Thank you for your submission to PeerJ. I am writing to inform you that in my opinion as the Academic Editor for your article, your manuscript "Rape and the Prevalence of Hybrids in Broadly Sympatric Species: a Case Study using Albatrosses" (#2014:02:1478:1:0:REVIEW) requires a number of major revisions before we could accept it for publication.

My comments are as follows:

Editor's comments

Thank you for your revision. First I should apologize for the long time it has taken to make a decision. Given your disagreement with several of the more substantive reviewer comments, I felt it important to evaluate each point carefully and finding time to do that took longer than I anticipated. I find merit in some of your arguments, but also find more merit in the reviews than you perhaps give them credit for. At this time, I do not believe the paper is suitable for publication, but I do think it could be quite easily and invite another round of revision should you be willing to address my comments, as described below. My comments follow the order of your response letter.

1. In your response you state that your fundamental goal is only to seek “the beginnings of an answer to the question of why hybrids vary so much in frequency between broadly sympatric”. Nonetheless, language in the discussion implies (to me, at least) that you are doing more than this – and in my view go beyond the evidence (e.g., the statement “heterospecific rapes are likely responsible for generating F1 hybrids” in ducks, among several examples). I suspect that this is the basis for the reviewer concerns. Recasting these claims more clearly as untested propositions could easily resolve the differences.

I think we have toned down the language appropriately.

Equally, stating the “fundamental point of the paper” more clearly in the introduction (i.e., along the lines that it is stated in the response letter) might help.

I have revised the introduction to more clearly develop the hypothesis.

2. I agree with you that the exact test does support a difference. But a small sample size still warrants great caution when it comes to interpretation. Type I errors remain a concern, especially when just one more individual could change the inference. Indeed, this issue is one that has received substantial attention in the recent literature on the difficulty that scientists have replicating statistically significant results.

We take your point and have modified the text toward caution in interpreting results.

3. In response to the reviewers’ concerns that your data do not test your predictions, you argue that, although you cannot prove that heterospecific rape is the cause of unidirectional hybridization, to meet the goals of this paper it is sufficient that your data are consistent with that hypothesis. This is a fair point, as long as it is acknowledged that the test is not an especially strong one. The problem I see – and the one that I suspect resulted in the reviewers’ concerns, is that the paper goes beyond suggesting that the data are consistent with the hypothesis. E.g., the abstract states “We show heterospecific rape to be the probable source of hybrids …”, the start of the conclusions also suggests that hybrid albatross “result from heterospecific rape”, etc. Yet, there is no evidence presented to address alternative hypotheses, such as the possibility that a small number of female albatrosses choose heterospecific pairings (e.g., perhaps because an egg got laid in the wrong nest and the resulting chick was raised by parents of the other species, which it then imprinted on – I’m not saying this is a likely scenario, just that alternatives of this type are not considered). For the paper to be accepted, the manuscript text needs to be consistent in tone/caution with the arguments made in the response letter.

I have attempted to further qualify this statement in the abstract.

4. I have gone back and forth over the issue of whether “rape” is an appropriate term to use. Like the reviewers, my inclination is to err away from it because of the difficulty of inferring motivations and because of the term’s human connotations. After reviewing various definitions, however, I’m not sure I can justify my instinctual stance. Moreover, “forced copulation” (which is what I too would have suggested as an alternative) suffers from the same problems concerning motivations. So, I leave this to your discretion as I do not feel that current conventions in the literature are sufficient reason for me to object.

Thanks!

5. I do not know enough about IM analysis to judge whether your sample sizes are adequate. I would, however, say that the fact that others have published based on similar sample sizes is not a good enough justification – plenty of work gets published in which the statistics are questionable. Are there theoretical grounds for considering your sample adequate? A response on this issue that directly addresses the paper cited by the reviewer would be helpful.

The paper cited is the standard documentation distributed with the IMa program. It does not say anything about the necessary sample size needed to run the IMa analysis adequately. We are unsure what the reviewer is referring to. However, we can refer to the simulation study by Choi and Hey (Hey is the creator of the program IMa2) that shows the distance from the true assignment using 2 populations and varying number of loci (1-10). In this simulation study, IMa began to recover the true population assignment with 7 or more loci and 20 individuals per population. Our study includes 10 loci and >20 individuals per population.

6. Perhaps my biggest concern about the results relates to the model comparisons. Table 3 only presents a small subset of the models compared and without seeing the entire model set it is difficult to make inferences about which variables are more important than others. Moreover, the 6 models shown all have ΔAICs <2 and low Akaike weights, suggesting that they are functionally equivalent in their ability to explain the data and that none of these models really stands out over the others. Given that the models shown appear to include both models with unidirectional and bidirectional gene flow (if I am reading the first column correctly), the statement that “Our results show that models with unidirectional gene flow are consistently ranked the highest” is not very well supported (i.e., within this group there is no basis for concluding that the 5th model is any better than the 1st).

The basis of this conclusion is based upon how AIC penalizes extra parameters:

AIC = 2k – 2 log(P). We have explained this further in the revised manuscript and cite the model inference book by Burham and Anderson 2002.

This conclusion is consistent with the inferences one would make from the overlapping confidence intervals. Moreover, if the next X models are only marginally worse than these 6 (say, all have ΔAIC <4-5, then the story could be even murkier). Although ΔAIC < 2 is a common cut-off, it is a debatable one and it is helpful to know more about the full set of models. To feel confident about the conclusion here I would want to see, at least, all models with delta AIC < 10.

We have presented the full set of models in the revised manuscript.

Model-averaging across all models considered (not just the subset shown) might also provide a clearer signal, and is an appropriate and standard extension of the information theoretic multi-model approach.

We show weighted AIC scores across all models in the revised manuscript.

7. This comment from the reviewers was not addressed in the response and I do not see text changes that address it. Please clarify if I am missing something:

• A summary of PCR conditions should be included (at minimum, magnesium concentrations and annealing temperatures for each primer).

The PCR conditions are included in the Methods of the revised manuscript.

8. The legend for Tables 1 has been used for Table 2 as well.

Finally, I have made a number of additional comments (many minor editorial points) directly on the manuscript, using track changes to facilitate your ability to examine them. I will send this file under separate cover as PeerJ has no mechanism for me to attach a document to this response letter.

These changes were all accepted or made with minor modifications of wording. In my files the table legends were correct, so i presumably made some error in uploading them.

Despite my lengthy comments I continue to find this paper to be an interesting contribution and hope that you will feel comfortable making the suggested changes. Should you decide to submit a further revision, please address each of the comments in this letter, and those on the manuscript in your response.

Please be aware that we consider these revisions to be major, and your revised manuscript will probably have to be re-reviewed.

If you are willing to undertake these changes, please submit your revised manuscript (with any rebuttal information\*) to the journal within 60 days.

\* A rebuttal letter and any tracked changes can be added to the file uploads page under the "Revision response files" section. Please also upload a clean untracked version for production purposes to the primary files section and replace your previous manuscript. Accepted formats for the rebuttal letter and tracked changes are: docx (preferred), doc, and PDF.

Chris Elphick

Academic Editor for PeerJ