Review by Arnaud Estoup of the revised/resubmitted PCI EvolBiol ms by Vastrade et al (round 2)

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 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.

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 convince 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.

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).

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.

5/ Very minor point: L502: "Moreover" might be better that "However".