

Decision for round #2: *Revision needed*

Minor revision

This preprint is suitable for a recommendation but the second reviewer makes a number of suggestions which I would like to give the authors a chance to respond to. In particular further attempt to make the discussion more general and more concise would be an improvement.

Dear recommenders,

We are grateful for this second review of our work. We carefully addressed their new comments. As for the previous revision, all comments made by each reviewer have been put in bold, while our answers are not.

by [Chris Jiggins](#), 19 Nov 2021 13:03

Manuscript: <https://doi.org/10.1101/590836>
version Version 2

Review by [Simon Henry Martin](#), 09 Nov 2021 11:47

The authors have gone to great lengths to address the reviewer comments, including performing additional cloning, reanalysis of all the data, and re-writing most of the manuscript. My main concerns with the previous version were the exclusion of singletons and the interpretations of distinct origins in the Discussion. Both of these have been addressed: singletons are now included, and the Discussion provides a very thoughtful and cautious interpretation of results, with a call for better sampling and the use of genome-wide data in the future.

Although this study is an exploration of one unusual genus, it has broader relevance in that it reveals the complex ancestries that can arise through atypical reproductive modes, and the challenges in species delimitation that can result. I therefore think the manuscript will be relevant to many biologists, and that it is suitable for publication.

We would like to thank this reviewer for his thorough revision during the previous round, and for his appreciation of the additional analyses performed and the approbation of the revised manuscript.

Review by [Arnaud Estoup](#), 04 Nov 2021 10:57

I found that the authors did a very nice job to answer and take into account most of the 4 reviewer's (and the recommender) requests, including my own requests. As a result, their paper is much better now. I sincerely congratulate the authors for their efforts and I look forward to see their paper recommended in PCI Evo Biol. I however have four (last) comments that might be useful to further improve the ms.

1/ Title: The new title is fine but key notions/words allowing a better understanding of the content of the ms (and making it more appealing to potential readers) are lacking in it.

- “reproduction system” = androgenetic + sexual as the populations studied include both reproduction systems

- Cross-species mixing (even if the concept of species is somewhat “dubious” here.

- Invasive (native) species/populations

- “Extensive” mixing...not sure you did really measure the intensity of mixing events so that the term might be exaggerated: may be “substantial “(or frequent?) which is more vague but still indicates that it appends relatively frequently is a better term.

So, a possible title might be something like:

“Substantial cross –species mixing among androgenetic and sexual lineages in the native and invasive clams of the genus Corbiculata”

Of course, this is just a suggestion that can be twisted and adapted.

We thank this reviewer for his many constructive comments during this reviewing process which improved our manuscript tremendously, we also considered this comment on the title. We agree the term “Substantial” may be more adequate, and we indeed refer now to the two reproductive systems (sexual and androgenetic) in the new title. However, we wish to keep a rather concise title and therefore did not include the notions of “invasive and native lineages”, but this is developed in the abstract. We avoid the use of species in this genus complex, therefore also in the title.

2/ Echoing Point 2) of the recommender: I recognize that the authors made an effort to fulfill the recommender request (which is important for me too)) but I have the feeling that the authors should make an additional step in this direction = please state/formalize more thoroughly the hypotheses or predictions being tested (in the last section of the introduction). I reckon that this is not necessarily easy to do owing to the relatively descriptive nature of the ms. But I am convinced that this is possible. Start a section such as: We more specifically address the following questions” or “Owing to the reconstruction method of the relationships between the populations and species studied we predict that:” For instance: if little or no mixture occurs then we should observed ...”this type of network”. If mixture occurs frequently then we should observed...”this type of network”. This is just an example. You also could make (similar) predictions on the other key results/focal points of the ms.

The end of the introduction was modified (from line 189) following the recommendations of this reviewer with the hypotheses of this study being developed (lines 205 – 216).

3/ Result sections: the sections remain too descriptive. Of course you need to describe (finely) your results and I know how laborious it can be to describe phylogeographic data/results. But I do not see any clear general “take-home message” in such sections that you could put at the beginning or at the end of the sections. Such “take-home message” would allow general readers to find their grail in the ms without having to read all very focal descriptive points. Actually the reader has to wait the discussion section to find such take home messages (which by the way necessitate a kind of redundant redaction of the results with recurrent mention to a large number of figures (and table) to reach the

message). This is intriguing as one does not expect to see many citations of figure or table illustrations in the discussion (as the job has been done in the result sections). The problem might be the general structure of the ms = it might be optimal in term of redaction (especially for the results and the discussion) to follow a plan resting on the hypothesis or predictions or questions that you addressed upstream in the ms (end of introduction) than a plan based on figure descriptions based on methods. A symptomatic fact of this is that the results include only two sections (one –very - small section and one huge section).

We modified the subtitles of the Results and the Discussion sections to provide the take-home message in each title. Reading all the sub-titles provides the main results of this article. In addition, a take-home sentence or paragraph is given at the beginning or the end of each sub section in the results. We also re-structured the results to include more sub-sections to facilitate the understanding. We thank the reviewer for this pertinent comment.

The Discussion section follows the same plan, with the answer to the main questions addressed in the introduction being emphasized at the beginning of the Discussion (lines 546-548 and 567-568), then following subsections including more details for specialized readers.

4/ I wonder to which extent the fact that many population/species samples are obviously missing (i.e. were not sampled or analyzed) in this fascinating complex story might change the interpretations of the deduced evolutionary relationships between populations/species. This might be particularly the case for the recent invasive population/species history for which a number of population/species locations are missing in the invaded range. According to my own experience, missing or “ghost” populations in a dataset are often a problem in deciphering safely the introduction history of an invasive species. I know that it is impossible to get and analyze all samples in such a system but I would be happy to see the authors adding a comment about this in the discussion section.

May be that (more generally) a section in the Discussion devoted to “Limiting factors of our study” would allow to point and discuss the main limiting factors that weakens the interpretation of the data. This section would include the point mentioned just above + various other points such as the fact that the reproduction system is unknown in several studied population/species. This might help structuring the discussion section.

In this study, the limiting factors are all discussed within their specific context rather than within a general section where they would all be listed and explained. The consequences of ghost populations are discussed on lines 589-591, and the need to increase the sampling is discussed on lines 573-575 (South-Est Asia), on lines 591-593 (Africa) and on lines 632-634 (invasive range). The need to clarify the reproductive system of several populations is discussed on lines 591-593 (Africa), on lines 608-615 (*C. sandai*) and we also emphasize the use of genome-wide markers on lines 608-615 and 632-634.

5/ Very minor point: L502: “Moreover” might be better that “However”

This was modified. Thank you.

Review by anonymous reviewer, 07 Nov 2021 17:34

The extensive revision has successfully addressed the majority of the points raised during the previous round of reviews. Especially the introduction is now aimed at a much more

general readership, and most of the technical issues regarding data and analyses have been clarified. A few points might need further attention in an additional minor revision.

First, the distinction between mitochondrial and nuclear markers could be taken even further. In particular the Circos plot (Fig. 4) apparently mixes nuclear and mitochondrial markers and therefore two very distinct genetic mechanisms: Sharing of mitochondrial alleles between lineages can be explained by androgenetic males using eggs from distinct lineages, whereas sharing of nuclear alleles is evidence for hybridization/nuclear capture. Perhaps consider presenting two circos plots, one for the mitochondrial and the other for the nuclear alleles, so that the two could directly be compared. Of course, hybridization/nuclear capture could also lead to sharing of mitochondrial alleles, so that the comparison of the figures does not directly present a comparison of the two mechanisms. This would need to be explained in more detail.

The aim of the Circos plot is to present the introgressions that occurred between distinct lineages. These introgressions can either represent cytonuclear mismatches (the association of the nuclear genome of one lineage with the mitochondrial genome of another lineage) or hybridizations (the association of the nuclear genome of one lineage with the mitochondrial and nuclear genomes of another lineage). As these two processes have similar consequences in terms of introgression, they are considered together in the Circos plot. Mitochondrial and nuclear markers are separated in the other analyses that aim to define allelic pools.

Second, the discussion is still rather system-specific. It is easy to become lost in the many lineages and comparisons discussed in great detail (and also with some repetition). I think the attractiveness of the preprint to a general readership could be improved by placing stronger emphasis on the major points and by adding a general conclusion that circles back to the main questions raised in the introduction.

Following a comment of Arnaud Estoup (reviewer 2), the end of the Introduction and the Results sections were modified to emphasize the take-home messages of the manuscript. Notably, the main conclusions of the study, referring to the description of the main aims (lines 203-205) are now summarized in the titles of the sub-sections of the Results (lines 363, 374-375, 437 and 467) and Discussion (lines 532, 616 and 667). We hope that the hypotheses put forward in the Introduction (lines 205-216) and the sub-titles in the Results and Discussion will help the general reader to get the major points.

A few additional minor comments:

- **Initially, the description of androgenesis (L. 59-60) sounds as if a reductional meiosis takes place during spermatogenesis.**

We clarified this by mentioning the “transmission of unreduced genomes” (line 63).

- **The statement that males or hermaphrodites “hijack” eggs of other individuals (L.64) is at odds with the statement that all androgens are hermaphrodites and that they can use their own eggs. Also, given the modified meiosis during oogenesis, which leads to eggs without any chromosomes, I do not completely follow the notion of “parasitism” (at least not during androgenetic “self-fertilization”): These eggs could not be used for any other form of reproduction, so the “parasite” appears to only confer benefit rather than**

harm to the “host”. Is there an added value to the manuscript of raising the issue of egg parasitism?

This notion of “egg parasitism” was redefined in the paper (line 70-72). This is an important aspect of androgenesis and we hope we have clarified it.

- **On L. 85-87 it is unclear how the orientation of the meiotic axis can lead to the formation of two polar bodies in a single meiotic division. Either explain better or discard (the argument doesn't appear to be crucial here).**

This information is represented in Figure 1, so it is considered important. Precisions were added (line 91-96 in the new version) and we hope it is clearer now.

- **There appears to be a contradiction between the first two sentences of the paragraph starting on L 117 (the first sentence says that sexuals occur in Africa, Asia, Australia, and Middle East, the second that they are geographically very restricted).**

The first sentence was split into two sentences for clarification and the previous assumption was mitigated (currently lines 125-127).

- **L. 154: unclear if “self-fertilize” refers to androgenesis within hermaphrodites or to sexual self-fertilization.**

This was referring to hermaphrodites, with self-fertilization through androgenesis. This was clarified (line 154-157), thank you.

- **L. 433: “Fig 4” (instead of “Fig. 3”).**

Thank you, we were indeed referring to Figure 4 (currently line 458).

- **Perhaps it is worth mentioning that nuclear capture/hybridization events are probably much more easily detected if they occur between distant lineages than within lineage. (The latter could be severely under-detected or go un-noticed altogether).**

This is a good point. More generally, this is also the case of androgenesis (mentioned line 73), but we added a specific focus on nuclear captures at the end of the manuscript (lines 718-720).