011, Jordano 2016), we also found the existence of forbidden motif positions and those represented almost two thirds of the total possible combinations of indirect interactions. This concurs well with the general idea that plant-pollinator networks have a high proportion of non-realized direct interactions (Jordano et al. 2003, Bascompte and Jordano 2013), and as a result, also a high proportion of non-realized combinations of indirect interactions. Some of this non-observed interactions are probably due to under-sampling (Dorado et al. 2011, Jordano 2016). However, a large fraction of the possible interactions do not take place because of life-history constraints such as species co-occurrence, trait matching processes and phenological uncoupling (Olesen et al. 2011, Bartomeus et al. 2016, Jordano 2016). Despite the advances in the understanding on the ecological processes that influence species pairwise interactions (Bartomeus et al. 2016, Peralta et al. 2020), we still know little of how these processes determine competitive and facilitative interactions in mutualistic networks. Thus, as suggested by Simmons et al. (2020), exploring the proportion of these realized network motifs that are competitive or facilitative is a necessary next step to improve our understanding of the architecture of biodiversity (Bascompte and Jordano 2007).

Despite the obvious usefulness of functional groups in summarizing species ecological roles (e.g., similar reproductive biology and behaviour of pollinators), delimiting species into functional groups is not often trivial. For instance, taxonomic diversity and trait availability limit our ability to properly represent the wide range of reproductive strategies (Dellinger 2020) and different clustering analysis can lead to distinct outcomes (e.g., k-means and hierarchical clustering). Further, here we divided floral visitors on the taxonomic rank level because it represented adequately the pollinator species ecological roles without the need of the tremendous effort of compiling trait data for over 1000 species of floral visitors. However, further groupings with traits could help revealing greater relevance of trait-matching processes on indirect interactions. We believe that this could be investigated by exploring which proportion of over- and under-represented motifs

present high and low trait-matching, respectively. Moreover, we analysed network motifs focusing exclusively on presence-absence interactions which assume that all interactions are equally important but as suggested by Simmons et al. (2020), further work should account for quantitative measurements of the strength of plant-pollinator interactions. Finally, despite the wide geographical coverage of our plant-pollinator dataset, there are major gaps for some regions of the world with a clear bias towards Europe, America and islands which have long been interested ecologists (MacArthur and Wilson 1963, Warren et al. 2015). These biases are also found in global datasets of networks where the data is highly fragmented across space (Poisot et al. 2021) and more efforts are needed to capture global patterns of plant-pollinator interactions.

In conclusion, by following the framework of Simmons et al. (2020) we have been able to explore competitive and facilitative interactions in plant-pollinator networks worldwide and found that indirect interactions are non-random and are more strongly related to neutral processes than any other ecological process. Further, we found that motif combinations of functional groups are strongly constrained and that the distinct motif positions are associated with different functional groups with different ecological roles (e.g., generalization versus specialization). Although in this study we do not predict indirect interactions, our work is a first step to elucidate novel interactions beyond pairwise interactions.

Acknowledgements - This study was supported by the European project SAFEGUARD (101003476 H2020-SFS-2019-2). We thank all researchers that made their data openly available and sent it upon request. We also thank Bryony Wilcox, Greg Bible, Mercedes Sanchez-Lanuza and David Ragel for their help with data collection. Finally, JBL also thanks the University of New England for the funding provided to carry out this work.

Author contributions

IB designed the idea. JBL collected the data. AA-P lead the analysis with contributions of JBL and guidance of IB. JBL wrote the manuscript with contributions of IB and AA-P.

Data availability statement

All data and code would be deposited at Dryad Digital repository upon acceptance.

References

- Allen-Perkins, A. et al. 2021. Community level individual-based plant-pollinator networks unveil pollen flow dynamics and plant reproductive success. - bioRxiv in press.
- Baker, N. J. et al. 2015. Species' roles in food webs show fidelity across a highly variable oak forest. - *Ecography* 38: 130–139.
- Bartomeus, I. et al. 2016. A common framework for identifying linkage rules across different types of interactions. - *Functional Ecology* 30: 1894–1903.
- Bascompte, J. and Jordano, P. 2007. Plant-animal mutualistic networks: The architecture of biodiversity. - *Annu. Rev. Ecol. Evol. Syst.* 38: 567–593.
- Bascompte, J. and Jordano, P. 2013. *Mutualistic networks*. - Princeton University Press.
- Bascompte, J. et al. 2003. The nested assembly of plant–animal mutualistic networks. - *Proceedings of the National Academy of Sciences* 100: 9383–9387.
- Bastolla, U. et al. 2009. The architecture of mutualistic networks minimizes competition and increases biodiversity. - *Nature* 458: 1018–1020.
- Bates, D. et al. 2015. Fitting linear mixed-effects models using lme4. - *Journal of Statistical Software* 67: 1–48.
- Bergamo, P. et al. 2021. Pollinator-mediated indirect effects on plant fecundity revealed by network indices. in press.
- Blüthgen, N. et al. 2006. Measuring specialization in species interaction networks. - *BMC Ecology* 6: 9.
- Carvalho, L. G. et al. 2014. The potential for indirect effects between co-flowering plants via shared pollinators depends on resource abundance, accessibility and relatedness. - *Ecology Letters* 17: 1389–1399.
- Chamberlain, S. et al. 2020. Taxize: Taxonomic information from around the web.
- Cirtwill, A. R. et al. 2018. Between-year changes in community composition shape species' roles in an arctic plant–pollinator network. - *Oikos* 127: 1163–1176.
- Coux, C. et al. 2016. Linking species functional roles to their network roles. - *Ecology Letters* 19: 762–770.
- Dellinger, A. S. 2020. Pollination syndromes in the 21st century: Where do we stand and where may we go? - *New Phytologist* 228: 1193–1213.
- Devaux, C. et al. 2014. Constraints imposed by pollinator behaviour on the ecology and evolution of plant mating systems. - *Journal of Evolutionary Biology* 27: 1413–1430.
- Dolnicar, S. et al. 2014. Required sample sizes for data-driven market segmentation analyses in tourism. - *Journal of Travel Research* 53: 296–306.
- Dorado, J. et al. 2011. Rareness and specialization in plant–pollinator networks. - *Ecology* 92: 19–25.
- Dormann, C. F. et al. 2009. Indices, graphs and null models: Analyzing bipartite ecological networks. - *The Open Ecology Journal* 2: 7–24.

- Fenster, C. B. et al. 2004. Pollination Syndromes and Floral Specialization. - *Annual Review of Ecology, Evolution, and Systematics* 35: 375–403.
- Földesi, R. et al. 2021. Larger pollinators deposit more pollen on stigmas across multiple plant species—a meta-analysis. - *Journal of Applied Ecology* 58: 699–707.
- Gong, Y.-B. and Huang, S.-Q. 2009. Floral symmetry: Pollinator-mediated stabilizing selection on flower size in bilateral species. - *Proceedings of the Royal Society B: Biological Sciences* 276: 4013–4020.
- Goodwillie, C. et al. 2010. Correlated evolution of mating system and floral display traits in flowering plants and its implications for the distribution of mating system variation. - *The New Phytologist* 185: 311–321.
- Gower, J. C. 1971. A general coefficient of similarity and some of its properties. - *Biometrics*: 857–871.
- Greenleaf, S. S. et al. 2007. Bee foraging ranges and their relationship to body size. - *Oecologia* 153: 589–596.
- Gu, Z. et al. 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. - *Bioinformatics* 32: 2847–2849.
- Guimarães, P. R. J. 2020. The structure of ecological networks across levels of organization. - *Annual Review of Ecology, Evolution, and Systematics* 51: 433–460.
- Jordano, P. 1987. Patterns of mutualistic interactions in pollination and seed dispersal: Connectance, dependence asymmetries, and coevolution. - *The American Naturalist* 129: 657–677.
- Jordano, P. 2016. Sampling networks of ecological interactions. - *Functional ecology* 30: 1883–1893.
- Jordano, P. et al. 2003. Invariant properties in coevolutionary networks of plant–animal interactions. - *Ecology letters* 6: 69–81.
- Krishna, A. et al. 2008. A neutral-niche theory of nestedness in mutualistic networks. - *Oikos* 117: 1609–1618.
- Laliberté, E. et al. 2014. Package ‘FD’. - Measuring functional diversity from multiple traits, and other tools for functional ecology in press.
- Lanuza, J. B. et al. 2021. Trade-offs among plant reproductive traits determine interactions with floral visitors. - *bioRxiv* in press.
- Lázaro, A. et al. 2020. Linking species-level network metrics to flower traits and plant fitness. - *Journal of Ecology* 108: 1287–1298.
- MacArthur, R. H. and Wilson, E. O. 1963. An equilibrium theory of insular zoogeography. - *Evolution*: 373–387.
- Milo, R. et al. 2002. Network motifs: Simple building blocks of complex networks. - *Science* 298: 824–827.
- Moeller, D. A. 2004. Facilitative interactions among plants via shared pollinators. - *Ecology* 85: 3289–3301.
- Olesen, J. M. and Jordano, P. 2002. Geographic Patterns in Plant–Pollinator Mutualistic Networks. - *Ecology* 83: 2416–2424.
- Olesen, J. M. et al. 2007. The modularity of pollination networks. - *Proceedings of the National Academy of*

- Sciences 104: 19891–19896.
- Olesen, J. M. et al. 2011. Missing and forbidden links in mutualistic networks. - *Proceedings of the Royal Society B: Biological Sciences* 278: 725–732.
- Ollerton, J. 2017. Pollinator diversity: Distribution, ecological function, and conservation. - *Annual Review of Ecology, Evolution, and Systematics* 48: 353–376.
- Ollerton, J. et al. 2009. A global test of the pollination syndrome hypothesis. - *Annals of botany* 103: 1471–1480.
- Peralta, G. et al. 2020. Trait matching and phenological overlap increase the spatio-temporal stability and functionality of plant–pollinator interactions. - *Ecology letters* 23: 1107–1116.
- Poisot, T. et al. 2015. A continuum of specialists and generalists in empirical communities. - *PloS One* 10: e0114674.
- Poisot, T. et al. 2021. Global knowledge gaps in species interaction networks data. - *Journal of Biogeography* in press.
- R Core Team 2021. R: A language and environment for statistical computing. - R Foundation for Statistical Computing.
- Rodger, J. G. et al. 2021. Widespread vulnerability of flowering plant seed production to pollinator declines. - *Science Advances* 7: eabd3524.
- Sargent, R. D. and Ackerly, D. D. 2008. Plant–pollinator interactions and the assembly of plant communities. - *Trends in Ecology & Evolution* 23: 123–130.
- Simmons, B. I. et al. 2019a. Motifs in bipartite ecological networks: Uncovering indirect interactions. - *Oikos* 128: 154–170.
- Simmons, B. I. et al. 2019b. Bmotif: A package for motif analyses of bipartite networks. - *Methods in Ecology and Evolution* 10: 695–701.
- Simmons, B. I. et al. 2020. Niche and neutral processes leave distinct structural imprints on indirect interactions in mutualistic networks. - *Functional Ecology* in press.
- Stouffer, D. B. et al. 2007. Evidence for the existence of a robust pattern of prey selection in food webs. - *Proceedings of the Royal Society B: Biological Sciences* 274: 1931–1940.
- Stouffer, D. B. et al. 2012. Evolutionary conservation of species’ roles in food webs. - *Science* 335: 1489–1492.
- Strauss, S. Y. 1991. Indirect effects in community ecology: Their definition, study and importance. - *Trends in Ecology & Evolution* 6: 206–210.
- Suweis, S. et al. 2013. Emergence of structural and dynamical properties of ecological mutualistic networks. - *Nature* 500: 449–452.
- Tur, C. et al. 2013. Linking plant specialization to dependence in interactions for seed set in pollination

networks. - PloS one 8: e78294.

Warren, B. H. et al. 2015. Islands as model systems in ecology and evolution: Prospects fifty years after MacArthur-wilson. - Ecology Letters 18: 200–217.

Waser, N. M. et al. 1996. Generalization in Pollination Systems, and Why it Matters.

White, D. and Gramacy, R. 2009. Maptree: Mapping, pruning, and graphing tree models. R package version 1.4-5.