


Volatility of network indices due to undersampling of intraspecific variation in plant–insect interactions

Jonas Kuppler^{1,2} · Tobias Grasegger¹ · Birte Peters¹ · Susanne Popp¹ · Martin Schlager¹ · Robert R. Junker¹ 

Received: 19 September 2016 / Accepted: 27 December 2016 / Published online: 1 March 2017
© The Author(s) 2017. This article is an open access publication

Abstract Appropriate sampling effort of interaction networks is necessary to extract robust indices describing the structure of species interactions. Here we show that time-invariant variation in the composition and diversity of interaction partners of plant individuals of the same species explains volatility in aggregate network statistics due to undersampling. Within a multi-species pollinator–plant interaction matrix, we replaced the interactions observed on multiple individuals of a single plant species (*Sinapis arvensis*, pooled interactions) with the plant–insect interactions observed on a single plant individual. In the resampling approach, we considered the interactions of 1 to 84 *S. arvensis* individuals in different combinations. For each resampled network, several commonly applied aggregated statistics were calculated to test how intraspecific variation affects the properties of a multi-species network. Our results showed that aggregate statistics are sensitive towards qualitative and quantitative intraspecific variation of flower–visitor interactions within a multi-species network, which may affect the ecological interpretation about the properties of a community. These findings challenge

the robustness of commonly applied network indices, confirm the urge for a sufficient and representative sampling of interactions, and emphasize the significance of intraspecific variation in the context of communities and networks.

Keywords Community ecology · Plant–pollinator network · Robustness · Specialization

Introduction

Network theory strongly contributed to the understanding about the structure of biotic interactions and their importance for community dynamics, function, and stability (Heleno et al. 2014). Classically network statistics summarize trophic interactions at the species level and thus consider the mean interaction frequency of a species. However, each species is represented by a number of individuals that interact with individuals of species of the other trophic level. Although recent studies emphasized the pronounced intraspecific variability in interactions in flower–visitor networks (Dupont et al. 2011, 2014; Gómez et al. 2014; Tur et al. 2014; Hoffmeister et al. 2015; Kuppler et al. 2016), we lack understanding of the relevance of variation among individuals within species in the context of flower–visitor communities and interaction networks.

The structure of interaction networks is commonly described by aggregate statistics such as connectance, complementary specialization, or modularity (Olesen et al. 2007; Blüthgen et al. 2008). These properties may be explained by the abundance distributions of plants and animals, phenology, morphological matching, and other factors (Vázquez et al. 2009; Junker et al. 2013). An on-going methodological debate in community ecology is the impact of sampling effort on the structure and aggregate metrics of

Handling Editor: Kristine Nemeč.

Electronic supplementary material The online version of this article (doi:10.1007/s11829-016-9493-1) contains supplementary material, which is available to authorized users.

✉ Robert R. Junker
robert.junker@sbg.ac.at

¹ Department of Ecology and Evolution, University of Salzburg, Hellbrunnerstrasse 34, 5020 Salzburg, Austria

² Department Biology, Institute of Sensory Ecology, Heinrich-Heine-University, Universitätstrasse 1, 40225 Düsseldorf, Germany

networks (Blüthgen et al. 2008; Gibson et al. 2011; Fründ et al. 2016). The aggregate metrics are based on the richness, relative abundance, and the composition of interactions for each plant and animal species. Sampling bias usually decreases with increased sampling effort (Chacoff et al. 2012). Thus, a sufficient sampling is required to gain a precise evaluation of network properties (Gotelli and Colwell 2001). However, insufficient sampling has been suggested to be common in network studies and often the observed interactions are incomplete (Chacoff et al. 2012). So far the bias induced in network statistics due to sampling has been mainly discussed from a methodological perspective (e.g., Blüthgen 2010; Gibson et al. 2011; Vizentin-Bugoni et al. 2015; Fründ et al. 2016). However, the underlying ecological causes for the need of sufficient sampling effort remain unknown.

We evaluated how intraspecific variation in flower-insect interactions affects the volatility of aggregate network statistics and thus urges for increased sampling effort. Therefore, in a resampling approach we replaced the species mean of interactions of a single plant species with the interactions observed on plant individuals of the same plant species and tested how this variation affects the properties of a multi-species network. Our simulation study thus aims at identifying intraspecific variability in interactions as a cause for volatility in aggregate network statistics resulting from undersampling.

Materials and methods

Flower–visitor interactions

Flower–visitor interactions were recorded on a 30×30 m meadow plot in the Botanical Garden of the University of Salzburg, Austria in June 2015. We observed all flowering plant species ($n=13$ species) and additionally 20 potted individuals of *Sinapis arvensis* L. (Brassicaceae) that were homogeneously distributed in the plot. To facilitate a homogenous sampling, the flower–visitor interactions were gathered in four subplots (15×15 m, each) by random walks on four days between 9:00 and 16:00 h (1 or 2 h in the morning and 1 h in the afternoon) resulting in a total observation time of 10 h per subplot (i.e., total observation time = 40 h). All insect interactions on flowers were recorded and visitor taxa were assigned to easily recognizable groups, e.g., honey bees and large and small solitary bees or syrphid flies (Kuppler et al. 2016). Additional to the sampled flower–visitor interactions, we used observations of flower–insect interactions of 64 plant individuals of *S. arvensis* that we recorded in a common garden setting in 2013 on the same meadow during the same time of the year (Kuppler et al. 2016, for a short description of the

method see supplementary material). In the further analysis, we thus used the data resulting in a total of 84 *S. arvensis* individuals.

Network resampling and statistics

To test the relative effect of time-invariant, non-random intraspecific differences between plant individuals (Kuppler et al. 2016) on commonly calculated network indices and thus the ecological interpretation of the properties of multi-species networks, we generated interaction matrices containing the interactions of the 13 non-potted plant species and the interactions of one or several ($n=1, 2, \dots, 84$) randomly sampled *S. arvensis* individuals. We resampled 81,169 networks $N\binom{n}{k}$ with n being the total number of *S.*

arvensis individuals (84) and k being the number of included *S. arvensis* individuals (1, 2, ..., 84). For each k between 2 and 82, we resampled 1000 combinations without replacement. For $k=1$ and $k=83$, we resampled all possible unique combinations (i.e., 84 combination); for $k=84$, we used only one combination representing full sampling. Each network consisted of the interaction frequency (per flower per hour) of the 13 plant species and the mean interaction frequency of $N\binom{84}{k}$ randomly selected *S.*

arvensis individuals (Figure S1). Thus, the network size did not strongly vary between resampled networks as the number of species in all networks was constant and the number links varied only marginally between *Sinapis arvensis* individuals. Further, as quantitative network statistics are sensitive to differences in interaction frequencies, the resampling was repeated with a constant interaction frequency for each *S. arvensis* individual (proportional interaction strength per *S. arvensis* individual times the mean number of interactions of all *S. arvensis* individuals).

For each resampled network $N\binom{84}{k}$, we calculated common aggregate statistics characterizing its structure: one at species level (complementary specialization d') and five at community level (complementary specialization H'_2 , connectance, nestedness (WNODF), interaction evenness, and QuanBiMo modularity (Dormann et al. 2009)). The QuanBiMo modularity algorithm (Dormann and Strauss 2014) assigns species to modules based on their similarity in quantitative interactions. To test whether the intraspecific differences in interactions in *S. arvensis* affect the assignment of species into modules, we used the QuanBiMo algorithm (steps = 1,000,000) for each of the resampled networks $N\binom{84}{k}$ ($k=1-84$). The QuanBiMo algorithm uses an iterative proposing process to detect modules,

and thus, two separate runs of the algorithm may not result in the same module composition (Dormann and Strauss 2014). Therefore, we used $n=1000$ QuanBiMo runs for each of the further analyses to reduce/avoid random effects in module composition caused by the algorithm. For each of these resampled networks ($n=84$), the plant species' assignment to modules in the $n=1000$ QuanBiMo runs was compiled in a presence/absence matrix with plant species as rows and identified modules as columns (number of columns = sum of modules of all 1000 QuanBiMo runs). Based on this matrix, we calculated Sørensen-distances between plant species (small distances between species pairs thus indicate a common co-occurrence in one module). To visualize the plant species' affiliation to modules relative to the other species, we performed a principal coordinate analysis (PCoA). To quantify the deviation of the module composition (represented as PCoA) based on the networks containing the interactions of one or several $N\binom{84}{k}$ and all *S. arvensis* individuals $N\binom{84}{84}$, we compared the deviation of the PCoAs using Procrustes analysis (Peres-Neto and Jackson 2001; Oksanen et al. 2016). Procrustes sum of squares ss were used as measurement for the deviation in the module composition. Further, to investigate how the impact of intraspecific differences in interactions in *S. arvensis* affects commonly used network statistics (complementary specialization d' and H'_2 , connectance, nestedness (WNODF), and interaction evenness), we calculated the range of values (i.e., difference between maximum and minimum value) of network indices for all networks $N\binom{84}{k}$ with $k=S. arvensis$ individuals included. All analyses were performed with the statistical software R 3.2.2. (R Core Team 2016) and the packages *vegan* (Oksanen et al. 2016) and *bipartite* (Dormann et al. 2009).

Results

In total, we observed 5411 interactions with a mean interaction frequency per flower per hour of 1.88 ± 0.65 (SE) per plant species including *S. arvensis*. Aggregate statistics were sensitive to the intraspecific variation in interaction patterns in *S. arvensis*. The range of variation across different networks $N\binom{84}{k}$ was considerable for all aggregate statistics that describe the properties of the whole network (Fig. 1) as well as for the species-specific complementary specialization d' (networks including one *S. arvensis* individual $N\binom{84}{1}$, each: full range of d' for *S. arvensis*: 0.71; for all other plant species: full range between 0 and 0.54;

Fig. 1f, electronic supplementary material, figure S2). The sensitivity decreased with the number of included individuals to calculate the mean interaction frequency of *S. arvensis* in the networks ($k=1-84$). The modularity analyses clearly showed that plant species that were part of the networks including the visitation frequencies of only one *S. arvensis* individual $N\binom{84}{1}$ were assigned in variable combinations to modules (Fig. 2, for $N\binom{84}{k}$ figure S6). The QuanBiMo algorithm applied to detect modules usually produces varying results in different runs (Dormann and Strauss 2014). Therefore, for each network the algorithm was repeated 1000 times to avoid arbitrary results (figure S7). However, the number of modules detected was constant across different networks $N\binom{84}{1}$ (number of modules = 5, except for one network $N\binom{84}{1}$ with only four modules).

For the resampled networks with a constant interaction frequency for *S. arvensis*, we found the same patterns of variation for all network indices except for species-specific complementary specialization d' which show less variation for some plant species (networks including one *S. arvensis* individual $N\binom{84}{1}$, each: range of d' for *S. arvensis*: 0.73; for all other plant species: range between 0 and 0.32; see electronic supplementary material).

Discussion

Our simulation study revealed that intraspecific differences in interactions of one plant species (*Sinapis arvensis*) in a multi-species network are a cause for the volatility in aggregate statistics as a result of undersampling. The aggregate statistics considered in this study showed a considerable range of variation towards changes in individual visitation patterns. With increasing numbers of individuals included in the network, however, the variation decreased. Accordingly, aggregate statistics were quite robust in networks with a high sampling effort (i.e., number of included individuals). Our findings demonstrate that the results of aggregated statistics can be strongly influenced by the intraspecific variation in interaction patterns and differential sampling of individuals.

In a previous study, we showed that the intraspecific variation in interaction patterns of *S. arvensis* were non-random and driven by differences in functional traits, which suggest that plant individuals occupy specific functional position within communities (Kuppler et al. 2016). This notion is confirmed by the resampling approach where the complementary

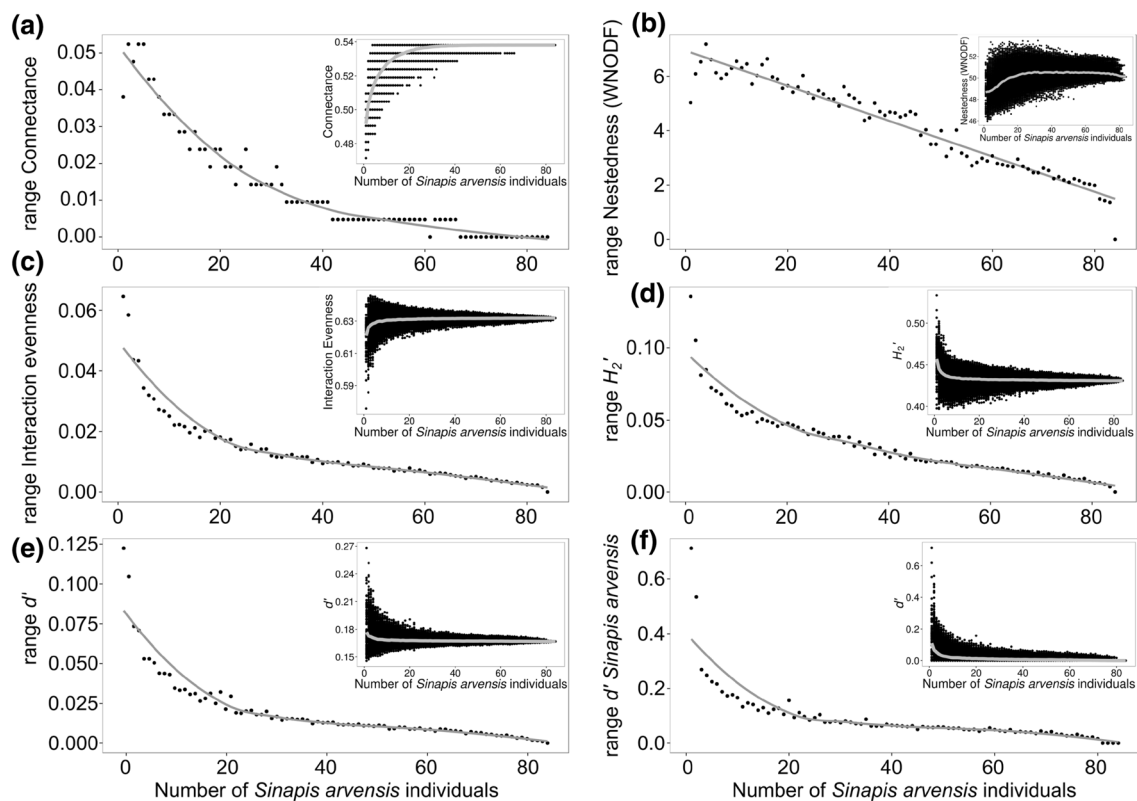


Fig. 1 Variation in commonly applied network statistics as a function of the number of *Sinapis arvensis* individuals included in the interaction matrix. **a–f** Range of values of network indices calculated for all networks $N\binom{84}{k}$ with $k = S. arvensis$ individuals included. Insets vis-

ualize the raw (points) and mean (lines) values. Non-parametric local polynomial regression curve or linear regression lines (for nestedness) depict the trends in changes in the range of the aggregate network statistics

specialization (d') of *S. arvensis* individuals strongly varied. This shows that different individuals of one species may appear to be generalized or specialized in a community context and therefore may occupy different positions in the network (Gómez and Perfectti 2012). Differences in the interactions between individuals within a network can result in differences in their reproductive success (Gómez et al. 2011; Gómez and Perfectti 2012). Hence, the composition of individuals of a species and their interactions influences the reproductive performance of a local population, which ultimately affect the species composition and structure of the community.

Apart from the ecological implication of individual variation, our results show that small quantitative intraspecific differences in interactions may alter the ecological interpretation of aggregate network statistics at community level. For example, modularity is used to address the topological role of species within the network and indicate their importance for the integrity of the network structure (Kaiser-Bunbury and Blüthgen 2015). These roles are defined by the module affiliation and the links to other modules of each species (Olesen et al. 2007). While the number of modules was constant, the module composition of the entire network

and likely the topological roles (Olesen et al. 2007) of most species were altered by different *S. arvensis* individuals. Thus, variation in interactions within species may affect modularity to a similar degree as interspecific variation. This variation in module affiliation and other aggregate statistics is most likely to be even more pronounced when including intraspecific variation of several plant species.

In our simulation, we varied the number of *S. arvensis* individuals, a generalized species in terms of flower visitors (Kunin 1993) that is characterized by a large intraspecific variation in interaction partners (Kuppler et al. 2016). Generally, it is suggested that interactions in flower–visitor networks are rather generalized (Waser et al. 1996; Olesen and Jordano 2002) and a certain degree of intraspecific variation has been shown in several natural systems (Herrera 2005; Gómez et al. 2011; Gómez and Perfectti 2012; Dáttilo et al. 2014, but see, 2015). Therefore, the effect of intraspecific variation on the volatility of aggregate statistics is likely an important factor to consider in many network studies. In contrast, in highly specialized plant species like orchids (Pemberton 2010) intraspecific variation usually is low, which reduces the effect on aggregate statistics.

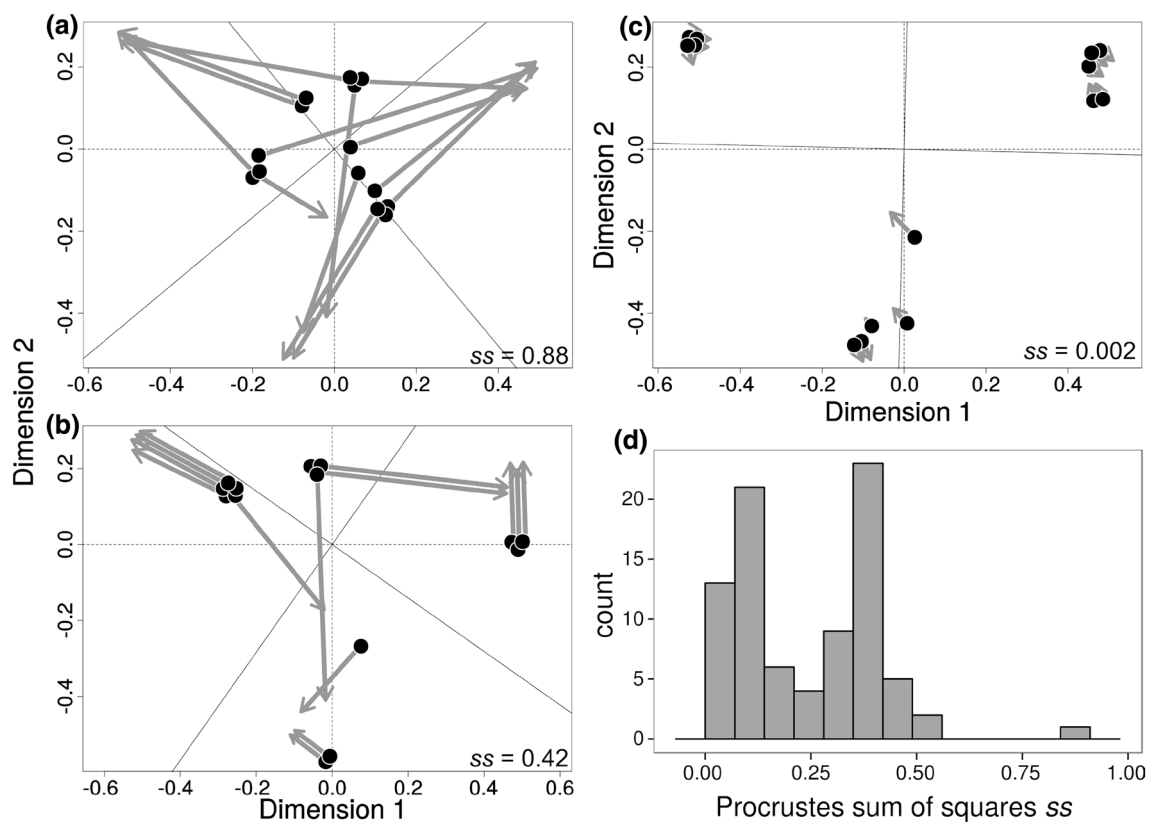


Fig. 2 Influence of different *Sinapis arvensis* individuals on module composition. **a–c** Representative procrustes plots. Each plot shows the comparison (procrustes analysis) of two principal coordinate analyses (PCoA). The PCoAs were based on Sørensen-distances that reflect the module affiliation of each plant species within network, for networks including only one *S. arvensis* individual $N\binom{84}{1}$ (black points) or all individuals $N\binom{84}{84}$ (arrowheads). A long arrow reflects

a large deviation in module affiliation for a plant species within both networks $N\binom{84}{1}$ (black points) and $N\binom{84}{84}$ (arrowheads). Procrustes sum of squares ss are shown as measurement for the deviation in module composition for all plant species ($n = 14$) **d** Histogram of the deviation in module composition of networks including one *S. arvensis* individual ($n = 84$) to species mean (ss -values) $N\binom{84}{84}$

Thus, the impact of intraspecific variation on aggregate statistics may vary between systems.

The sampling effort, which is necessary to properly characterize a species, increases with its intraspecific variability in interactions with animals. Species with a high intraspecific variability in interactions (often generalized) require a larger sampling effort to reliably record the occurring variation, while for species with a low variability (often specialized) less sampling effort may be sufficient. Plant individuals can strongly vary in their interactions despite spatial proximity of the individuals (Kuppler et al. 2016), and therefore, an insufficient number of sampled individuals can result in non-representative aggregated statistics. Further, the spatial arrangement of individuals within a community can additionally increase the variation in interactions. Individuals or patches of individuals with large distances to flowering conspecifics or other plant species may differ in their interactions from less isolated individuals (Steffan-Dewenter and Tschamtkke 1999; Steffan-Dewenter 2003). Besides the

density of conspecifics, the overall species composition and diversity can influence the visitation patterns (Hegland 2014) and especially in heterogeneous habitats a larger number of sampled individuals are needed for robust results of aggregated statistics. Thus, the pollination system, the spatial distribution of a plant species, as well as the habitat structure can influence the sampling effort needed to obtain robust results. Due to the high variability in all three factors, the number of individuals sampled of each plant species in a given habitat should be carefully considered and may even vary between one species in two different habitats.

In summary, our study indicated that aggregate networks statistics are sensitive to small quantitative changes in the interactions of one plant species within a multi-species network. This shows that intraspecific variation can induce volatility in aggregate statistics as a result of undersampling and emphasizes the importance of sufficient and representative sampling of individuals within a community. Further, the conclusions derived from aggregated networks statistics

as indicators for community stability, diversity, or impact of environmental change (Hagen et al. 2012; Olesen et al. 2012) may also be influenced. Thus, intraspecific variation is both pitfall and opportunity for a detailed understanding of community structure using network theory.

Acknowledgements Open access funding provided by Paris Lodron University of Salzburg. We thank the Botanical Garden of the University Salzburg for logistical support and Maren K. Höfers and Lisa Wiesmann for help with fieldwork. Marcos Méndez and anonymous reviewers provided valuable comments on earlier versions of the manuscript. The study was supported by the Graduate School “Evolutionary Networks: Organisms, Reactions, Molecules” (E-Norm) of the Heinrich-Heine-University, Düsseldorf, Germany and the Deutsche Forschungsgemeinschaft (DFG JU 2856/2–2).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Blüthgen N (2010) Why network analysis is often disconnected from community ecology: a critique and an ecologist’s guide. *Basic Appl Ecol* 11:185–195.
- Blüthgen N, Fründ N, Vázquez DP, Menzel F (2008) What do interaction network metrics tell us about specialization and biological traits? *Ecology* 89:3387–3399
- Chacoff NP, Vázquez DP, Lomáscolo SB et al (2012) Evaluating sampling completeness in a desert plant-pollinator network. *J Anim Ecol* 81:190–200
- Dáttilo W, Fagundes R, Gurka CAQ et al (2014) Individual-based ant-plant networks: diurnal-nocturnal structure and species-area relationship. *PLoS ONE* 9:e99838
- Dáttilo W, Aguirre A, Quesada M, Dirzo R (2015) Tropical forest fragmentation affects floral visitors but not the structure of individual-based palm-pollinator networks. *PLoS ONE* 10:e0121275
- Dormann CF, Strauss R (2014) A method for detecting modules in quantitative bipartite networks. *Methods Ecol Evol* 5:90–98
- Dormann CF, Fründ J, Blüthgen N, Gruber B (2009) Indices, graphs and null models: analyzing bipartite ecological networks. *Open Ecol J* 2:7–24
- Dupont YL, Trøjelsgaard K, Olesen JM (2011) Scaling down from species to individuals: a flower-visitation network between individual honeybees and thistle plants. *Oikos* 120:170–177
- Dupont YL, Trøjelsgaard K, Hagen M et al (2014) Spatial structure of an individual-based plant-pollinator network. *Oikos* 123:1301–1310
- Fründ J, Mccann KS, Williams NM (2016) Sampling bias is a challenge for quantifying specialization and network structure: lessons from a quantitative niche model. *Oikos* 125:502–513
- Gibson RH, Knott B, Eberlein T, Memmott J (2011) Sampling method influences the structure of plant-pollinator networks. *Oikos* 120:822–831
- Gómez JM, Perfectti F (2012) Fitness consequences of centrality in mutualistic individual-based networks. *Proc R Soc B* 279:1754–1760.
- Gómez JM, Perfectti F, Jordano P (2011) The functional consequences of mutualistic network architecture. *PLoS ONE* 6:e16143
- Gómez J, Perfectti F, Klingleberg C (2014) The role of pollinator diversity in the evolution of corolla-shape integration in a pollination-generalist plant clade. *Philos Trans R Soc Lond Ser B* 369:20130257
- Gotelli NJ, Colwell RK (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecol Lett* 4:379–391
- Hagen M, Kissling WD, Rasmussen C et al (2012) Biodiversity, species interactions and ecological networks in a fragmented world. *Adv Ecol Res* 46:89–120
- Hegland SJ (2014) Floral neighbourhood effects on pollination success in red clover are scale-dependent. *Funct Ecol* 28:561–568
- Heleno R, Garcia C, Jordano P et al (2014) Ecological networks: delving into the architecture of biodiversity. *Biol Lett* 10:20131000
- Herrera CM (2005) Plant generalization on pollinators: species property or local phenomenon? *Am J Bot* 92:13–20
- Hoffmeister M, Wittköpper N, Junker RR (2015) Herbivore-induced changes in flower scent and morphology affect the structure of flower-visitor networks but not plant reproduction. *Oikos* 125:1241–1249
- Junker RR, Blüthgen N, Brehm T et al (2013) Specialization on traits as basis for the niche-breadth of flower visitors and as structuring mechanism of ecological networks. *Funct Ecol* 27:329–341
- Kaiser-Bunbury CN, Blüthgen N (2015) Integrating network ecology with applied conservation: a synthesis and guide to implementation. *AoB Plants* 7:plv076.
- Kunin WE (1993) Sex and the single mustard: population density and pollinator behavior effects on seed-set. *Ecology* 74:2145–2160
- Kuppler J, Höfers MK, Wiesmann L, Junker RR (2016) Time-invariant differences between plant individuals in interactions with arthropods correlate with intraspecific variation in plant phenology, morphology and floral scent. *New Phytol* 210:1357–1368
- Oksanen J, Blanchet FG, Kindt R, et al (2016) *Vegan: Community Ecology Package* version 2.4–1. <https://cran.r-project.org/web/packages/vegan/index.html>
- Olesen JM, Jordano P (2002) Geographic patterns in plant–pollinator mutualistic networks. *Ecology* 83:2416–2424
- Olesen JM, Bascompte J, Dupont YL, Jordano P (2007) The modularity of pollination networks. *Proc Natl Acad Sci USA* 104:19891–19896
- Olesen JM, Dupont YL, Hagen M, et al (2012) Structure and dynamics of pollination networks: the past, present, and future. In: Patiny S (ed) *Evolution of plant-pollinator relationships*. Cambridge University Press, Cambridge, UK, pp 374–391
- Pemberton RW (2010) Biotic resource needs of specialist Orchid pollinators. *Bot Rev* 76:275–292
- Peres-Neto PR, Jackson DA (2001) How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. *Oecologia* 129:169–178
- R Core Team (2016) *R: A language and environment for statistical computing*, version 3.3.1. <https://www.r-project.org/>
- Steffan-Dewenter I (2003) Importance of habitat area and landscape context for species richness of bees and wasps in fragmented orchard meadows. *Conserv Biol* 17:1036–1044
- Steffan-Dewenter I, Tscharrntke T (1999) Effects of habitat isolation on pollinator communities and seed set. *Oecologia* 121:432–440
- Tur C, Vigalondo B, Trøjelsgaard K et al (2014) Downscaling pollen-transport networks to the level of individuals. *J Anim Ecol* 83:306–317
- Vázquez DP, Chacoff NP, Cagnolo L (2009) Evaluating multiple determinants of the structure of plant-animal mutualistic networks. *Ecology* 90:2039–2046
- Vizentin-Bugoni J, Maruyama PK, Debastiani VJ et al (2015) Influences of sampling effort on detected patterns and structuring processes of a Neotropical plant-hummingbird network. *J Anim Ecol* 85:262–272
- Waser NM, Chittka L, Price MV et al (1996) Generalization in pollination systems, and why it matters. *Ecology* 77:1043–1060