Dear Henrique,

We thank your for your positive feedback and your additional constructive comments. We also thank the reviewer for his/her new work on our manuscript. Below is our reply to your comments and the comments of the reviewer.

Xuyue Yang, Martin Lascoux and Sylvain Glémin

<u>Answer to editor's comments</u>

I've now received one review of the revised manuscript you sent us, and I'm happy to write that I will recommend it for publication. This said, there are a few issues that you could address in a last round of revisions to strengthen the manuscript.

The last point of the reviewer is pertinent. It could be made clearer in the discussion that vegetative traits may be involved in direct competition between individuals, even if you have not measured them, while reproductive traits may be involved in reproductive assurance and/or colonization ability. If you could show that rosette surface is a fitness component, wouldn't the difference in the responses to your treatments be indicative that there is a trade-off between colonization and competition?
We have rephrased this part of the discussion (see answer to reviewer) but we don't see well how we could conclude about a competition/colonisation trade-off with this dataset. So we have not elaborated on this point.

On page 11, at least in the version I have, the equation for the competitive index is missing.
We hope the equation appears correctly now.

And in the last sentence of the paragraph you use the term "accession", shouldn't it be "species"?
You're right. This has not been corrected from the previous version where we initially estimated the competitive index for each accession separately.

Page 14, the genetic diversity in China is of 0.0015, right?
Yes, thank you for checking. It's corrected now.

On page 16, you mention one Figure 5, typo?
We meant Table 5: corrected.

I am not entirely sure that the statistical analyses you perform are well explained. First, I feel that some of the models are likely prone to collinearity. For example, in the model presented in table 3, I don't understand the rationale to include "rosette surface at t2" and "rosette" as separate factors; "growth rate" would be more appropriate? Can you test for collinearity even if in the end you don't present it in the manuscript? Stating only that you prevent over parameterization by dropping factors and their interactions does not seem to be sufficient.
Sorry for this confusion. Rosette stands for "rosette surface at t2" in interaction terms; so there is only one parameter. We have corrected this in Tables 3 and 5. We introduced rosette surface as a covariable in flower numbers analysis following a reviewer's suggestion.

Second, I think that the analysis and presentation would be much clearer and simple if you use the competitive index as the dependent variable throughout. At several places you write that a particular factor "explains" or "causes" differences in the competitive index, but in reality you never test it. A figure with the competitive index estimates (including confidence or credible intervals) would be nice.
We have suppressed the analyses on the competitive indices following the suggestions of a reviewer on our first version. The idea was that a test directly on interaction terms was more appropriate (that's what we did using contrast of contrasts using the testInteraction function). We have rewritten the corresponding text to be more explicit:
P12: "The differences among species can be summarized by the competition index (Figure 1) and tested by pairwise contrasts on the species x treatment interactions (Table S3)"
P14: we replaced "competition index" by "sensitivity to competition"

Third, the models testing for the effects of genetic diversity should include it as a covariate and not as a fixed factor. The main reason for this is that you did not manipulate genetic diversity levels and thus cannot test for interactions between what you think are the independent variables. It's not clear to me if area shouldn't be tested as a covariate or as a fixed factor. The model of Table 3 is also confusing in this regard.
OK. We have removed interactions with genetic diversity to keep only the term as a covariate. We have changed Table S6 accordingly.

One last comment is that I find that you easily equate levels of genetic diversity with the presence of deleterious genetic loads, which in turn will result in the short-term maintenance of selfing. Alternative and mutually-exclusive explanations are that selfing populations are initially better able to purge ancestral inbreeding depression, and thus better survive potential bottlenecks during range expansion than outcrossing populations, or because they actually have increased additive genetic variance for fitness traits, due to expression of dominance and epistasis, and thus respond faster to selection in novel environments. The observed levels of genetic diversity do not contradict these ideas. You suggest at one point in the discussion that measuring inbreeding depression in these populations would be important but I'm not sure that the reader will really appreciate these alternative explanations.

We have tried to be more explicit on this point in the introduction by adding/modifying the two following sentences:

"Selfing species are expected to rapidly purge inbreeding depression caused by strongly deleterious recessive alleles but can accumulate weakly deleterious mutations on the long term because of linked selection effects (Wright *et al.* 2008; Glémin & Galtier 2012). From a demographic perspective, the purging of inbreeding depression can help survive bottlenecks associated with initial colonisation of new habitats but weakly deleterious mutations can also accumulate later on during demographic expansion, the so-called "expansion load" (Peischl *et al.* 2013; Peischl & Excoffier 2015)."

Finally, in the code of conduct, it is specified that:

-Data for preprints must be available to readers after recommendation, in the preprint or through deposition in an open data repository, such as Zenodo, Dryad or institutional repositories, for example.

Raw data will be deposited on Dryad. Note that the manuscript will need to be accepted/recommended before being deposited on Dryad as PCI Evol Biol is not listed in the medias where deposition before acceptance is possible.

-Details of the quantitative analyses (e.g. data treatment and statistical scripts in R, bioinformatic pipeline scripts, etc.) in the recommended preprints must be available to readers in the text or as appendices or supplementary materials, for example.

-Details of the experimental procedures in the recommended preprints must be available to readers.

The R script used for all analyses and plotting has been added as a supplementary file.

Note also that a bioRxiv manuscript we were citing is now published: Park, D. S., A. M. Ellison, and C. C. Davis. 2018. Mating system does not predict niche breath. Global Ecology and Biogeography.

Answer to reviewer's comment

The paper by Yang et al. has been largely and adequately modified and corrected following our suggestions. We thus think that the paper is now ok for recommendation.

Minor comments

- The conclusions of the abstract must also be softened and more accurate: We suggest that the authors add the information that they detected competition effect on a reproductive traits but not on vegetative traits

Now we mention the two different results for vegetative and reproductive traits in the abstract.

- For the information of the authors, there are statistical procedures adapted to cases where there are many zeros in a dataset: zero-inflated models. See for ex. https://stats.idre.ucla.edu/r/dae/zinb/

Thanks for this information. Instead of a zero-inflated model we preferred a hurdle model where the probability of flowering is modeled independently of the number of flowers given that flowering occurred. We though this was more appropriate because it allows having different treatment effects for the probability of flowering and for the number of flowers instead of having a single inflation term. In addition, the distribution of the number of flowers is bimodal with either 0 or a much higher number (> 100) with no value in between. Actually we also tested a zero-inflated negative binomial model with the *glmmADMB* package but convergence was not properly achieved and the fit was pretty poor.

- p11: something is missing after the ":" or something is wrong with the beginning of a paragraph

Thanks for checking. Indeed the equation disappeared during pdf formatting. It should be OK now.

- p13 & p16: We are not sure to agree with the authors' statement that since weak differences were obtained for rosette surface then it is a not a good proxy for fitness (but maybe we misunderstood what the authors meant, which calls for a clarification). The authors' results may also mean that competition effectively does not affect fitness through vegetative traits. The authors' statement can look like an adhoc argument. An alternative conclusion might be (even if, of course, more reproductive and vegetative traits should be measured) that only reproductive traits are affected. A question would thus emerge: why would selfer vs. outcrosser species mostly differ by reproductive traits rather than vegetative traits? Are there any theoretical reasons for that?

We have suppressed the comment on fitness proxy on P13 (in the results part). In the discussion part (P16), we have rephrased it to explain that competition may not act through this specific vegetative trait (modifications bolded):

"A possible caveat that could explain the difference between vegetative and reproductive traits is that rosette surface estimates were less precise than flowers number estimates. ***In addition, rosette surface is a less integrated fitness component than flowers number****. Alternatively, although rosette surface is positively correlated with flowers number (Table 5)* ***this could reflect the fact that rosette surface is not differentially affected by competition in this experiment****. Indeed, access to light was likely not the main driver of competition in the experiment as differences in rosette surface could not explain (or if it did, only very weakly) the difference in flowers number between treatments (Table 5). In the experiment, most of the competition likely occurred at the root levels as we purposefully chose rather small pots.* ***Unfortunately, underground biomass is very difficult to measure when roots are highly entangled****.* ***A more detailed analysis of functional traits remains to be done to understand how competition is mediated and whether there is any difference between selfing and outcrossing species."***