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.

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. The Methods regarding the pedigree construction are extremely confusing.

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.

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). 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?

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.

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). 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 data-scale 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).

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).

2. L50 – drop 'the weight of'?

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.

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

5. L60 – "only a few studies *have* quantified…"

6. L61 – "and these have been conducted"

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 'self-recruitment' in Table S1, which is a review of the heritability of fitness.

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): 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.

9. L62 – "scarcity of quantitative genetic studies in.."

10. L64 – hampers/constrains our ability?

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

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!

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

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.

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.

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.

17. L101 – define evolvability

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

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

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

21. L118 – "allows us to"

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

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?

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.

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?)

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.

27. L134 - "the same individual"

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

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

30. L142 : Overall, I found this paragraph on the measurement of individual fitness very confusing, especially with regard to the two-yearly measures. How long does a clownfish live for? If you look at their offspring production every two years, then is that across a whole lifetime, or not?

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

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

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

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.

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

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.

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?

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.

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?

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

41. L188 – "and we reported"

42. L195 – the whole issue of the estimatpin 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)"

43. L203 "clownfish live in strong association with their anemone"

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

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

46. L228 – "linear mixed model"

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.