Review for the PCI Manuscript "Cryptic species and hybridization in corals: challenges and opportunities for conservation and restoration" by C Riginos *et al.*

As this manuscript is a review article, some of the questions in the PCI checklist don't apply. Here below are my answers to ones I deemed relevant:

- **Title and abstract**
  - Does the title clearly reflect the content of the article? Yes
  - Does the abstract present the main findings of the study? Yes
- **Introduction**
  - Are the research questions/hypotheses/predictions clearly presented? Yes
  - Does the introduction build on relevant research in the field? Yes
- **Materials and methods**
  - Are the methods and analyses sufficiently detailed to allow replication by other researchers? Yes
  - Are the methods and statistical analyses appropriate and well described? Yes and No see below
- **Discussion**
  - Have the authors appropriately emphasized the strengths and limitations of their study/theory/methods/argument? Yes
  - Are the conclusions adequately supported by the results (without overstating the implications of the findings)? Yes

General comments

The manuscript is really well written and addresses an important issue in coral reef studies. Through its analysis of the literature, it also stresses a point I strongly back, that corals are very nice models for the study of speciation processes.

However, the read would have been totally enjoyable and timely if another review on precisely the same topic had not been published very recently ("Integrating cryptic diversity into coral evolution, symbiosis and conservation" by Grupstra *et al.* published online in February 2024 in Nature Ecology & Evolution https://www.nature.com/articles/s41559-023-02319-y). Overall, the manuscript under evaluation shares more than a quarter of its references with the published review, and even if the authors of the present manuscript stress that they focused their analyses on NGS based studies, this proportion of shared studies rises to more than 40% for coral related citations. A good point for the present manuscript is that it could profit from the more recent publications that per force were not analyzed in the already published review (a strong signal of the current high activity in the domain).

Unfortunately, this leads to very similar discussions and conclusions in the two articles on several topics, leaving involuntarily the somewhat false impression that NGS based approaches

did not bring about new insights in coral crypticity. But, again, this feeling arises when reading the two articles in a row and is not due to any fault of the authors of the present study. They just had the bad luck of publishing their analysis in second…

On the other hand, the present manuscript develops more issue around speciation processes, and goes back on previous thermotolerance literature in the light of the high occurrence of cryptic species, two original and valuable points.

The Nature Ecol. Evol. Review being already published, I would recommend to the authors, with the objective to gain in added value for their manuscript, to slightly amend their text in stressing more strongly the conclusions that are original to their study (especially their focus on genetic data analysis), and in reducing the parts that are now redundant to the published review (see more detailed comments below).

Line by Line comments:

Lines 215-217: As stated above (and on line 190), this focus on NGS based approaches is original to this manuscript. I would therefore recommend that conclusions based on earlier techniques be reduced if not skipped, as they are developed in the published review.

Line 261: the statement that PCA and hierarchical genetic clustering methods are “unsupervised machine learning methods” seems to me a bit bold, and would have at least needed more justification…

Lines 296-314: If I understood correctly, these lines describe the way the literature survey was conducted. They therefore arrive late in the manuscript, in the sense that this part is included in section 2 whose title is already a conclusion of the survey. This part would have gained to be presented in its own section, as it describes the data mining protocol.

Line 299: The authors restricted their review to publications using Structure and ADMIXTURE as model-based clustering analysis. I wonder why they excluded snmf analyses (Frichot & François 2015) from their choice, especially when focusing on NGS based approaches (considering that snmf clustering is particularly adapted to large datasets).

Line 312: The authors also restricted their survey of model-based clustering analyses to outcomes when individuals were assigned to only two groups (K=2). Isn’t this threshold too stringent, as the number of cryptic species found in a given dataset is strongly dependent on the sampling scheme (i.e. it can be relatively easy in some places to sample more than two cryptic species even in few individuals)?

Lines 357-408: This whole section, through based on recent publications, is quite redundant in its conclusions to the corresponding section in Grupstra *et al.* (which also included some of these publications). It could maybe therefore be reduced to develop more its last part (lines 402-408) that stress the need to go beyond single marker approaches for the characterization of the symbiont population.

Lines 384-387 A verb seems to be missing.

Lines 410-478: The discussion on depth structuration, mostly based on less recent papers, is also well developed in Grupstra *et al.* and would gain by differentiating itself from the published review. In that prospect, the Figure 4 is quite interesting and welcome.

Lines 524-540: This whole section is totally redundant to its counterpart in Grupstra *et al*.

Lines 542-547: This section, and Textbox 2, on the other hand, are totally original and definitely interesting (through disturbing!).

Lines 578-615: Same comment, these sections are original and interesting and would therefore gain in being more developed.

Lines 762-804 Textbox 1 relies heavily on the conclusions of Pante *et al.* 2015 (and not 2014, the online publication date), cited elsewhere but not here, and Carstens *et al*. 2013, not cited at all)