**Determinants of lifetime reproductive success in a wild clownfish population: strong habitat and weak genetic effects**

Océane C. Salles, Glenn R. Almany, Michael L. Berumen, Geoffrey P. Jones, Pablo Saenz-Agudelo, Maya Srinivasan, Simon R. Thorrold, Benoit Pujol,and Serge Planes

**Responses to Reviewer Comments**

Dear Philip Munday,

We have revised our manuscript in response to the comments we received from the two reviewers. We thank them for their constructive comments that improved the manuscript. Specific replies to all of their comments are listed point by point below. We took note of all the suggestions and made a series of clarifications in the revised manuscript.

We now differentiated "self-recruitment" (*i.e.*, the process by which offspring return to the population of their parents to further reproduce) from Lifetime Reproductive Success (*i.e.*, which is an individual measure and a component of fitness restricted to the local population). As suggested by one of the reviewers, we moved the Delifing results and discussion to the Supplementary Information. We also explained the difference between size effects observed for the permanent environmental effects affecting LRS and DL.

We modified the *Figure 2* and added the distributions (histograms) for the Lifetime Reproductive Success and the variance components on both latent-scale and observed data-scale (pie charts). We also updated the estimates of fitness heritability and maternal effects on fitness in wild populations in the revised *Table S1*.

We trust that you will find our revised manuscript satisfactory. If you require any further information please do not hesitate to contact us. Looking forward to hearing from you.

Sincerely on the behalf of our coauthors,

Océane C. Salles and Benoit Pujol

**Reviewers' comments**

*Reviewed by Loeske Kruuk, 2018-10-03 01:47*

**Salles *et al.* Habitat variation of wild clownfish population shapes self‐recruitment [1] more than genetic effects**.

There is increasing interest in, and ability to determine, the genetic basis of variation in phenotypic traits in wild populations. However the literature to date suffers from a massive constraint in being highly taxonomically restricted: as the review in this paper’s Table S1 and other similar publications show, estimates of the heritability of fitness are almost all in terrestrial, temperate‐zone birds and mammals. This makes these data from a wild clownfish population a really valuable addition to the literature. I cannot emphasise this enough: it is really exceptional to have such measures for a wild marine fish, and very exciting to see them used for quantitative genetic analyses. So I would very much like to see this published! But also, precisely because they’re such rare/valuable data, it would be good to see much more detail about them. I also have several major questions about the analyses, and the paper does needs considerable work on presentation.

> Thank you for emphasizing the originality of these results. We considerably worked on the presentation of the paper and modified the text following all your comments.

**General/major comments**

The paper needs careful proof‐reading and attention to the language. Overall, it felt rather unpolished, and almost like a series of notes rather than a coherent argument. It is also quite confusing for someone who doesn’t know the system. I’m not sure if there was a rush to get this published, but it feels like some more time needs to be taken with careful editing.

> We restructured the paper and polished the text to clarify its logical narrative which is now more linear. Coherence was greatly improved. We also had the paper proof-read with careful attention to the language by multiple native English speakers, who are also coauthors of the paper.

The Methods regarding the pedigree construction are extremely confusing.

>We added precisions in the *§Methods* (Lines 192 to 212) to clarify how the pedigree was constructed based on genotype data. Furthermore, supplementary information can be found in the pedigree paper that we published previously and to which we refer in the text (Salles *et al* 2016, PNAS).

I am not a fish biologist, and I appreciate that this may be standard terminology, but I found the term “self‐recruitment” puzzling. Would “local recruitment” be clearer, as really you mean recruitment to the local population, not to yourself? Almost every study of a wild population that one can think of considers only this “self” recruitment, but doesn’t call it such – hence (see my comment below regarding Table S1) you run into difficulties if you refer to it as one thing in the text, but other studies are calling it something else.

> As noticed by both the reviewers, the term "self-recruitment" is commonly and widely used in marine biology to refer to the process by which offspring return to the population of their parents to further reproduce. Natal philopatry can also be used in that sense. We now clarify this aspect and define the various terms in the *§Introduction* (Lines 93 to 101), in the *§Methods* (Lines 215 to 216). In this paper, we investigate the evolutionary potential of this mechanism by estimating the genetic and environmental variation of the reproductive success of local individuals by taking only into account the offspring that came back to settle in their population of origin. We clarified how this parameter characterizes the relative fitness of local individuals in the *§Methods* (Lines 216 to 222). As suggested by Loeske Kruuk, we now use the term “Lifetime Reproductive Success” to avoid “difficulties if [we] refer to it as one thing in the text, but other studies are calling it something else”, which clarifies the fact that this is an individual measure and a component of fitness [restricted to the local population].

Please show the distribution of your measure of offspring production – at the moment it’s very difficult to get a handle on it (or on how many observations you have per individual).

> We added the distributions (histograms) for the biennial estimate of the Lifetime Reproductive Success in *Figure 2*.

****

**Figure 2. Sources of variation in the biennial estimate of the Lifetime Reproductive Success (LRS) of the Kimbe Island orange clownfish.** Distribution of the biennial **estimate** of the LRS (histograms). Variance components on both latent-scale and observed data-scale (pie charts) for the biennial **estimate** of the LRS explained by Additive genetic (VA), Maternal (VM), Natal Habitat (VNH), Resident Habitat (VRH), Permanent Environmental (VPE) and Residual (VR) variances. These proportions were calculated from the values of the posterior modes of a quantitative genetics generalized linear mixed model analysis (for details see Table 1).

Individuals were in average sampled 2.88 ± 0.04 times (mean ± SE) over the 6 sampling periods (1% of individuals persisted over the 10-yr survey. We clarified this in the *§Methods* (Lines 180 to 181)*.*

This is fascinating data given the potential for individuals to change sex, but I can’t quite see how this is handled – is sex considered in the model at all?

> We did not consider sex in the model. We are conscious that estimating the additive genetic variance for LRS in separated sexes (and cross-sex genetic covariance) would improve our understanding of the population rate of adaptive evolution. We have recorded the sex of every individual: male sub-adults, male breeders or female breeders. These are the three possible categories in the population. However, in the absence of a very large data-set, estimates of sex dependent additive genetic variance might be biased because some effects will be hard to disentangle as the clownfish only changes sex in one direction (from male to female) and this sex change is always associated with a change of sexual partner. Up to date, we only have replicated measures for different sexes in the same individual for a small portion of the population (n=41). We wish to tackle this aspect in the next few years with longer-term data. Longer-term data should indeed allow us to obtain more data for individuals before and after their change of sex. We clarified this in the *§Methods* (Lines 175 to 176 and 184 to 190)*.*

The ‘delifing’ approach has obviously received a lot of attention since its publication in 2006, but my impression is that people have become increasingly aware of its potential complications in being defined by the population size (look up its citations for various methodological papers about it). This complexity is reflected here in that its largest variance component being ‘resident habitat’ – is this meaningful from an evolutionary point of view? The contrast between the two measures is potentially confusing, and not addressed in the Discussion – why would there be a huge permanent environment effect for SLR but not DL etc.? I also find it very odd (in fact inconceivable) that there is effectively no residual variance for DL. I would actually suggest move the delifing analysis to Suppl Info, and focusing the paper on the SLR counts.

> We followed Loeske Kruuk’s suggestion to move the Delifing results and discussion to the *Supplementary Information* (Lines 222 to 229 and *Supplementary Information*). We agree that caution must be taken in interpreting the Delifing results. We also think it is important to keep the Delifing results available to the readership for comparison and replication purposes since only two out the fifteen studies that estimate the additive genetic variance of fitness in wild populations made it available. We point out that the Delifing results corroborate our conclusions that little to no genetic variance and large habitat effects shape the contribution of local individuals to the local population growth.

> We have also wondered: “why would there be a huge permanent environmental effect for SLR but not DL etc.?” In the absence of a modeling approach where different scenarios would be simulated in order to compare DL and LRS, I’m afraid that our answer will remain speculative at best and should not be included in the text of the paper. One might speculate that precision issues play a role in the observed differences between the LRS and DL approaches. The comparison of effects based on different estimates characterized by different units and large confidence intervals might artificially increase the perceived contrast in the relative weight of the variance components between the different approaches. With that in mind, it is still surprising to observe that the permanent environmental effect has shrunk drastically in the Delifing approach. The resident habitat explained eight times more variance in the DL approach than the LRS approach, the natal habitat nearly the double. Both the residual and the permanent environmental components were drastically reduced (from around 30% to very low values) in the DL approach. In the LRS approach, both the permanent environmental effect (which accounts for repeatability between individuals and years) and the residual variation likely explained some degree of similarity between unrelated individuals. In the LRS approach, it is also likely that some of the effects were confounded with the temporal heterogeneity in population growth and survival (which is a heterogeneity accounted for by DL). In terms of biology, one interpretation of the contrasted results between LRS and DL could be that environmental similarity explained to some extent the variation of the LRS. When using DL, the explanatory factors might therefore have not captured this temporal heterogeneity linked to the habitat which generated similarity between unrelated individuals. Under such scenario, it is possible that the habitat variation in DL explained variation that was explained in the other model by the residual and permanent environmental effect.

I am also not clear as to the derivation of the heritability (or other proportion) parameters from the GLMMs. From Table 1 and Figure 2, it seems like you have used QGglmm to translate the variance components from the latent scale to the data scale, and then have just taken the proportions of these (as a proportion of the sum of those data scale components).

> This is indeed what we have done.

But this means that, for SLR, the only residual variance is actually the over‐dispersion term in the Poisson model latent scale – which ignores the additional residual variance from the scatter around the expected data scale mean (see equation 3 in de Villemereuil 2016). I am therefore puzzled at such high proportions of variance on the data scale in Figure 2, as my understanding of the transformation from the latent to the observed scale is that an additional source of variance is introduced (theta in equation 3c in de Villemereuil 2016). How does your sum of data‐scale components in Table 1 compare with the observed datascale phenotypic variance (and how do the QGglmm estimates of data‐scale mean and variance compare with your raw data)? There is some discussion of these issues in Wolak *et al.* 2018. I realise this won’t change your conclusions about the very low heritability of SLR on the data scale, but it has implications for the other variance components: overall, I expect all of the %s in Fig 2a to be smaller. I realise these are complex issues, but you need to be clear about the issues in de Villemereuil 2016. This is also another reason for presenting information about the actual phenotypic data (means, overall variance etc).

>We have followed the suggestion and clarified this aspect of the calculation of the residual variance component on the data-scale in the text in *§Methods* (Lines 272 to 287) after a discussion with Pierre De Villemereuil. We also present the mean and phenotypic variance estimated on raw data (Lines 294 to 295) and the estimates (and relative proportions) produced by MCMCglmm on the latent data scale and by QGLMM on the observed data-scale in *§Results* (Table 1 and Figure 2). As mentioned in Loeske Kruuk’s comment, $V\_{R}$ on the observed data-scale is indeed calculated on the basis of the additive over-dispersion term calculated at the latent data-scale by the nonlinear model. This cannot be interpreted as a residual variance term *per se*. There is not a unique way to estimate the "residual variance" in a GLMM, and it is true that the estimation of $V\_{R}$ on the observed data-scale calculated by QGLMM is not the classical type. It is important to note that its calculation is mathematically correct. It is also important to note that the sum of variances estimated on the observed data-scale is not expected to be additive and is therefore not expected to sum up to the value of the phenotypic variance that was estimated on the basis of the raw data. As a result, $V\_{P}$ estimated on the observed data scale by QGLMM is expected to be overestimated. As a result, using $V\_{P}$ calculated on the raw data and $V\_{A}$ estimated by QGLMM on the data scale to estimate *h*² would be wrong. The advantage of this method is that intraclass correlation coefficients can be rightfully interpreted as the predictive capacity of random effects on the observed data-scale (*e.g.*, the estimated repeatability is exactly the correlation between successive samplings of similar individuals). Furthermore, the ratio between the various variance components estimated on the observed data-scale by QGLMM is reliable. The most accurate estimate of *h*² that we can calculate for this type of data is therefore expected to result from ratios between variance components estimated by QGLMM on the observed data scale. In Table 1, we show the results on both the latent and the observed data-scale for the sake of clarity and comparison.

**Detailed comments**

1. The title is potentially confusing: I think you mean that ‘self‐recruitment is shaped more by habitat variation than by genetic effects’, but at the moment it could be read as if ‘habitat variation shapes self‐recruitment more than it shapes genetic effects’ (which is obviously nonsense).

> We have modified the title.

2. L50 – drop ‘the weight of’?

> Done.

3. L53 – end this paragraph with a sentence saying that if local recruitment is going to be increasingly important, we need to understand what drives variation in it.

> Done. Furthermore, we now explicitly present the evolutionary inferences that can be made on the basis of this knowledge (population potential for microevolutionary response to selection).

4. L56 – drop ‘draw a’ ? (there’s something odd with the syntax of this sentence)

> Done.

5. L60 – “only a few studies \*have\* quantified…”

> Done.

6. L61 – “and these have been conducted”

> Done.

7. L62 – you reference Table S1 in the text as being about studies of genetic variation in self‐recruitment in wild populations, but then don’t mention the term ‘selfrecruitment’ in Table S1, which is a review of the heritability of fitness.

> We clarified this issue in relation with your comment above. Self-recruitment is indeed a population measure whereas what we measured is what people usually call LRS. We therefore called it so in the text and Table S1 to avoid confusion and allow for comparison.

8. It would be worth referencing Hendry et al. 2018, which also contains a review of estimates of heritability of fitness in the wild (with values from a few more studies):

> Thanks for the reference. The review of Hendry *et al.* (2018) allowed us to update our revised Table S1. Two studies cited in Hendry *et al.* (2018) were missing in our Table S1: (1) Réale and Festa-Bianchet (2000) on bighorn sheep in St Kilda and (2) Wheelwright *et al.* (2014) on savannah sparrow at Kent Island. We noticed that the study of Calsbeek *et al.* (2015) on brown anole lizard in Kidd Cay was missing in the review of Hendry et al. (2018).

You might want to reference Postma 2014 and the review of wild heritability estimates in that as further evidence for the lack of data on wild fish.

> Thanks, we added this reference too in the text. As a result, the total number of publications on the heritability of fitness went from thirteen to fifteen. We now emphasize in the text the lack of data in wild fish in the *§Introduction* (Lines 89 to 101).

9. L62 – “scarcity of quantitative genetic studies in..”

> Done.

10. L64 – hampers/constrains our ability?

> We replaced "challenges" by "constrains".

11. L66 – measurement of total reproductive contribution to the next generation.

> Done.

12. L65 etc – Note that the problem that we may not be measuring all lifetime reproductive success of all individuals born as young in a study population is an issue for any study of any wild population!

> We have clarified in the text throughout the paper that we measured LRS, mentioned the reason behind the scarcity of such data in wild fish population in the *§Introduction* (Lines 89 to 101). This is mainly because migration between populations limits our ability to obtain a pedigree and an exhaustive measurement of LRS. As mentioned by the reviewer, this is the case for any other pedigree based quantitative genetic study of a wild population (*§Discussion* Lines 438 to 439).

13. L68 – say where Kimbe Island is (and also in the Abstract)

> Done

14. L75 – this is a slightly odd conflation of Fisher’s Fundamental Theorem, which states that the additive genetic variance of relative fitness determines the rate of adaptation in the population, and the Robertson‐Price covariance, which states that the additive genetic covariance of a trait with relative fitness determines the rate of change of a trait.

> We have restructured the text of the paper, in particular the first two paragraphs of the *§Introduction* in order to clarify that the “additive genetic variance in relative fitness determines the rate of adaptation of the population” and that this is the question we wanted to address in this revised version of the paper.

15. P72‐97 : this is a bit simplistic, and also completely ignores the review of exactly this issue that you’ve presented in Table S1! There are obviously a range of complex arguments here regarding the expectation and measurement of the relevant parameters: the Hendry AREES review below is a good starting place, but as the authors know, there are very different expectations regarding levels of heritability vs additive genetic variance vs evolvability.

>We have modified the text in the *§Introduction* and clarified how measurements of additive genetic variance, evolvability and heritability bring about different messages and expectations.

16. P92 – this is a confusing jump, as the discussion up to now seems very general and centred on evolutionary biology theory, whereas the sentence suddenly plunges into thinking about different anemone species (how is this relevant?). It would be clearer if you kept the general theoretical arguments in one paragraph, and points about the specific study system in another. I would find this much clearer with some basic information on the study system and species first before any of the methods.

> We have restructured the *§Introduction*, developed general theoretical adaptive evolution arguments in the first paragraphs and moved the points about the specific ecology of the study system and challenges in coral reef studies towards the end of the introduction.

17. L101 – define evolvability

> The term evolvability is now defined in the *§Introduction* (Line 79 and throughout the text).

18. Self recruitment – what about offspring that recruited to other bits of the island?

> Every anemone inhabited by clownfish that was found on the entire reef surrounding Kimbe island was included in the analysis. We have clarified this aspect in the *§Methods* (Lines 139 to 141). We also present how we separated the reef into three geographic areas accordingly with the different coasts (Figure 1, Lines 156 to 159).

19. L113 – so how many “microhabitats” did you have? This seems like crucial information.

> We clarified this in the *§Methods* (Lines 150 to 159) that we used six different habitats in the analysis that were defined by the combination of host anemone species and geographical areas.

20. L116 – how many clownfish (on average) do you get per anemone?

> Three to five - we clarified this in the *§Methods* (Lines 161 to 162).

21. L118 – “allows us to”

> Done.

22. L120 – how do you deal with sex change in the context of a pedigree?

> In the context of pedigree reconstruction, we did not have to include whether the parent was a male or a female. This is the same for hermaphroditic plants. Details on which methods were used for pedigree reconstruction are given in the *§Methods* (Lines 193 to 210). As mentioned earlier, more details can also be found about this in the pedigree paper that we published previously and to which we refer in the text (Salles *et al* 2016, PNAS).

23. Methods: ‐ overall, I found it difficult to get a handle on which individuals were being genotyped and included – what age/size did a fish have to be to be considered a recruit? How old would new recruits be?

> We added text in the *§Methods* (Lines 172 to 190) to clarify the sampling approach, how fish were measured and included. New-recruits were between 10 to 458 days old (unpublished data from otolith).

24. L130 – I know you reference the other studies, but as this paper needs to be stand alone, please give some rough indication of how you assigned parentage.

> We added information in the *§Methods* (Lines 193 to 210) about how we assigned parentage.

25. L 131 – “The pedigree contains only offspring pairs of parents assignments” : why? The Introduction (and L137) said it was 5 generations, but this sounds like only parent‐offspring pairs? (or do you mean something else?)

> Sorry for the confusion, we added information to clarify this misunderstanding in *§Methods* (Lines 193 to 210). The pedigree included 121 families, 987 paternal, 987 maternal, 1809 full-sib, 412 maternal half-sibs, 248 paternal half-sib, 135 maternal grandmothers, 135 maternal grandfathers, 278 paternal grandmothers, 278 paternal grandfathers and 218 cousins. This information is now available in the text.

26. L132: “Because of the presence of overlapping generations in this system, all assignments to single parents were rejected to avoid possible false assignments” I read this sentence several times, but still don’t understand it.

> We modified this sentence to “Assignments to single parents were rejected. We kept assignments of parental pairs. In the context of overlapping generations, we used the year of first sampling and the anemone of each parental couple as information to avoid possible false assignments. As a result, sibship links could not be confused with parental links.” in the *§Methods* (Lines 198 to 201). After settlement, an *A. percula* fish stays the rest of its life in the same anemone, rarely straying more than a few centimeters beyond the periphery of its anemone, which is a helpful property of this system when building the pedigree. We hope that this clarified this point.

27. L134 ‐ “the same individual”

> Done.

28. Fig S1A – how many of the individuals in the pedigree had phenotypes?

> We added the number of individuals in the pedigree which had phenotypes (n=1735) in the Fig. S1 legend.

29. Fig S1B: great to see a power analysis, but it obviously indicates very little power to detect anything other than pretty high heritabilities.

> The power analysis revealed that we had a good probability to detect values above 20%. As a result, it is possible to discuss the fact that we did not detect a value above 20%. As a complementary information, we discussed in the text credible intervals from our main model to evaluate the upper bounds of estimates.

30. L142 : Overall, I found this paragraph on the measurement of individual fitness very confusing, especially with regard to the two‐yearly measures.

[1] How long does a clownfish live for?

[2] If you look at their offspring production every two years, then is that across a whole lifetime, or not?

> We clarified these two points in the *§Methods* and added a more precise description of the clownfish lifecycleto avoid any confusion Lines 161 to 170.

[1] We did not know the exact age of adults because we have no information about when exactly these individuals settled or when they died in between capture events. In a previous paper, we evaluated the relative longevity of Kimbe Island clownfish captured between 2003 and 2013. We found that one fish was at least 10yo, two were at least 8yo, three were at least 6yo, four were at least 4yo, and so on (Salles *et al.* 2016b). We added text Lines 180 to 181 to clarify that individuals were in average sampled 2.88 ± 0.04 times (mean ± SE) over the 6 sampling period (1% of individuals persisting over the 10-yr survey).

[2] Our measurement is the biennual estimate of the Lifetime Reproductive Success. Using the LRS as such would be incomplete because we do not know if breeders sampled in 2003 produced offspring before the monitoring began. We also do not know the future success for fish still alive at the end of our monitoring. I guess this is a similar issue in every long term survey. Jarrod Hadfield suggested to use the biennual estimate to overcome this concern and control for the effect of repeated measures by *V*PE. We had 42.8% of individuals sampled during the 10-year period, we have access to their reproductive success across a whole lifetime (*i.e.*, individuals that settled in Kimbe Island after 2003 and died before 2013). We used biennial measurements so that we could compare reproductive success on a similar timescale, which is like a snapshot of the population reproductive success every two years. This corrects for the issue of individuals surveyed over incomplete lifespan which would insert a directional bias at the beginning and the end of the pedigree. It also protects the comparison of the reproductive success between individuals from possible biases caused by changes in total recruitment some specific years (*e.g.*, producing two offspring when everybody did produce one is best but looks bad if we compare it to an individual from another time who produced 15 on a year when the average was 20).

31. L144 – “recruiting into the breeder population”: is that the entire breeding population around the whole island?

> Yes, indeed. The breeder population refers to all the adults living around the whole island. We have clarified the text throughout the *§Methods*.

32. L145 – do you mean biennial, not biannual (Benoit, in case you’re wondering, Wambam was always inappropriately named; it should have been Wambem).

> We changed "biannual" into "biennial".

33. L144: I don’t really understand the biennial measure: why is this lifetime production?

> Strictly speaking, it is a biennial estimate of the lifetime reproductive success rather than a total "Lifetime Reproductive Success". We are happy to follow your suggestion to call our biennual measurement of the individual reproductive success LRS in order to connect with other studies measuring genetic variance for fitness. We would not like for reader to be confused. We explain that we use a biennual estimate in the *§Methods* and remind the reader about it in the following sections. We believe this is appropriate but are open to suggestions.

34. L147 “Success to Local Recruitment” is an odd term. Why not just Local Recruitment Success? As above, I am still not clear what time‐scale this is over.

> We now use LRS for the Lifetime Reproductive Success as suggested.

35. L148: Delifing needs a reference as soon as it is measured. I am not convinced it is a useful measure of individual fitness, given that it incorporates such a massive component of environmental heterogeneity.

> Done.

36. L152: Delifing has rather dropped out of fashion, maybe in part due to the above concerns. It may therefore not be familiar to people, so you need to explain what it does (not least as the paper is initially set up as talking about production of offspring, but now you are incorporating an aspect of individual survival with delifing). One of the main benefits of it – that it followed a more Gaussian distribution than ‘raw’ measures of breeding success – is no irrelevant given the availability of generalized mixed models. I would therefore strongly suggest dropping it, as I don’t think it adds anything, whereas if you do use it, it needs much more explanation.

> Delifing approach was moved to the *§Supplementary*. We detailed how we estimated the Individual's contribution to biennial population growth rate. To facilitate the understanding, we added an equation and defined all the parameters used for the calculation of Delifing.

37. L158 – explain the “repeated” records, given that above you say the phenotype is an individual’s total contribution of offspring across its lifetime (so repeated measures don’t seem possible?). How many measures did you have on how many individuals of each sex?

> We clarified this aspect in the *§Methods* Lines 215 to 222. As mentioned earlier in our responses, fish that were captured in 2003 might have reproduced earlier, and fish that were captured in 2013 might have reproduced afterwards. I guess this is an issue that affects all wild population survey data sets. To avoid a bias caused by this aspect, and thereby allow for the comparison of the reproductive success between all fish, we used a biennual measurement. As mentioned earlier in our responses, and added in the methods: Individuals were in average sampled 2.88 ± 0.04 times (mean ± SE) over the six surveys, with 1% of individuals persisted over the 10-yr period.

38. L162 – what is the parameter V\_sub(SLR,DL)? This is the usual notation for denoting a covariance between two traits, i.e. Cov\_sub(X,Y). Do you mean the variation in the two measures of local recruitment (V\_SLR and V\_DL)? L165 – the animal model is typically not a generalized linear mixed model, but just a LMM.

> This was clarified in the *§Methods* Line 246: $V\_{LRS}$= $V\_{A\_{}}$+ $V\_{M\_{}}$+ $V\_{NH\_{}}$+ $V\_{RH\_{}}$+ $V\_{PE\_{}}$ + $V\_{R\_{}}$ DL variance decomposition appears in the supplementary. We also clarified that the animal model is typically a LMM (*§Methods*) Lines 239 to 240.

39. L172. So if you do have repeated measures on individuals, this implies at different times of their lives, when you would presumably know if they were male or female? So why not fit sex as a fixed effect?

> This is technically true, we indeed know the sex of each individual at the time of each measurement. However, we only have replicated measures for different sexes in the same individual for a small portion of the population (n=41). We will test for sex effects when we are sure that we can correctly estimate its effects (including cross-sex correlation).

40. L184. Given that variance components are set to be positive, 95% credible intervals will always be positive.

> We have re-written this sentence: “Effects of variance components were considered statistically supported if their posterior distributions do not overlap zero.”

41. L188 – “and we reported”

> We have made the requested revision.

42. L195 – the whole issue of the estimation of variance components on the latent scale and then back transformation onto the data scale is complex, and there are few studies that have embarked on presenting both. This makes this study valuable, and it is very useful that you present both the latent scale and the data scale estimates in Table 1. L203 “HDinterval package within MCMCglmm (ref)”

> Thank you. We have added the reference for the HDinterval package (Meredith & Kruschke 2016).

43. L203 “clownfish live in strong association with their anemone”

> Done.

44. L207 ‐ It would be useful to see the distributions of SLR and DL.

> We have added the distributions of the LRS (Figure 2) and DL (Figure S2).

45. L209 “the habitats where the breeders that contribute more to the local replenishment of the population are located”

> We have made the requested revision.

46. L228 – “linear mixed model”

> Done.

Signed

Loeske Kruuk

**References**

Hendry, A. P., et al. (2018). "The Contemporary Evolution of Fitness." Annual Review of Ecology, Evolution, and Systematics 49(1): 457‐476.

Postma, E. (2014). Four decades of estimating heritabilities in wild vertebrate populations: Improved methods, more data, better estimates? Quantitative Genetics *in* the Wild. A. Charmantier, D. Garant and L. E. B. Kruuk. Oxford, Oxford University Press: 16‐33.

Wolak, M. E., et al. (2018). "Sex‐specific additive genetic variances and correlations for fitness in a song sparrow (Melospiza melodia) population subject to natural immigration and

inbreeding." Evolution 72(10): 2057‐2075.

*Reviewed by Juan Diego Gaitan-Espitia, 2018-10-03 02:40*

The manuscript “Habitat variation of wild clownfish population shapes self-recruitment more than genetic effects” by Océane C. Salles et al explores two interesting questions in evolutionary ecology: 1) What’s the role/contribution of self-recruitment to local population replenishment?; and 2) to what extent additive genetic, maternal and environmental (i.e., habitat) variation contribute to phenotypic variation in components of self-recruitment?. The logic behind this study is that environmental changes (warming, habitat loss/fragmentation), are causing global declines in reef fish populations by affecting larval development and dispersal capacity. The main demographic/ecological consequence of this is (will be) a decrease in the ability of reef fish populations to replenish themselves and therefore the contribution of self-recruitment is expected to increase. By assessing the influence of genetic and environmental components of phenotypic variation on self-recruitment, the authors aimed to reveal the potential of fish populations to respond to selection on this trait. For this, Salles and co-authors used a unique (and amazing!) dataset (pedigree) of 5 generations of the orange clownfish within a 10-year time frame. As expected, the authors found extremely low to negligible additive genetic variation for self-recruitment, and most of the variation in this trait was explained by the habitat (quality, availability, distribution of host-anemones). Overall, the manuscript is very interesting, well designed/written and easy to follow, with a great potential for publication in a good journal. The Ms includes a good theoretical framework, hypothesis and methods are nicely described. Results are fully detailed and discussion is clear.

> Thank you for your comment.

However, before recommending this MS, I invite the authors to address the following comments and clarify some points that are confusing (at least for me). I hope these comments and suggestions will serve to improve the quality and impact of this MS:

Intro (Lines 54-57): In general, I got the idea of the importance of self-recruitment. However, it is still difficult to see the main effect of this trait on Darwinian fitness (at the individual level). The authors explained the existence of self-recruitment at different levels in this section, but did not explain whether this trait shows variation within/among populations, or its connection with fitness. The question in Line 57 is valid but needs to clarify these points before moving in that direction.

> We modified the text throughout the paper and in particular in the *§Introduction* to clarify that the term "self-recruitment" is commonly and widely used in marine biology to refer to the population process by which offspring return to the population of their parents to further reproduce. Natal philopatry can also be used in that sense. We clarified its link to Darwinian fitness throughout the text. The individual contribution to the population to self-recruitment is the reproductive success of local individuals regarding only their offspring that came back to settle in their population of origin. This characterizes the fitness of local individuals in terms of their relative contribution to the replenishment of the local population. In this paper, we study the evolutionary potential of this mechanism by investigating its genetic and environmental variation. We now use the term “Lifetime Reproductive Success” (LRS) as suggested by reviewer Loeske Kruuk to avoid “difficulties if [we] refer to it as one thing in the text, but other studies are calling it something else”. We clarified that this is an individual measurement and that a component of fitness [restricted to the local population].

There is an implicit assumption (reading between lines) across the MS that this trait is an adaptation shaped by past selection. I really can’t see so far the adaptive role and fitness connection of this trait in the current version of the MS. Why should we expect evolutionary adaptive change in self-recruitment? Is it because directional selection is acting on self-recruitment per se? or because correlational selection? Making a clear connection of self-recruitment to fitness will help to visualise this issue (it is assumed along the MS but it is not clear).

> We reorganized the introduction to clarify the direct link between the individual contribution to self-recruitment, which is an estimate of the reproductive success and fitness. Our modifications throughout the *§Introduction* should help to visualize this connection.

Lines 57-58: What do you mean with “at the individual level”? individual species? Populations? Or organisms within populations?.

>Individual level meant organisms within population. We clarified the text in the *§Introduction*, in particular in the first paragraph.

For the general audience it would be good to have a clear distinction between “the ability of populations to replenish themselves” and “self-recruitment”, which in principle sound related. There is a good definition of self-recruitment in Lines 51-52, but it would be easier for the reader to see something similar for the first concept.

> We clarified these notions throughout the *§Introduction*.

Line 72: add “additive” to “gen***e***tic variation”. The presence of other sources of genetic variation doesn’t play the say role in adaptive evolution.

> Done

Lines 76-78/86-87: Why do you expect very low levels of genetic variation and contribution to fitness (Va and maternal) for self-recruitment? I am not talking about a potential cause (78-81) but the reasoning of these expectations.

> In a nutshell, no genetic variation for fitness should be expected in a population at equilibrium and most studies document very small estimates of genetic variance for fitness. We added text in the *§Introduction* (Lines 89 to 93) to state in more detail that low additive genetic variation for LRS is generally found when measured, and in the *§Discussion* (Lines 387 to 398) to discuss our result; the low to negligible evolvability and heritability of LRS. Later paragraphs in *§Discussion* discuss the possible implications and underlying mechanisms.

Lines 78-81: I understand the idea of this sentence but can’t follow its connection with the trait of interest here. What is a “less fit genetic variant” in terms of self-recruitment? Past selection and evolutionary equilibrium… are there other more plausible options?? (e.g., environmental influenced, no real effect on fitness)

> We have removed this formulation “less fit genetic variant” from the text and now discuss the low VA that we found for LRS in the *§Discussion* as mentioned in the previous response. We hope that our text changes will have clarified this point. We also added and reorganized the text to clarify the major role played by environmental effects throughout the paper and in particular in the first paragraph of the *§Introduction* and the first two paragraphs of the *§Discussion*.

Lines 88-89: Not only these factors as the presence/absence of genetic correlations and trade-offs are relevant for evolutionary responses.

Lines 90-91: Phenotypic plasticity can also enhance microevolutionary responses to selection (e.g., genetic assimilation).

> We have simplified this part of the text while improving the logical narrative of the paper. We removed the mention of genetic mechanisms (*e.g.*, genetic correlations) which are not studied in detail in this paper and cited environmental mechanisms (*e.g.*, plasticity, genetic assimilation) that can affect the response to selection in the *§Introduction* (Lines 66 to 71) because we measured strong environmental effects .

Methods: The study seems to be well designed and the statistical analysis seems valid, as well as the mixed-effect approach that the authors used.

> Thanks.

Results: This section is a bit odd. [1] There are parts that seem more relevant for the intro or discussion than for results (e.g., Lines 235-242). [2] In addition, it looks repetitive when the authors report the values in the text but also in tables and figures of the same section. [3] Minor comment here: try to keep uniform “additive/Additive” “maternal/Maternal” (e.g., Lines 230-233).

> We have numbered all of the points in the comment to respond to each one individually.

[1] We moved text to the *§Discussion* and modified the remaining text of the *§Results* to make sure that it all belonged to the *§Results*.

[2] In Figure 2, the distribution of the data and the proportions (size effects) calculated on the basis of estimates on both the latent scale and observed data scale are represented. Only the size effects calculated on the observed data-scale are reported in the text. These are the important values that we want to see stated in the text.

In Table 1, 20 estimates are presented. Only three of these are presented in the text. These are the three important values that we want to see stated in the text at the same time than the process they relate to (adaptation rate and breeder’s equation).

[3] We have made the requested revision and kept capital letters only when used to explain the correspondence with the A of *V*A and the M of *V*M.

Discussion: Lines 291-292: Assuming that selection is in fact acting on this trait. Moreover, here you are reporting results.

> This aspect is now clarified in the revised manuscript. As mentioned in the introduction and discussed in other places. LRS is a proxy for fitness. As a trait of a given individual fish, it estimates its number of offspring that settled in their parental population. It is therefore highly related to fitness and selection. Values that were relevant to the *§Results* section were removed from the text of the discussion. Some values are still presented as percentages in the *§Discussion* to be precise and give the reader the possibility to take the measure of the size effect.

Lines 316-326: migration was indeed a big component of population renewal (44%). However, selection against migrants may also be involved in the low Va detected, especially when migrants are from very contrasting habitats (See Nosil et al., 2005; Hendry et al., 2002).

Moreover, it is true that gene flow is expected to increase genetic variation. However, this is mainly neutral genetic variation which is not always linked to additive genetic variation, and therefore it is selectively neutral and tells us nothing about the adaptive or evolutionary potential of a population or a species (Holderegger et al., 2006).

> We totally agree with the first part of the comment. In fact, we already mentioned it in the text by stating that “selection for self-recruitment has to be strong to keep the population at evolutionary equilibrium”. We added the following test: “thereby against migrants” in the middle of this sentence (*§Discussion* Line 410) to make sure that this aspect is crystal clear to the readers.

> However, we do not agree with the second part. It is not true that gene flow is not expected to increase additive genetic variation and the adaptation rate. In fact, the expectation is quite the contrary and there is empirical evidence for it (*e.g.*, Lavergne & Molofsky 2007, Facon *et al* 2008).

Line 327, 329-: But you showed here that maternal effects do not contribute to the genetic variance of self-recruitment.

> We removed “(*e.g.*, maternal effects)” from the former version of the *§Discussion* Line 327. About the following paragraph about maternal effects (*§Discussion* Lines 420 to 434): We indeed found that the maternal variance component was low to negligible (explaining up to 2.7% of LRS). In this paragraph we discuss the possibility that this variance component reflects the maternal genetic effect whereas the effect of the habitat of birth might reflect the maternal environmental effect.

In some parts of the text, there are jumps in the focus to climate change instead of discussing to a greater extend the low Va of this trait, the existence of selection (type, direction, etc), and the potential for genetic correlations influencing the adaptive landscape.

> We have restructured the *§Introduction* to clarify the logical narrative of the paper which is now more linear. As a result, the focus is now on $V\_{A}$ and adaptive potential because this is a more fundamental question of evolutionary ecology. Climate change comes later when we discuss questions that are more specific to the study system. There should be no more jumps there. We also have clarified the structure of the *§Discussion*. We discuss the “susceptibility to habitat change”, the “low to negligible evolutionary potential” where we present the maximum microevolutionary increase in LRS in response to selection that could be expected (10 generations to increase the average reproductive success of the population by 1 progeny), the “connectivity as a plausible cause”. However, we do not feel comfortable discussing other selection pressures when no clear evidence is available or gradients of selection (fitness-trait correlation) quantified. We discuss additive genetic variance in the context of the rate of adaptive evolution but we do not feel comfortable discussing the role of genetic correlations influencing the adaptive landscape in the absence of a proper adaptive landscape approach (which we recognize would be an interesting topic to investigate) and in the absence of other traits that we could estimate genetic correlations for.

The single population focus is also a limitation. Trying to move your conclusions to a broader scale may help to understand the evolutionary dynamics of self-recruitment in your study system.

> We agree that this is limiting. We clarified in the revised version of our manuscript that our findings can be discussed at the broad scale of studies conducted on other species that investigated the genetic variation of LRS. There are only 15 studies available that tackled the same question than us, and none in a marine population (*§Discussion* Lines 436 to 450).

Population surveys at the individual level require data acquisition, molecular and analytical work that is already very challenging in one population, in particular for marine fish. The logistical challenge associated with conducting this type of work in multiple populations will probably be overcome in future years after more results have been obtained in single populations (15 to date) that have allowed to advance our knowledge on the rate of adaptive evolution in wild populations.

**References:**

Hendry, A.P., Taylor, E.B. & McPhail, J.D. 2002. Adaptive divergence and the balance between selection and gene flow: lake and stream stickleback in the Misty system. Evolution 56: 1199–1216.

Holderegger, R., Kamm, U., & Gugerli, F. (2006). Adaptive vs. neutral genetic diversity: implications for landscape genetics. Landscape Ecology, 21(6), 797-807.

Nosil, P., Vines, T. H., & Funk, D. J. (2005). Reproductive isolation caused by natural selection against immigrants from divergent habitats. Evolution, 59(4), 705-719.

**References for our response to reviewers comments.**

Calsbeek, R. *et al.* (2015). *J. Evol. Biol.*, 28, 1975-1985

Facon, B. *et al.* (2008). *Curr. Biol*., 18, 363-367.

Hendry, A. P. *et al.* (2018). *Annu. Rev. Ecol. Evol., Syst.* 49(1): 457-476.

Lavergne, S. & Molofsky, J. (2007). *Proc. Natl. Acad. Sci. USA*, 104, 3883-3888.

Meredith, M., & J. Kruschke (2016) HDInterval: highest (posterior) density intervals. R package version 0.1.3. [https://CRAN.R-project.org/package=HDInterval](https://CRAN.R-project.org/package%3DHDInterval).

Réale, D. & Festa-Bianchet, M. (2000). *Heredity, 85, 593-603.*

Salles, O.C. *et al.* (2016a). *Proc. Natl. Acad. Sci. USA*, 46, 13245-13250.

Wheelwright, N. T. (2014). *Evolution*, 68, 3325-3336.