Dear Ainhoa Magrach,

Your preprint, entitled **Interaction network structure maximizes community-level plant reproduction success via niche complementarity**, has now been reviewed. The referees' comments and the recommender's decision are shown below. As you can see, the recommender found your article very interesting, but suggests certain revisions.

We shall, in principle, be happy to recommend your article as soon as it has been revised in response to the points raised by the referees.

When your revised article is ready, please:

1) Upload the new version of your manuscript onto your favorite preprint server;

2) Follow this link https://ecology.peercommunityin.org/user/my_articles or logging onto the *PCI Ecology* website and go to 'your contributions' section and then to 'Your submitted preprints' subsection in the top menu;

3) Make your changes to the title, summary, link to the article (or its DOI) and keywords if necessary by clicking on the 'Edit Article' button;

4) Write, copy/paste or upload (as a PDF file) your reply to the recommender's and reviewers' comments by clicking on the 'Write, edit or upload your reply to recommender' button. You can also upload (as a PDF file) a revised version of your preprint, with the modifications indicated in TrackChanges mode;

5) When you are ready to submit your new version, click on the 'Save & submit your reply' button.

Once the recommender has read the revised version, he/she may decide to recommend it directly, in which case the editorial correspondence (reviews, recommender's decisions, authors' replies) and a recommendation text will be published by *PCI Ecology* under the license CC-BY-ND.

Alternatively, other rounds of reviews may be needed before the recommender reaches a favorable conclusion. He/she also might decide not to recommend your article. In this latter case, the reviews and decision will be sent to you, but they will not be published or publicly released by *PCI Ecology*. They will be safely stored in our database, to which only the Managing Board has access. You will be notified by e-mail at each stage in the procedure.

Thanks in advance for submitting your revised version.

Yours sincerely,

The Managing Board of *PCI Ecology*

Interaction network structure maximizes community-level plant reproduction success via niche complementarity

Ainhoa Magrach, Francisco P. Molina, Ignasi Bartomeus

https://www.biorxiv.org/content/10.1101/629931v1
Submitted by Ainhoa Magrach 2019-05-07 17:03

Abstract

Declines in pollinator diversity and abundance have been reported across different regions, with implications for the reproductive success of plant species. However, research has focused primarily on pairwise plant-pollinator interactions, largely overlooking community-level dynamics. Yet species do not interact in isolation, they are embedded within larger networks whose structure can affect pollinator functional roles and, ultimately, the pollination services they deliver to plants. Here, we present one of the first efforts linking pollinator visitation to plant reproduction from a community-wide perspective using a well-replicated dataset encompassing 16 well-resolved plant-pollinator networks and data on reproductive success for 19 plant species from Mediterranean shrub ecosystems. We find that, for prediction purposes, information on simple visitation metrics is sufficient. Contrastingly, a mechanistic understanding of the pathways through which differences in pollinator diversity translate into changes in reproductive success, requires additional information on community structure, particularly that reflecting niche complementarity between pollinators.

Keywords: nestedness, niche complementary, fruit set, pollination, plant-pollinator interactions

Round #1

Author's Reply:

Decision

- *by Cédric Gaucherel, 2019-06-17 19:18*
  Manuscript: https://www.biorxiv.org/content/10.1101/629931v1

We ask for a major revision

Recommendation: Major revision
This paper is well written and addresses interesting ecological questions. Yet, as stated by all reviewers, it deserves additional analyses and requires the writing to be more rigorous.

The R3 has made an impressive work by listing all locations at which the authors should add justifications and explanations for readers. Many other precisions are required to follow the study and make it fully reproducible. The R4 added the point of view of an experimenter and suggests to include more details on the measurements (be it in the supplementary materials). R2 has mentioned the possible biases coming from the non-exhaustive sampling of the species network. This point may be generalized by the question of the possible biases of the study. How to reduce the uncertainty coming from all the possible co-variables? Elevation has been mentioned by reviewers, as well as other species (outside the most common plant sp.) or possible evolutionary effects. In addition, R4 reminds that wind-pollination and self-pollination may bias the study, and should be addressed or at least argued.
While R2 mentioned the danger to use the term “prediction”, which was not discussed in the paper, R1 mentioned the danger to use the term “mechanistic” when based on correlative studies. Authors have no choice but to remove these terms or to complete their analyses and provide all the details required to convince the readers. As an ecologist, I am also surprised that authors did not comment the possible autocorrelation between sites. Even with a 7km averaged inter-distance in a forested landscape, I guess we can easily hypothesize a partial redundancy between sites due to spatial links. Unfortunately, autocorrelation is one of the most difficult issues in statistics and in particular in GLM models (see Dorman et al. papers). Authors should mention this point, as some other limitations: in particular, GLM are assuming linear relationships whereas ecological relationships are often non-linear. Authors should probably control this issue too. Finally, one point I am particularly aware is that ecological networks are not static at all, although they are often hypothesized so. Even on a short term as in this study, networks may change their structure (not only their fluxes, but also species and species interactions involved) due to frequent local extinctions and invasions. Furthermore, what if a specific relationships is not stable (shifting from positive to negative for a while)? This observation is neglected by most ecologists today and should be at least commented in the discussion.

Overall, this paper seems to be a relevant attempt to include community analysis into more traditional species-centered studies. For this reason, it should be considered. But reviewers mentioned a large number of points that authors should address before recommendation. For this reason, on behalf of PCI ecology, I suggest a Major revision to be sent.

Best wishes. CG.

Dear Cédric,

We want to thank you and the reviewers for the detailed comments and the encouragement received. Our understanding is that linking interaction networks to community functioning is in its early developments, and hence, several unknowns are still present. Those include methodological questions such as which indices are better at describing the mechanisms, as well as conceptual issues on how to incorporate other acting mechanisms such as stochasticity or evolution. Unfortunately, with our data we cannot possibly explore all these possible options. However, we believe that our efforts constitute a solid step forward in the field, by confronting empirical data with the expectations emerging from a clear question: "By adding specific information on interaction network structure that describes indirect effects amongst plant species, are we better able to understand plant functioning in a plant-pollinator network?" Nevertheless, we were highly stimulated by the complexity of the problem as highlighted by the reviewers and yourself and we have added a full paragraph in the discussion (line ~443) to reflect future lines of research and knowledge gaps.

We also addressed most technical comments or suggestions by:

- Adding extra justifications and explanations, including a description of plant mating systems and possible confounding factors.
- Adding R^2 measures, as a proxy of the predictive power, and reinforcing how the metrics used relate to specific ecological mechanisms.
- Testing for autocorrelation in the data and showing that spatial autocorrelation values are low in our system.
- Releasing the code and data to make the analysis fully reproducible.

Reviews

Reviewed by anonymous reviewer, 2019-05-10 22:31

The manuscript by Magrach et al. addresses how community-level plant reproduction is linked with pollinator visitation. The authors found evidence that visitation alone can well explain some pattern, while niche complementarity offers a mechanistic understanding of what ecological process is responsible for determining the differences in reproductive outputs. The paper is clearly written and I find the research questions to be well-defined and of broad interests. As I am a theoretical ecologist, my review would only focus on the theoretical part of the paper.

We thank the Reviewer for the positive comments and for taking time to review our paper.

One major concern I have is the robustness of the results. As stated in the abstract, this paper has focused on community dynamics, which is a broad and complex issue. The authors have only considered niche complementarity and nestedness, and it is unclear to me why the authors picked only these two possibilities. Furthermore, nestedness, as the authors have admitted, may not be a meaningful metric for this study because of the small network size. Therefore, it is not entirely convincing that the authors have proved their main claim. I suggest the authors test other metrics as well, such as the ones proposed in 10.1101/604868. Of course, the authors are free to choose any other metrics, but the bottom line is that more

metrics must be tested to convince the readers. Again, I am not doubting that correctness of the findings in the manuscript, but rather encourage the authors to run more tests.

The reason why we selected only two network metrics is that we wanted to maintain our analyses simple and with clear prior hypothesis. In observational works that have limited sample sizes, too often too many metrics are tested with the hope of finding one that is relevant. Our understanding is that trying several metrics on the same limited dataset has the same problem as multiple testing. This exercise may be ok for an exploratory analysis, but we believe we have already clear expectations on which metrics may have a stronger influence. Niche overlap gives us an idea of the complementarity of functions performed by different individuals, which we hypothesize should be important for reproductive success as, a greater complementarity might mean that less pollen from other species is brought to a focal plant. However, most plant-pollinator interaction networks are thought to present nested structures, which involves a certain degree of redundancy between the functions performed by different species. Thus, although nestedness is thought to promote diversity and stabilize networks, it might not be the best descriptor for plant reproductive success. We could test for other metrics, but we believe that keeping it simple is a strength of our paper. We have now specified this in the main text and added a discussion about the limitations of testing only two metrics.

P6-L107-114: “Specifically, our study focuses on how the interplay between niche complementarity and redundancy determines reproductive success. Plant reproductive success requires of the delivery of conspecific pollen and thus of a certain degree of niche complementarity (Blüthgen & Klein 2011). Yet, greater values of nestedness, which imply redundancy in species functions, are thought to promote species diversity (Bastolla et al. 2009) and stability (Thébault & Fontaine 2010) within plant-pollinator networks. At present, we do not know how either of these network characteristics affects the functions performed by pollinators.”

P21-L442-455: “Our study illustrates the complexity of linking network structure to ecosystem function empirically, because measuring both structure and function is challenging. For example, there is an ongoing debate as to which network metrics better reflect classic ecological mechanisms, such as niche partitioning or competition (Delmas et al 2018). Here, we focus on testing two specific hypotheses, but other structural properties can be explored when more data becomes available. Furthermore, the structure of plant-pollinators networks is dynamic due to ecological and evolutionary reasons, but so far, we are only able to characterize it for single snap-shots. Moreover, different aspects of functioning may be important, such as the presence of non-linear relationships or the need to consider the functioning of both trophic levels (Godoy et al 2018). In terms of plant reproductive success and the functions performed by pollinators we can measure different aspects, ranging from pollen deposition (the direct pollinator function), to its final effects on plant fitness. Here, we focus on an intermediate stage including fruit quantity and quality, which is of clear ecological importance.”

Other major concern I have is the repetitive use of "prediction". One of the central claims of the paper is that `information on simple visitation metrics is sufficient for prediction

purposes'. Yet, it falls short in two important aspects. First, regression analysis of in-sample inference does show its predictive power (which has to be based on the out-of-sample forecast). I cannot find any forecast analysis in both the main text and the supplementary material. Second, no R^2 or any other similar statistic is reported. Table 1 and 2 only show if a variable is statistically significant, but it does not show if the model explains the data well. Please correct me if I miss something,

Following the Reviewer´s very interesting comment we are now providing R^2 values for all of our models thus showing a proxy of their predictive power apart from their fit. We also tone down the text to highlight the trade-off between model R^2 and model complexity. Unfortunately, our limiting number of sites does not allow us for a proper validation of the predictive power on independent samples. See throughout Results section.

In addition, we have removed the word prediction from the abstract.

P2-L29-35: "We find that models including information on simple visitation metrics alone are good in explaining the variability in reproductive success observed. However, insights into the mechanisms through which differences in pollinator diversity translate into changes in reproductive success require additional information on network structure, particularly that reflecting niche complementarity between pollinators. Specifically, we find a positive effect of increasing niche complementarity between pollinators on plant reproductive success."

Minor points:

- Despite the clarity in the flow of the paper, many points need more explanations. For example, What is the definition of the asymptotic number of species (line 135)? What is the Morisita index (line 159)? Please excuse me if these questions are well-known concepts in the empirical field, but I believe adding more explanations would never hurt the paper.

  We have now provided extra information for these two concepts in particular and provide references of the original concepts.

  P8-L164-168: "In order to evaluate the completeness of our sampling of the pollinator and plant community as well as that of their interactions, we estimated the asymptotic number of species present (Chao et al. 2009), a non-parametric estimator of species richness for abundance data. This estimator includes non-detected species and allowed us to calculate the proportion detected with our original sampling data."

  P10-L193-195: "Niche overlap was calculated as the average overlap in pollinator species visiting a focal plant and each of the other plants in the community using the Morisita overlap index, a measure of similarity between two sets of data (Zhang 2016)."

- Title. Maximization may not be entirely appropriate here. I think the authors found evidence that niche complementarity increases plant reproduction success, but not maximizes.

Following the suggestion by several reviewers we have now changed the title to "Niche complementarity between pollinators **increases** community-level plant reproductive success"

- Line 50. 10.1111/ele.13091 and 10.1002/ecy.2708 are relevant recent works on the importance of indirect interactions via shared resources.

  We have now included one of the references suggested by the Reviewer in this point.

- Line 98. I suggest to remove 'us'.

  Removed.

- Line 127 and many other similar places. Please specify what is the error around the mean. I assume it is standard deviation, but it does not hurt to be precise.

  We have now specified that the error term around the mean is the standard deviation throughout the text.

- Line 176-184. Recently 10.1111/1365-2656.12963 have come up with a new algorithm to compute the normalized NODF. Although the difference between the two algorithms appears to be small (10.1111/1365-2656.12964), I suggest the authors to re-compute the normalized NODF for the sake of accuracy (the code can be found in github.com/CHoeppke/maxnodf).

  Thanks for the reference. We have now recalculated our NODF values using this new methodology. The recalculation provides values that are indeed very similar to the ones we had originally calculated and thus we reach the same conclusions. However, while re-calculating NODF had a problem with three sites in particular where the method suggested by Simmons et al was not working because the condition that the number of links>number of species was not met. We have now specified this in the main text.

  P11-L220-223: "To calculate max(NODF) we used a recently corrected version of the algorithm (Simmons et al 2019) in all but three sites, where the condition that the number of links>number of species was not met and thus precluded us from using this new version."

  Line 199. Is it "generalized" or "general"?

  It is generalized and we have changed this accordingly.

- Line 249. Perhaps 'small' is better than 'low'?

  Changed.

- Line 349. I am not sure if James et al. 2012 have claimed that nestedness destabilizes the dynamics. Instead, I think they have claimed nestedness has an almost null effect on the dynamics.

  Changed.
  P17-L394-397: “In our study, we did not find an effect of nestedness for reproductive success in any case. This metric, widely used across network analysis, and which is deemed to stabilize natural communities (Bastolla et al. 2009 but see James et al. 2012), does not seem to play a direct role in ecosystem function measured as plant reproductive success.”

Reviewed by anonymous reviewer, 2019-06-10 21:21

I reviewed the manuscript of Magrach et al. submitted to PCI Ecol. The authors analyze a number of network metrics and their relationship with plant reproductive success in an attempt to decipher the putative link between network theory and community-level processes. In my opinion, the authors, after extensive fieldwork provide strong evidence that some network metrics, such as centralization and niche overlap, have a significant effect on variation in fruit set and seed number per fruit. However, other metrics often calculated in network-based studies (e.g., nestedness) do not provide significant association with plant reproduction, hence lacking easy interpretation in ecosystem function. This is an interesting result.

My major concern is the complete lack of reference to classic niche and competition theory. Indeed, one can easily interpret the results by omitting network jargon and focusing on the theory developed in the 60s-70s. As I suppose network theory provides a new way to reveal and understand community patterns, I would expect that new patterns should emerge. For example, the observation that "niche complementarity is key in determining differences in reproductive outputs. Indeed, we find that communities where there is less overlap in the niches occupied by pollinator species had greater values of reproductive success, both greater fruit set values and larger numbers of seeds per fruit." (Lines338-341) is entirely accountable by niche and competition theory. Likewise, "At the community level, however, we find that niche complementarity between pollinators, a measure of the overlap in the niches of different species in terms of plant coverage, has an important effect for average fruit set." (Lines 324-327). Again, by definition the extent of resource partitioning among consumers is expected to have a substantial impact on resource dynamics and fitness.

In consequence, I would suggest that in order to improve the quality of their manuscript, the authors should try to reinforce the idea that some findings are in line with seminal advances in competition and niche theory by MacArthur, Pianka, Levin, (e.g., niche complementarity, niche overlap, resource partitioning, competition release, etc).

Thanks for pointing this out. We completely agree that network theory is mainly a tool to scale-up classical concepts of community ecology to full multi-trophic communities in an elegant way. We have now clarified this in text and added some relevant refences in different places. See, for example:

P4-L71-74: “This line of research, rooted in niche theory and revamped by food-web studies (MacArthur & Levins 1967; May & MacArthur 1972, Tilman 1982, Godoy et al 2018), has greatly advanced theory, but these ideas have not yet been tested using empirical data (but see Poisot et al. 2013).”

Reviewed by anonymous reviewer, 2019-05-22 14:53

The paper clearly has several merits. Approaching the functioning of plant-pollination network using species and community based indices in parallel is clever. The specific objective addressed (exploring the explanatory power of community structure as opposed to more classical diversity indices) is also stimulating.

I however see several limits and needs for additional justifications.

A major limit is, according to me, the mismatch between the main objective, the results, and the conclusions. It is claimed that using community structure will help to shed light on overlooked mechanisms. Yet, those mechanisms are still very unclear after several readings. The study remains correlative and how species interact and what and how potential mechanisms contribute to increase the reproductive success is still very speculative and not really inferred from the results.

The general conclusion, therefore, is I think very expected (eg end of the abstract or L328-335). Indeed, the fact that other metrics (than simple visitation) are needed to understand community level processes is not surprising. But this does not provide in itself “a mechanistic understanding of the pathways through which pollinator diversity translate into changes in reporoductive success”. Dispersion, competition, trades off between specialization and pollination success, habitat filtering, co-evolution are few among the many processes that could be studied explicitly for such a mechanistic understanding. The metrics in itself are only tools to capture potential processes. In this respect, that additional metrics are needed to capture multiple processes is not making a very strong case.

We agree that we are not testing mechanisms in an experimental way. However, we are testing clear hypothesis that link ecological mechanisms (e.g. niche partitioning) with proxies for these mechanisms (i.e. measured niche overlap). We have rephrased several sentences to show that rather than testing the mechanism, we provide correlative evidence that the observed patterns are consistent with the mechanisms proposed. For example:

P6-L118-124: “Our results suggest that models including information on simple visitation metrics alone are good in explaining the variability observed in reproductive success. However, insights into the mechanisms through which differences in pollinator diversity translate into changes in reproductive success require additional information on network structure, notably information on the complementarity between the functions performed and the niches occupied by different pollinator species. Specifically, we find a positive effect of increasing niche complementarity between pollinators on plant reproductive success.”

A potentially key missing information for this framework are the environmental variables. Unless I have missed something 16 sites are described and equated to 16 networks. Therefore the pure effect of site is confounded with the network property. If an environmental co-variate both influence the plant and insect diversity, then part (and potentially most) of the results derive from the spatial distribution of this co-variate, not from any network functioning. It is likely, for instance, that altitude (Figure 1) is both affecting plant and pollinator diversity similarly, driving the correlation between the two components of the network. I was surprised that this is completely neglected.

This is a fair point, but the study was designed to compare habitats with similar species composition and hence minimizing confounding variables. For example, altitude does not vary widely across sites and ranges from 50m to 150m above sea level, and the area covered by all sites is small enough so that all sites have the same soil characteristics and experience the same climatic variables (e.g. rainfall). Landscape composition is the only variable which varies across sites, but the effects of local habitat are much more prevalent on community structure than indirect effects derived from landscape composition. Hence, we do not identify any key confounding variables that should be included. We highlighted this in methods.

In addition, we ran mantel correlograms to identify the presence of spatial autocorrelation in the variables included in our analyses and found that this would not be an issue.

P7-L129-132: “All sites were located within similar elevations (ranging from 50 to 150 m a.s.l.), similar habitat and soil types, and presented similar plant compositions (plant mean Sørensen beta-diversity among sites = 0.41), reducing potential confounding factors.”

P9-L180-182: “In addition, we checked for potential spatial autocorrelation in our data by means of Mantel correlograms. Autocorrelation values were low for all variables included in our analyses (Figure S1) and hence we treat each site as independent in our analysis.”

The paper remains largely inconclusive for many relationships. More precisely, for important descriptors of reproductive success (fruit and seed weight, L282) and for community level, model 1 and 2 ie (ie including simple visitation metrics or information on community structure) are equally good (L291) or model 1 is even better for weight variables. This, together with other intriguing results (eg negative effect of pollinator diversity on fruit set) cast some doubts on the robustness of the conclusion that major gains of information and a better understanding of the mechanisms involved are provided by considering metrics of community structure. It probably does. But the results are not that clear on that perspective. My understanding is that even more confusion is produced.

Overall, it seems that this paper is a nice methodological contribution with an important message: approaching network functioning necessitates integrative descriptors. But the extent to which it helped to access to a better understanding of ecosystem functioning as claimed, is at this stage overstretched and deserve more justifications. My feeling is that the authors need to better delineate their strategy towards a more rigorous test of alternative candidate processes or develop the methodological side of the paper further. But the mixture is I think not really convincing or would deserve more nuanced conclusions and more justifications.

We would also like to present more conclusive and straightforward results, but ecology is complex, and several factors act simultaneously. Our robust dataset may open up some questions we cannot answer yet, but we believe this is positive. However, following the Reviewer´s suggestions we have now reinforced the main message that the addition of interaction network pattern increases the explanatory power of models, but that this improvement is not always enough to compensate for the increased model complexity. Hence, in this trade-off the researcher should value what to do. For example:

P18-L373-380: "One of the first conclusions we can extract from the fact that in most cases both of the models we considered (i.e., the simple model based on visitation metrics and the more complex one including network structure metrics) were equally good, is that the added complexity of measuring the full network of interactions may not pay off for rapid assessments. Hence, simple visitation metrics, such as pollinator richness, might be enough to describe general patterns (Garibaldi et al. 2013, 2015). Yet, adding network level information may inform us of the potential ecological mechanisms underlying the processes driving the observed patterns."

Reviewed by Michael Lattorff, 2019-06-13 19:47
Download the review (PDF file)

The manuscript "Interaction network structure maximizes community-level plant reproduction success via niche complementarity" by Magrach et al. published as a pre-print on bioRxiv deals with pollinator community level effects on plat reproduction. The experimental effort and the data analysis are state of the art and beyond, however, some of the assumptions and interpretation of the results are not well thought about and are overemphasizing what could be interpreted from it.
The assumptions that are made are very clear, all plant species in the study rely to 100% on insect-mediated pollination. This is rarely the case. Most plant species show a mixture among the three most common plant reproduction mediating effects (self-, wind-, insect mediated pollination). Even if a high number of plant species (approx. 75% are dependent on insect-mediated pollination) still are number of these species might be able to use self- (often a lower extend) and wind-pollination alongside with insect-mediated pollination. For me it is not understandable why this has not been considered, as taking into account the effects of wind-pollination might strengthen some of the results.

We provide mating system for all plant species studied. All studied species are at least partially self-incompatible and 68% rely completely on pollinators for reproduction (self-incompatible), and hence all plants studied depend on pollinators to maximize their reproduction, see Table S4.

Most of the results can be explained by Model1 which incorporates individual-based metric like number of flower visitations. Thus, the title claiming that community structure is heavily important is just overemphasizing the results.

We believe that we are not overemphasizing the results for two reasons. 1) Most of the models show nearly identical values of AIC for both models which tells us these models are

equally good and thus one cannot be selected over the other, suggesting that both simple and network metrics are needed to understand the patterns observed and its choice depends on level of model complexity you are willing to get and 2) in closer inspection of the predictive ability of the models we actually see that model 2 shows a substantial better predictive power than model 1 for both fruit set and the number of seeds per fruit at the species and site level. We have now included these R^2 values in the main text, which we hope will help substantiate our claim.

That pollinator species diversity is known to enhance delivery of pollination services. For me it is not understandable why this is again studied here without adding more information, e.g. why diversity is enhancing delivery of pollination services, which could be due to complementarity or redundancy in function of pollinator species. This would really add value when considering such ideas instead of studying species diversity over and over again.

We believe that this is exactly the opposite of what we are saying in our manuscript. In this case, we are saying that just focusing on species diversity is not enough and that you need to understand how that diversity is structured in order to understand its relationship with function. In this case, we show that it is functional complementarity that is driving the effects of diversity on function.

Based on these two issues, I think the manuscript, although robust in all parts of the methodology and data analysis, should be critically revised incl. a change of the title.

Based on the suggestions of several of the reviewers the title has now been changed and reads: “Niche complementarity between pollinators increases community-level plant reproductive success”.

Minor issues:

L117-121 capturing rare observations outside of the observation period
This needs some explanation and I suggest to even include some analysis in order to show that there is no effect induced by including such observations. However, my understanding is that plots are selected in order to study plant-pollinator interactions in reliable and replicable spatial areas using a time frame (here 30 min) that will allow to have comparable data. Although the field sites were exactly set-up to have comparability this is then weakened through adding additional “rare” observations. Usually we are interested in the major effects that occur in (eco)systems, then I would like to understand more clearly what the value of “rare” observations is and why they will improve my dataset and subsequently my analysis.

This is explained in text.

P7-L143-148: “Furthermore, we included some interactions between plant and pollinator individuals that were not observed during the sampling but that were opportunistically recorded immediately before or after the sampling periods, as some of these interactions are difficult to document and might be important to define network structure (Jordano 2016). These opportunistic interactions represented 22.96% of all interactions recorded.”

Moreover, we now provide all the results of analysis excluding rare observations, which do not change the main conclusions (see Tables S12A-H)

L227-230 how did you derive these thresholds of 0.25, 0.5, and 0.75? Could be nice to apply tools to derive such a threshold directly from the data (the distribution of the original data).

This is a good point, originally, we were cutting at 0.25, 0.50 and 0.75 raw values, but we have now changed this to reflect the 25, 50 and 75th percentiles, which for our data are 0.63, 0.80 and 0.91. We have repeated all the analyses with these new values and obtain very similar results, although the negative effect of pollinator richness for the 0.75 value now disappears.

P13-L268-272: "We therefore calculated a measure of equity in reproductive success at the site level as the proportion of species with normalized (between 0 and 1) average fruit set values that were above the 50th percentile. As any selected threshold is arbitrary, we repeated this using the 25th and 75th percentile thresholds (Byrnes et al 2014)."

L262/263 …1472 pollinator individuals belonging to 57 species of plants… there is something wrong some rewording is needed

This was a mistake and it has now been amended.
P15-L307-308: "Within our sampling we recorded 1,472 plant-pollinator interactions involving 277 pollinator species and 57 plant species."

L290-292 both models are equally good but model 2 is more complex and hence you decide to report data for this model only.
This for sure needs more explanation than what is given here. I think it is generally accepted to go for the simplest possible explanation which would be the less complex model. Also known as Occam's Razor.

We have now included R^2 values for all models and show that having the same fit (very similar AIC values), model 2 shows greater predictive abilities and hence we keep that model. Note that we are using AIC to show the trade-off between model complexity and model performance, and not to select the simplest model. Moreover, as model 1 is nested in model 2, by reporting model 2, we are showing all relevant estimates.

L297-299 please when you mention such things then show the analysis (or at least link to it in some supplementary material)

This part has now been removed from the text following some changes to the analyses.

L308-309 the negative relationship between equity and pollinator species diversity using a threshold of 0.75 what does it mean considering that the threshold is set arbitrarily? This is not really clear, needs explanation. See also my comment above about choosing these thresholds.

The use of thresholds is common practice when evaluating equity (Byrnes, J. E. K. et al. Investigating the relationship between biodiversity and ecosystem multifunctionality: challenges and solutions. Methods Ecol. Evol 5, 111–124 (2014)). Selecting thresholds is often arbitrary, and as we lack *a priori* knowledge on which thresholds are ecologically meaningful (e.g. is having 50% of seed set in all plants enough for having viable

popuations?). Hence, using a variety of thresholds is often recommended. We clarified this in methods now.

P13-L268-272: “We therefore calculated a measure of equity in reproductive success at the site level as the proportion of species with normalized (between 0 and 1) average fruit set values that were above the 50th percentile. As any selected threshold is arbitrary, we repeated this using the 25th and 75th percentile thresholds (Byrnes et al 2014).”

Reviewed by Nicolas Deguines, 2019-05-28 14:18
Download the review (PDF file)

Reviewer comments to authors
General comments
The manuscript authored by Magrach A. and colleagues investigates the drivers of plant reproductive success in natural ecosystems of SW Spain using a plant-pollinator network approach. Specifically, this manuscript looks at whether ‘simple’ metrics (pollinator richness, number of visits) can be used to predict reproductive success or if more complex metrics informing of the network structure can be more informative and provide clues to the mechanisms involved in the patterns observed. For this, they use their own dataset collected from 16 sites in SW Spain, each surveyed for 7 visits in 2015. The dataset is impressive and seems adequate for the aims of this study. The research presented here is quite novel as looking into how exactly the structure of plant-pollinator networks affect plant reproductive success in natural plant communities, which would allow to empirically validate previous theoretical and experimental works on this issue.
Overall, this manuscript is very interesting; I am largely positive about both the questions asked and the data used, and I am convinced it will be a nice addition to the literature. However, on a less positive note, there are issues that need to be addressed. There is some substantial work to be done to clarify a number of different points (e.g. reducing the possible confusions to readers between overlap and complementarity, statistical analyses: e.g. data transformation and choice of distribution family for inclusion in glmm’s) and to provide more details (e.g. choice of focal plant species for measuring reproductive success). I also thought the discussion did not address all the points raised by the results. The abstract states that simple metrics can be used for prediction purposes but I felt this was not enough backed up by both the results presented and the discussion, and would need more arguments in the discussion. I below provide detailed comments that I hope will help authors improve this manuscript.

We thank the Reviewer for the thorough review of our paper and for the list of suggestions provided which we are sure will only improve our paper.

Detailed comments
Title
I wonder if the title could be improved because when reading ‘Interaction network structure maximizes’, I thought: ‘what exactly in the structure?’.
Perhaps: “Niche complementarity in interaction networks maximizes community-level plant reproduction success”?

Following the Reviewer´s advice we have now changed the title to:“ Niche complementarity between pollinators increases community-level plant reproductive success”

Abstract

L29-30: After reviewing the paper, I feel that this result (simple measures can be used for predictive purpose) is not enough backed-up and lacks specifics. Is it more suitable than complex ones or the latter do better? What's the direction of the prediction? What are the implications (e.g. can we infer plant reproductive success from pollinator richness estimates based on captures from pan traps?) ? I provide complementary comments about this below.

We have now provided answers to the queries presented here by the reviewer and added a line to the abstract stating the direction of the relationship that we find:

P2-L34-35: "Specifically, we find a positive effect of increasing niche complementarity between pollinators on plant reproductive success."

Introduction

L58: change 'but also they found that community structure had' to 'but community structure also had'

Changed.

L75-78: I suggest being more specific here about exactly what is known from the literature on network structures and pollination. For example, authors wrote that 'a prevalence of nested structures' and 'presence of asymmetric specialization as a pervasive feature' shape mutualistic interactions: how exactly? What are the effects of nestedness and asymmetric specialization on mutualistic interactions? More stable or robust to extinctions? More efficient plant pollination? Also, there would be a need to define nestedness (and perhaps asymmetric specialization? It may be self-explanatory, I'm not sure).

We have now provided further explanations to this part of the Introduction.

P5-L82-89: "Especially, in the case of pollination, we have ample knowledge on the attributes that shape these mutualistic interactions at the community level. Amongst them is the prevalence of nested structures, i.e., arrangements where specialist species interact with a subset of the species that generalists interact with (Bascompte et al 2003) and which is thought to promote species diversity (Bastolla et al 2009); or the relatively high extent of complementary specialization at the community scale, which may be directly related to key ecosystem functions (Blüthgen & Klein 2011).

L66-83: In this paragraph overall, I would suggest using less wordings such as 'considerable understanding' and instead try to be more specific on what is known. That would help readers understand what is not known and thus what's interesting in the current MS. For example, L81 mentions it's time to 'use this knowledge' but the paragraph does not provide clear facts.
Also, it should be here emphasized that a strength of the paper is to answer questions with empirical data from natural communities (compared to theoretical or experimental, as said L336-338). Also, how is the approach different from papers cited at L337-338 – Valdovinos et al. 2016 or Poisot et al. 2013)

We now have reworded part of this section following the Reviewer´s comments.

P5-L90-96: "The time is thus ripe to use the existing knowledge around plant-pollinator network structures to explore the relationship between network structure and

ecosystem functioning empirically, with special emphasis being placed on the underlying ecological mechanisms that drive these relationships."

Here, we present one of the first efforts linking pollinator visitation and plant reproductive success at the community level using empirical data on plant-pollinator interaction networks and plant reproductive success."

L87: change 'at Mediterranean' to 'in Meditereanean'.

Changed

L88: what is the meaning of 'area of influence' of the Donana National Park? Is it a legal concept (e.g. outside the core area of the park but still has to be managed according to some rules from the parks) or do you mean ecological influence (e.g. not far from the park)?

We have now explained this in more detail in the main text.

P6-L127-129: "Our study was conducted in SW Spain within the area of influence of Doñana National Park, i.e., within the limits of the Natural Space of Doñana as defined by the local government (Junta de Andalucía, Fig. 1)."

From start to L97: I think I got confused because I understood that authors use 'community' and 'network' for the same meaning, as on L97 mentioning 'community structure'. But from paragraph L66-83, nestedness and asymmetric specialization are characteristics or 'attributes' of interaction networks. I would suggest homogenizing, using rather network than community (the latter often used as a synonym of assemblage, e.g. bird assemblage or bird community).

Following the Reviewer´s suggestion we have now changed the term community for network or site referring to network structure and site-level metrics throughout the text.

L96-97: I think that about here somewhere, it should be clear to readers what measure of network structure authors will use.

We now introduce at the end of the introduction the type of metrics we will focus on.

P6-L107-114: "Specifically, our study focuses on how the interplay between niche complementarity and redundancy determines reproductive success. Plant reproductive success requires of the delivery of conspecific pollen and thus of a certain degree of niche complementarity (Blüthgen & Klein 2011). Yet, greater values of nestedness, which imply redundancy in species functions, are thought to promote species diversity (Bastolla et al. 2009) and stability (Thébault & Fontaine 2010) within plant-pollinator networks. At present, we do not know how either of these network characteristics affects the functions performed by pollinators."

Methods
L109: 'area of influence' again

We have now explained this in more detail in the main text.

P6-L127-129: "Our study was conducted in SW Spain within the area of influence of Doñana

National Park, i.e., within the limits of the Natural Space of Doñana as defined by the local government (Junta de Andalucía, Fig. 1)."

L113: by 'recorded', do you mean 'captured'?

Recorded means that we noted the species of pollinator visiting a plant species. This was done for species that could be readily identified by observation only. Those that could not be identified only were captured. We have reworded the text to clarify this.

P7-L135-140: "Along each transect, we identified all plant species and recorded all the floral visitors that landed on their flowers and touched the plant´s reproductive parts during each 30-min period. Only floral visitors (from now on referred to as pollinators) that could not be identified in the field were captured, stored and identified in the laboratory by FPM and experts in the different taxonomic groups (see acknowledgements)."

L114: 'legitimately' would need a definition here (the one used by fieldworkers when capturing pollinators). Indeed, there can be bumblebees robbing for example. However, how were considered flies from the Bombyliidae family or diurnal moths from the Sphingidae family for example? The two can take nectar from flowers without even landing and thus often without getting pollen grains on their bodies.

We considered a visit to be legitimate when we observed a pollinator landing on a flower. This has now been specified in the Methods section.

P7-L135-137: "Along each transect, we identified all plant species and recorded all the floral visitors that landed on their flowers and touched the plant´s reproductive parts during each 30-min period."

L115-117: How did authors know before the fieldwork what were the 19 most common species across the study area? Also, how could authors make sure they were following three individuals from the same species or that those three actually differed from other individuals without capturing them and identifying them? Finally, what information was recorded during those 3 minutes of focal observation / what was the purpose? update: I understand know that *plant* individuals were followed, not pollinators. Authors should precise this. Also, the final comment remains: precise the information collected during the 3 minutes and how it was used. Were the 'pollinator diversity' and 'total number of visits' taken from those three minutes of observation?

The authors have been working on the study area for the past years and thus know from previous surveys what the most common plants are. This has now been explained in the text. We have also clarified focal individuals were plants and what was recorded during that time.

P7-L140-143: "In addition, at each round we conducted 3 minutes of focal observations recording all floral visitors observed on 3 plant individuals per species belonging to the 19 most common (based on previous surveys) plant species across the study area (mean ± SD: 6.25 ± 1.73 species per site)."

L117-121: If not said later, readers would be interested in knowing how many such opportunistic interactions were recorded and included in the network (perhaps as an average % per site).

We have now included the total percentage they represent in our whole sample.

P7-L147-148: “These opportunistic interactions represented 22.96% of all interactions recorded.”

L109-125: This paragraphs lacks the ‘time of day’ information regarding sampling. Was each site sampled both in mornings and afternoons?

This information has now been included.

P8-L149-150: “Surveys were done during mornings and afternoons with the sampling order being established randomly.”

L127-132: I suppose there were more than 19 plant species in each transects. What are the species sampled? How were they chosen? I am not a plant expert but for practical reason (e.g. plants with long and differential flowering – not all flowers at the same time, i.e. fruits present over several weeks), could the choice of plant being sampled somehow bias the results if only some types of plants are sampled and appear to be those more visited by particular visitors? What about self-(in)compatibility of these plants?

These plants were selected based on their commonness. This has now been reflected in the text. We also provide more information on their compatibility in Table S4.

P8-L158-159: “Plants species were selected based on their availability with sampling being focused on the most abundant plant species.”

L153: I had trouble understanding ‘which covers’. Change to ‘which is an estimation of / which measures’ maybe?

Changed to “which estimates”.

L179-184: This sentence needs to be divided.? Perhaps stop the first sentence after first mentioning of ‘NODFc’. Then starts with ‘NODFc is calculated as:’

This text has now been re-arranged.

P11-L217-223: “This normalized value is calculated as $NODF_c = NODF_n/(C * log(S))$, where C is connectance and S is network size. $NODF_n$ is calculated as $NODF/max(NODF)$, which is independent of network size and thus comparable across different networks (Song *et al* 2017). To calculate max(NODF) we used a recently corrected version of the algorithm (Simmons *et al* 2019) in all but three sites, where the condition that the number of links>number of species was not met and thus precluded us from using this new version.”

L198-199: it is unclear to me to what refer ‘both at the species and the community-level’. The way it is, I thought it applied to reproductive success but in section ‘plant reproductive success’, there is no mention of a ‘community-level’. Then I suppose it may be the network metrics at species and community-level? But then, that would mean that L197 uses ‘community structure’ to encompass both species-level metrics and community-level metrics (see also my previous comments on using ‘community’ or ‘network’). This should be clarified.

We have now changed the community term for site across the text, which now reads

P11-L238-239: "…both at the species and the site level."

We have also included how the site level metrics were calculated in the section Plant reproductive success.

P8-L159-161: "The values at the species level were then averaged per site to calculate unique reproductive success measures at the site level."

L202-204: 'All response variables': so these are 'reproductive success', correct? I suggest having two sentences: the first specifying what are the response variable, and the second explaining that prior to be analyzed, each species reproductive success values were scaled (across sites then?).

We have now changed this because it was not reflecting what we actually did. Reproductive success measures include fruit set, the number of seeds per fruit and fruit and seed weight. We only scaled the number of seeds per fruit to allow for comparison across species with different life histories but did not scale fruit set values as they are readily comparable across species.

P12-L245-250: "At the species level, response variables included the fruit set for different individuals of each species analyzed using a binomial distribution, the average number of seeds per fruit analyzed using a normal distribution, and the average values of fruit and seed weight fitted to Poisson distributions. The number of seeds per fruit was centered and scaled (i.e., we subtracted column means and divided by standard deviation) to allow meaningful comparisons across species with contrasting life histories."

L205-206: some paper present results in terms of both richness and diversity (e.g. Shannon index of diversity) so here I strongly suggest sticking to richness (i.e. the number of species, and thus using 'Pollinator richness') if this is what is used by authors, and not use 'Pollinator diversity'.

We have now changed the term pollinator diversity for pollinator richness throughout the whole text.

L198 & L205: after reading the next paragraphs, data are first analyzed at the individual level: all individual values from all species are included in two models: model 1 and model 2. Then, shouldn't it be 'Models analyzing data at the individual level' instead of 'At the species level' which I first understood as one model per species.

We have multiple values per species as repeated measures (hence the use of random factors) but the focus of this analysis is the species as the network metrics calculated are at this level.

L208-209: what do the authors mean by 'In this case' ? Model 1? Model 2? Species level models 1 and 2?

Here we mean for both models. We have changed the text accordingly.

P12-L253-255: “For both models, we included plant species nested within site and site as random effects to account for the non-independence of several individuals measured per species and site. …”

L211-214: This should arrive prior to the explanatory variables: indeed, readers need to first understand that response variables include values at the plant individual level, taken from different species in different sites. Without this, the need for the random effect (L208-209) is unclear.

We have now changed the order to show first response variables and then the explanatory ones.

P12-13-L245-264: “At the species level, response variables included the fruit set for different individuals of each species analyzed using a binomial distribution, the average number of seeds per fruit analyzed using a normal distribution, and the average values of fruit and seed weight fitted to Poisson distributions. The number of seeds per fruit was centered and scaled (i.e., we subtracted column means and divided by standard deviation) to allow meaningful comparisons across species with contrasting life histories. As explanatory variables, model 1 included pollinator richness, and the total number of visits received by each plant species; while model 2 added the two network attributes calculated at the species level: average plant niche overlap and centrality. For both models, we included plant species nested within site and site as random effects to account for the non-indepedence of several individuals measured per species and site.

At the site level, we upscaled our species-level analyses. As response variables we had the average reproductive success per site (i.e., average fruit set analyzed using a binomial distribution, average number of seeds per fruit and average fruit and seed weight using a normal distribution). We thus had a single value per site and no random effects are needed in this case. In this case, model 1 included total pollinator richness and total pollinator abundance (i.e. number of visits received by all plants within the community) as explanatory variables. Model 2, in turn, added information on network structure by including nestedness and pollinator niche complementarity as explanatory variables.”

Also, I’m not sure I get what means the ‘average values of’, why not simply ‘average fruit and seed weight’? Actually, a previous section mentioned ‘the average seed weight per fruit’. Some clarifications are needed here.

Changed.

Finally, Poisson distribution works with strictly positive integer values; averages of weights that had first been scaled should produce some negative values and decimal values. Unless I misunderstood the verb ‘scaled’ that I took as ‘standardized’ (i.e. substracting the mean to each value and dividing by the standard deviation ; https://www.rdocumentation.org/packages/base/versions/3.6.0/topics/scale).

See also comments about L229 for the ‘binomial’ model with data that had been scaled (i.e. about ranging from -1 to 1 in my definition of scale, resulting from the above mentioned R function).

As stated before, this was an error in our writing and we have now changed this to reflect

what we actually did. Reproductive success measures include fruit set, the number of seeds per fruit and fruit and seed weight. We only scaled the number of seeds per fruit to allow for comparison across species with different life histories but did not scale fruit set values as they are readily comparable across species.

P12-L245-250: "At the species level, response variables included the fruit set for different individuals of each species analyzed using a binomial distribution, the average number of seeds per fruit analyzed using a normal distribution, and the average values of fruit and seed weight fitted to Poisson distributions. The number of seeds per fruit was centered and scaled (i.e., we subtracted column means and divided by standard deviation) to allow meaningful comparisons across species with contrasting life histories."

L220: 'average reproductive success' => authors mean 'per site', correct? I think here it needs to be clarified.

Changed.

L229: is 'normalized' used as 'scaled'? or did the authors normalized to a given range (as done by https://www.rdocumentation.org/packages/BBmisc/versions/1.10/topics/normalize)? Perhaps this should be clarified to become obvious to readers what was done by scaling or normalizing, especially to understand which (g)lmm's were used. For example, I questioned above the use of the Poisson distribution and I realize now I should also have questioned the use of the binomial family given that values had been 'scaled', which I took as resulting to values approximately being between -1 and 1.

This has now been explained in a previous comment. Only the number of seeds per fruit was scaled.

P12-L245-250: "At the species level, response variables included the fruit set for different individuals of each species analyzed using a binomial distribution, the average number of seeds per fruit analyzed using a normal distribution, and the average values of fruit and seed weight fitted to Poisson distributions. The number of seeds per fruit was centered and scaled (i.e., we subtracted column means and divided by standard deviation) to allow meaningful comparisons across species with contrasting life histories."

L256-258: different parameters can be extracted from a model and authors should precise to which parameter their expectation applies. I would suppose it's the beta estimates (i.e. the slope of the effect of the network predictor on equity).

We have now specified this in the text.

P14-L301: "…and extracted the model slope parameter estimates"

Statistical analyses: In all models, authors should mention how assumptions were checked and if they were met; additionally, Variance Inflation Factors [see (Zuur et al. 2009)] should be computed to check collinearity between explanatory variables, especially perhaps between "niche overlap" and "niche complementarity" which I expect to be negatively correlated, somewhat strongly?

We did check for collinearity between variables using variance inflation factors, but forgot to state it in the methods section. We have now done so. In addition, we ran residual diagnostics for our residuals and this has now also been explained in the methods.

P13-L274-276: “In all cases, we used variance inflation factors to check for collinearity between explanatory variables. Additionally, we ran residual diagnostics to check if model assumptions were met.”

Results
L262-263: ‘1472 pollinator individuals belonging to 57 species of plants’ appeared weird to me. I suggest instead ‘1472 plant-pollinator interactions involving 57 plant species and 277 pollinator species’.

Changed following Reviewer´s suggestion.
P15-L307-308: “Within our sampling we recorded 1,472 plant-pollinator interactions involving 277 pollinator species and 57 plant species.”

L264: are % corresponding to species or abundance? Perhaps provide both?

% here corresponds to individuals and has now been specified in the text.
P15-L308-310: “Within the pollinator community the distribution of individuals in different orders was: 87.84% Hymenoptera, 6.78% Diptera, 4.05% Coleoptera and 1.09% Lepidoptera.”

L291-292: this is of debate I suppose, but I tend to conclude the contrary: the added complexity is useless because it does not improve the AIC, and therefore the simpler the better. Instead of choosing between the two models, did authors consider starting with a full model (i.e. model 2) and prune out explanatory variables with no significant effects? Or simply keeping it whole so that each estimate is adjusted for other variables. That would also simplify the method to not have to present two types of models. What I suggest is only valid if there is no collinearity among predictor variables; if there was then, the approach of the authors must be revised as well.

As we have stated earlier in previous comments, we have checked for collinearity between predictors and this is not the case in our analyses and thus our approach is correct. We disagree here however with the Reviewer. Our use of AIC is to highlight the tradeoff between complexity and predictive power, not to select the simplest model. In this case our complex models show better or similar AIC values to the simple models but show better R2 and are thus considered globally as showing a better fit. We have now included all R2 values for our models in our results.

L291-296: typically, here and given Figure 3, could it be that niche complementarity and pollinator diversity are negatively correlated, so much that having them in the same model is not possible? This needs to be checked by computing VIF values.

This has been checked and is not a problem.

L297: Is the effect of total number of visits in table 2B significant? It’s not in bold there.

No it is not, this was an error from a previous version and reference to its significance has now been removed.

L300-303: I was not able to find Supplementary materials and could not check see those results. After reading the discussion, I strongly recommend moving those results in the main text. They are crucial to discuss the unexpected results of pollinator diversity and total number of visitation being negatively associated to fruit set.

We are now providing the Supplementary material which we intended to do in the 1st round but somehow missed. We have also included this information in particular in the main text now as part of Figure 4.

L308: 'However' is misleading and confusing here and seems not appropriate with 'also' just after. If there was an expectation, either it should have been clearer before or perhaps authors can shortly re-state it here. Or if authors want to insist on the fact that it's similar to the effect from pollinator diversity on fruit set, then I suggest stating it here simply.

"However" has now been removed and changed to "Also, we find…"

Additionally, information is missing: model 1 is best, but are there any significant effects besides pollinator species diversity using the 0.75 threshold?

This analysis has now been re-done following the use of new thresholds based on percentiles and this effect is no longer significant.

L313: I had no access to Supplementary Materials so I am unable to see Figure S4, which is a shame.

Supplementary material is now available.

Discussion
L318: 'a relationship' => 'relationships' given the generality of the sentence and current MS's results.

Changed.

L320: I need to admit here that I am not enough aware of the literature to confirm or not that mechanisms have been elusive 'until now'. However, has there been no papers since Thompson et al. 2012? Especially given that the sentence does not precise which function and which attributes of network structure. If need be, authors should precise a little bit (e.g. mutualistic interactions; or plant-pollinator networks and plant reproduction in natural communities, …). What of Valdovinos et al. 2016 for example?

There have been few attempts in general, and specially in plant-pollinator networks. Those are cited in the introduction, e.g., Fründ et al. 2013 using experimental communities.

L325: is niche complementarity between pollinators a measure of overlap or rather the contrary? The greater the 'Niche complementarity' indicator, the less overlap among pollinators, correct? If so, this sentence should be adjusted. Overall, I would suggest using more 'complementarity'

throughout the MS than overlap so that most sentences can be read and understood in the same 'direction' (e.g. 'complementarity increased community-level fruit set' instead of 'overlap decreased community-level fruit set').

We have now changed this sentence to improve clarity.

P18-L371-373: "At the site level, we find that greater values of niche complementarity between pollinators result in larger average fruit sets and number of seeds per fruit."

I'd like to insist on this point because both are used in the analyses but at different level: species-level, i.e. Table 1, test for an effect of 'Niche overlap' but community-level analysis, i.e. Table 2, test for an effect of 'Niche complementarity' (L208 & 219). Both consider plant niche (i.e. the first overlap in pollinator visitation, and the second complementarity in pollinator visits).
I realize I have been confused all along the MS and I here suggest choose to use overlap or complementarity in both (e.g. using "Morisita index x (-1)" to get a complementarity here as well?). I would further suggest to add 'plant' (leading to either 'plant niche overlap' or 'plant niche complementarity').

Throughout the tables and figures we have now specified that at the species level we measure plant niche overlap, so overlap in the use of pollinators between focal plants, while at the site level we focus on pollinator niche complementarity.

Additionally, L325: 'niche complementarity between pollinators' but L187-190: "[…] between plants": so which is it? Niche complementarity among plants (for pollinator visit) or among pollinators (for plant resources)?

Explained in previous comment.

L330-332: Given the materials provided, I'm not convinced by the predictive power of the simple visitation metrics. Readers would be very interesting in knowing such metrics can be used but I feel that more information is needed. For example, to predict Fruit set at the community level, is using Pollinator species diversity as efficient as Niche complementarity? If using those in bivariate regressions, what would be the R•? This is thus true for what's shown in Figure 3 and Figure 4.

R2 values have now been provided throughout the results section.

This paragraph, with more specific information on which of the simple metrics can be used for informing on what, would be better placed toward the end of the discussion as this is more of an 'application output' while the main message of the MS is about the role of Niche complementarity in plant reproductive success?

Additionally, more than the predictive power, there is a need to describe and explain the relationships between the 'simple' metrics and the 'network' metrics. In particular, they do not measure the same things, but I felt the paper presented the comparison as a test for using simple ones as substitute for more complex time-consuming ones.

This is not the case, note that the models are additive. So, network metrics can complement the simple metrics, but not substitute.

P12-L242-244: “These models are meant to be additive, so that the network metrics included are intended to complement rather than substitute the simple metrics traditionally used.”

L334-335: To avoid redundancy, I would erase the end of this sentence (i.e. stop right before “and, in particular”] because this is explained from L336.

Removed.

L342-346: In my opinion, the general idea in the literature is that diversity increases functions and services (many articles from D. Tilman, or (Cardinale et al. 2002) for example). Here however, pollinator species diversity is associated to lower fruit set and equity in fruit set. A convincing explanation is provided but I think it’s worth to note and discuss the discrepancy (from my knowledge of the literature) with previous research. Similarly, greater abundance of pollinators on flowers (nb of visitation) usually leads to greater fruit set (Garibaldi et al. 2013) but that is not what is found here (Fig. 3), so why is this?

We understand the point raised by the Reviewer and we have therefore tried to improved our explanation of the mechanisms we think lead to this difference.

P20-21-L426-442: “One of the unexpected results of our analyses is the strong negative effect of pollinator richness for fruit set at the site level. An explanation to this might be the fact that pollinator richness here includes all the pollinators recorded during our sampling efforts, i.e., it includes species that do not pollinate some of the species whose reproductive success was measured. More complex communities with more pollinators, but also with more plant species (Pearson correlation between plant and pollinator richness = 0.42 in our case) may require stabilizing mechanisms that reduce the competition exerted by the dominant plant species. A way to reduce the competition exerted by these dominant species, which are precisely those evaluated in this study, is by reducing their reproductive success (Lanuza et al 2018, Stavert et al 2019). These ideas open the door to exploring the positive or negative effects of the complete pollinator community on full plant species coexistence, which may be determined by density-dependence effects (Benadi & Pauw 2018). In our case, while fruit set is negatively related to pollinator richness, it is important to note that fruit and seed weight show the opposite relationship, indicating that this density-dependent effect might only be limiting fruit quantity and not fruit quality. Thus, taking into account the densities of co-flowering plant species may be the next step (Vanbergen et al. 2014).”

In particular, if ‘total number of visits’ can be used for predictive purposes, it is not in the direction that is usually thought (i.e. more visitation leading to better fruit set, not the contrary as found in Fig. 3B).

This effect is no longer shown once new analyses have been done using the new NODF metric as we had a miscalculation in the earlier version.

L341: remove “values” from “greater fruit set values” ? “fruit set” alone is understood.
Removed.

L342: in “species requires of the”, remove “of”

Removed.

L342 + L345-346: The first states that reproductive success needs conspecific pollen, and the second state a somehow opposite idea that there can be interference with conspecific pollen. This requires more explanation. This relates to a previous comment above about the choice of plants that requires more details. For example, did authors choose self-incompatible plant species, i.e. "the ability of plants to reject their own pollen" (Tovar-Mendez and McClure 2016)? In that same Dispatch paper from Tovar-Mendez & McClure, they argue that domesticated plants (i.e. crops) were selected for self-compatibility. Would that explain why the present MS found that pollinator diversity and total number of visits have negative effects on fruit set whereas Garibaldi & colleagues (2013) (among others) found the opposite for crops? I strongly recommend adding more information from a plant perspective in the introduction, the methods, and the discussion. It is critical for understanding the function (plant pollination) authors seek to explain with measures of reproductive success.

This paragraph has been reworded for clarity (see response above). In addition, we report information on self-compatibility for all plants analyzed. We refrain to speculate about differences between crops and wild species, as we don't have enough information to compare those.

L354-369: I feel this whole paragraph could be reduced as the point for looking at equity in fruit set was well-made before in the MS. I suggest to then lump it with the next one discussing what is found here.

We believe that given that the focus on equity is a novelty in our paper it deserves to be well explained in the text and therefore if there are no space constraints would rather keep this paragraph.

L373: unfortunately, I could not see this 'dramatic' effect. It is the first time I read about fruit set equity and I think it's a very important facet of the pollination function and the community-level plant reproductive success. There are already five figures; however, if possible, moving Figure S4 (which I did not see) to the main text could be useful given it is discussed and given the importance of fruit set equity. Unless authors think that with "only" six species maximum per community, it is not enough for making a strong point presently (in which case then perhaps using 'dramatically' is too much and more caution should be used).

We have now moved Figure S4 to the main text which shows the strong effect and is now Figure 5. We removed the word "dramatic".

L378-380: given my previous comments, I suggest discussing the unexpected results before. Then, only once this is 'cleared', discuss other things. Also, currently, it's not clear if there were significant effects when using a different threshold than 0.75.

We would like to keep the order as is, given that we try to provide a very thorough explanation of this unexpected result. We thus, want to first present all the "simpler" results and then focus on explaining this one.

L381: from the sentences at L205-207 and L284, I understood that pollinator species diversity was the one found visiting each focal plant species; as for total number of visits. But with L381, I become uncertain: did authors test for an effect of overall pollinator diversity in a community

with fruit set of each focal plant species (in its community)? I would understand testing the latter* but in which case, it would be very important to clarify this throughout the MS and show clearly results of both and explaining why it would be interesting to look at this.
*For example, it could be interesting when sampling pollinators with pan traps and therefore without the information of plant visitation, as a way to consider resulting diversity as an index of plant pollination.

This sentence has been changed to improve clarity. We do not refer to how can information on pan traps relate, as pan traps have other biases (e.g. pollinator body size) and do not link to visitation rates.
P20-L427-430: "An explanation to this might be the fact that pollinator richness here includes all the pollinators recorded during our sampling efforts, i.e., it includes species that do not pollinate some of the species whose reproductive success was measured."

L385-386: Could authors detail the mechanism(s) here by which there could be a reduction in the reproductive success of the dominant species? Is it related to plant self-(in)compatibility or type of reproduction?
Ok, I see one mechanism is introduced right after. Perhaps then modify this sentence, so that readers get there are mechanisms you will be discussing. Also, are there others or can there be more details on how density-dependence could play a role?

WE have now improved this explanation and provided extra references.
P20-21-L426-442: "One of the unexpected results of our analyses is the strong negative effect of pollinator richness for fruit set at the site level. An explanation to this might be the fact that pollinator richness here includes all the pollinators recorded during our sampling efforts, i.e., it includes species that do not pollinate some of the species whose reproductive success was measured. More complex communities with more pollinators, but also with more plant species (Pearson correlation between plant and pollinator richness = 0.42 in our case) may require stabilizing mechanisms that reduce the competition exerted by the dominant plant species. A way to reduce the competition exerted by these dominant species, which are precisely those evaluated in this study, is by reducing their reproductive success (Lanuza et al 2018, Stavert et al 2019). These ideas open the door to exploring the positive or negative effects of the complete pollinator community on full plant species coexistence, which may be determined by density-dependence effects (Benadi & Pauw 2018). In our case, while fruit set is negatively related to pollinator richness, it is important to note that fruit and seed weight show the opposite relationship, indicating that this density-dependent effect might only be limiting fruit quantity and not fruit quality. Thus, taking into account the densities of co-flowering plant species may be the next step (Vanbergen et al. 2014)."

L387: 'those evaluated in this study' => as previously mentioned, details needed in the method.

Rephrased

P20-L433-435: "A way to reduce the competition exerted by these dominant species, which are precisely those evaluated in this study, is by reducing their reproductive success (Lanuza et al 2018, Stavert et al 2019)."

L390-391: very interesting! I strongly suggest putting more results in the main MS rather than in the supplementary. At the very least in a table if there are too many figures.

We have now shifted some of the results from the Supplementary material section to the main text.

L400: it may be because I'm not a plant expert but I would need slightly more detains on what authors meant exactly by heritability. For example, which traits do they think researchers would need to focus on? Or is it about the effects of plant-pollinator network structure on the evolution of plant reproduction strategy?

We have now changed this paragraph and elaborated more on these ideas on the previous paragraph.

P22-L442-455: "Our study illustrates the complexity of linking network structure to ecosystem function empirically, because measuring both structure and function is challenging. For example, there is an ongoing debate as to which network metrics better reflect classic ecological mechanisms, such as niche partitioning or competition (Delmas et al 2018). Here, we focus on testing two specific hypotheses, but other structural properties can be explored when more data becomes available. Furthermore, the structure of plant-pollinators networks is dynamic due to ecological and evolutionary reasons, but so far, we are only able to characterize it for single snap-shots. Moreover, different aspects of functioning may be important, such as the presence of non-linear relationships or the need to consider the functioning of both trophic levels (Godoy et al 2018). In terms of plant reproductive success and the functions performed by pollinators we can measure different aspects, ranging from pollen deposition (the direct pollinator function), to its final effects on plant fitness. Here, we focus on an intermediate stage including fruit quantity and quality, which is of clear ecological importance. "

Missing from the discussion: How come at the 'species-level' (measures at the individual level used) there is a positive association between plant niche *overlap* and seeds per fruit (Fig 2B), but there is also a positive association between plant niche *complementarity* and seeds per fruit (Fig 4B). Niche overlap and complementarity should be negatively correlated and so these two results are opposite, right?

So in the discussion, on top of clarifying throughout the MS the overlap or complementarity wordings, here there is a need to discuss these results which appear to me as contradictory.

We have now included a reference to this in the Discussion section.

P18-19-L382-394: "Consistent with previous experimental (Fontaine et al. 2005; Fründ et al. 2013), theoretical (Pauw 2013), and empirical studies (Valdovinos et al. 2016, Poisot et al. 2013), we find that niche complementarity is key in determining differences in reproductive outputs. Indeed, we find that communities where there is less overlap in the niches occupied by pollinator species had greater values of reproductive success, both greater fruit set and larger numbers of seeds per fruit. This therefore reflects the fact that reproductive success in plant species requires the delivery of conspecific pollen and thus of a certain degree of specialization amongst pollinator species on a particular plant resource in order to avoid the negative effects of inter-specific pollen deposition (e.g., pollen loss, Flanagan et al. 2009) or

interference with conspecific pollen (Morales & Traveset 2008). However, we also find that some level of redundancy in these functions is needed as revealed by the positive effect of plant niche overlap on the number of seeds per fruit at the species level."

Tables

Table 1:

I am vaguely aware of some debate about providing or not p-values associated to mixed-effects models (or to any models nowadays), but I would prefer to have a p-value here as z. or t.values are not as straightforward. Given table 2A, I take it authors have no strong feeling against p-values. 95% confidence intervals would also be better than SE because they can be interpreted directly. This applies to Table 2 as well. If resulting tables become too wide to fit in the page, I suggest SE can be removed (provided there are 95% CI instead) and z or t values as well. Finally, unless this is against PCI Ecology editing rules, I would lump Table 1A and 1B and adding a 'Response variable' column to more easily what predictor affect which response. Same for Table 2A and B.

Given that it is not possible to get classical p-values with GLMM due to its nested structure and the current debate about the perils of using arbitrary threshold to stablish significances, we have removed all p-values and are only interpreting estimates of standardized variables and their uncertainty in order to compare the size of effects, which we believe is the most informative way of presenting results.

Table 2:

Some tested predictors are in bold but not others and it does not seem to correspond to being a significance? (Conversely to what is said in legend of table 1).

Bold letters represent strong effect sizes variables.

Figures

Figure 1:

I think authors should add a north arrow somewhere, for readership not familiar with the European geography. Also, size font for the scalebar and the Elevation range should be increased; for the elevation range and color scale, consider placing it in an inset as well or to add a black line framing each colored rectangles: currently, colors are not easily seen because not differentiated from the map

North arrow has now been included, font size changed and legend modified.

Figure 2-5: I am not familiar with partial residual plots and this was not described in the methods. When looking at values on x- and y-axes, it seemed to me the data were plotted and that the predicted model effect was added.

Partial residual plots are used to graphically show the effect of a single predictor by removing the effects of the others variables which are set to their mean value. This has now been added to the main text:

"**Figure 2**. Partial residual plot showing the effect of a single predictor for the relationship between A) plant species centrality and fruit set for each of the plant species considered and B) plant niche overlap and average number of seeds per fruit. Dots represent each of the individuals sampled for each species within each site."

Figure 2:
Font size for plant species is way too small for reading correctly. It would be better if imagining using a four panels-figure and spread the plant species across the full width and height that take panels A and B, but below (as if there were panels C and D)

More importantly, fruit set was said to be modelled with family=binomial but the effect of centrality goes beyond 1.0, which is not possible under this distribution. I suppose this is a plotting issue, as authors did use a glmm (as explained in the methods)?

This is a plotting issue with no points going over 1 in the graph.

To avoid any confusion, I suggest replacing, in B, 'seeds/fruit' by 'seeds per fruit'. Also in Figure 4

Changed.

Figure 3:
Y-axis should be kept equal across panels A to C ; as is, it's hard to compare the strength of each predictor.
In 'Pollinator sps diversity', might change depending on answer to a previous comment: but sps likely can be removed to have only 'Pollinator richness' or 'Pollinator diversity'. Also in Figure 5.

Changed.

Figure 4:
Is the effect of Total number of visits significant?

This effect is not significant and has been removed.

Y-axis should be kept equal across panels A to B; as is, it's hard to compare the strength of each predictor.

We have now changed the figure to make the Y-axis constant across graphs, although we have now merged Figures 3 and 4 and this is now Figure 3.

Supplementary information
I had no access to the supplementary either on PCI Ecology website or on biorxiv website. Did I miss them somehow or were they not available?

This was an error when copying the link. The supplementary material should now be available.

Data availability and details of quantitative analyses

According to guidelines of PCI Ecology to reviewers, I here state I did not see that data were made available on an open data repository yet. Similarly, I did not find a link to R (or similar) scripts for redoing the analyses.

We have now made all data and code available.

References used above

Cardinale, B. J. et al. 2002. Species diversity enhances ecosystem functioning through interspecies facilitation. - Nature 415: 426–429.
Garibaldi, L. A. et al. 2013. Wild pollinators enhance fruit set of crops regardless of honey bee abundance. - Science (80-.). 340: 1608–1611.
Tovar-Mendez, A. and McClure, B. 2016. Plant Reproduction: Self-Incompatibility to Go. - Curr. Biol. 26: R115–R117.
Zuur, A. F. et al. 2009. Mixed effects models and extensions in ecology with R. - Springer-Verlag.