Recommender's decision summary #1 *September 21, 2023*

Dear Dr. Riadi and Dr. Opazo,

Two reviewers have evaluated your preprint 'Evolution of ion channels in cetaceans a natural experiment in the tree of life' and have identified several issues that should be addressed. Please note that experiments (as suggested by Reviewer 2) are not necessary for your preprint to be considered for recommendation. Instead, I strongly suggest that the amount of speculation in the relevant section be reduced, as described below.

I have also evaluated the preprint myself and found it an interesting read. However, I have several concerns that should be addressed in your next revised draft. My key concern is that you have not convincingly shown that ion channel genes in particular have evolved differently in cetaceans vs. non-cetaceans. For instance, there are fewer genes in general in cetacean genomes, and it seems likely that overall cetacean genomes experience higher rates of gene turnover, rather than ion channel genes in particular. Similarly, it is unclear in the current draft whether ion channel genes display enriched signals of positive selection relative to other gene categories (and in cetaceans specifically compared to non-cetaceans). This would need to be shown to support the author's conclusions.

Additional analyses are needed to address these issues. I suggest that the authors reformat their manuscript to focus less on the speculative results (e.g., related to human-pathological variants and the $Na_v1.5/7$ sodium channels) and instead provide more convincing evidence for the claims for which they can be more confident.

Please respond to my specific points listed below in addition to the reviewers' comments and let me know if you have any questions.

Sincerely,

Gavin Douglas

Main comments

- I agree with Reviewer 1's point that the authors have not provided clear motivation for why they have chosen to study ion channels. Further details should be given to explain (and justify) the claim that physiological axes have diverged in cetaceans.

- Please provide justification for why so few non-cetaceans were included (given that there are many more mammals that could have been included). I think it would be much more convincing that cetaceans were major outliers genomically in mammals if a more diverse set was compared to. This is highly related to Reviewer 1's point 11.

- You report that there are significantly fewer ion channel genes in cetaceans vs. non-cetaceans (which is true, based on a t-test at least), but actually there appears to be a stronger signal of fewer protein-coding genes in general in cetaceans vs. non-cetaceans, which seems more relevant. The actual percentage of ion channels does not appear lower (and could actually be higher), as displayed in Figure 1. This suggests that many gene groupings are likely at lower absolute copy numbers, and not ion channels specifically. Is this true (e.g., if you look at genes grouped based on different protein

domains)? Or are ion channels specifically depleted? The current description of the results would be misleading unless the latter is true.

- In addition, I think showing the distributions of the gene counts of all protein-coding genes and for ion channels in the two lineages separately would help readers pick up that the overall numbers of genes are lower (although using a phylogeny-informed statistical test, as suggested by Reviewer 1, would be good).

- Starting at L242 you describe what functional categories genes displaying signals of positive selection are enriched for. A summary of the signals of positive selection identified is first needed. For instance, how many genes out of how many tested were significant? What were the effect sizes? How did the results differ based on the two sets of models compared? It would also be good to remind the reader what the general analysis was (e.g., that it was restricted to ion channel genes).

- The authors claim that their positive selection results "emphasize the importance of ion channel genes in adapting to diving" (L261-262), but this is not convincing as they only scanned for signatures of positive selection in ion channel genes. It is very possible that many kinds of genes show evidence of positive selection, and that ion channel genes are not especially enriched for this signal compared to the entire genome. The authors would need (1) to compare to other gene categories to convincingly show that ion channel genes in particular display evidence for positive selection, and (2) show that this signal is restricted to cetaceans rather than mammals in general.
    Related points:
    o My point #2 above is very similar to Reviewer 1's point 12: if the authors are making claims about higher levels of positive selection specifically in cetaceans, then this must be relative to other mammals, but based on the methods I do not believe that non-cetaceans were tested for signatures of positive selection.
    o What was the background set of genes used for the Enrichr analysis? Based on the methods it sounds like only ion channel genes were tested for positive selection.
    o Table 1 only a small number of unique genes are listed as associated, so this suggests that the signal of positive selection could be restricted to just a few ion channel genes (if there are similarly small gene sets for the other phenotypes discussed in addition to heart physiology).

- Clarification is needed that many of your results are bioinformatic predictions of phenotypes rather than actual observations of phenotypic differences. Two (non-exhaustive) examples are listed below.
    o L55 – 'seems to be sensitive to TTX' implies observed sensitivity (or at least could be interpreted that way). This sentence should be re-worded to clarify that you *predict* sensitivity.
    o L262 – 'adapting to diving' should be clarified that this is only a possible link. This is a hypothesis and that there is no direct evidence of a link between the elevated dN/dS in some genes and adaptation to diving.

- I agree with both reviewers that the section describing the scanning of human-pathogenic variants in cetacean ion channel genes is overly speculative. This information could be useful but there is very little evidence to support the speculations. Human-pathogenic variants could have entirely different effects in the different genetic and environmental background of cetaceans, so I do not think much can be concluded from this analysis. In addition, for at least some of the analyses (e.g., Table 2) the authors do not provide information on the distribution of these variants in non-cetaceans beyond humans, meaning that it is possible that the non-human-pathological state is recently derived, and the pathological state in humans is ancestral.

Formatting comments

- I could not find the link to your Zenodo repository within your preprint itself (although I could find it through the PCI Genomics portal). Please make sure this is included in a separate section titled "Data, script, code, and supplementary information availability".

- I appreciate that you provided a Word document accompanying your scripts that describes your exact bioinformatic steps, but I strongly suggest this be changed to plain text format as sometimes special characters can cause issues when commands are copied and pasted from Word.

- Please move your funding information from the acknowledgements to a separate subsection called "Funding".

- Also include a separate section called "Conflict of Interest disclosure" indicating any conflicts or confirming that you have none.

- Please use a consistent reference style in your reference section. Note that many journals, including Peer Community Journal, request that DOIs be included in the reference list.

Minor comments

- L53: Should be "a signal of positive selection", not "the signal".

- L56: Please write TTX out as tetrodotoxin in the abstract (and likely this would be better written out in the keywords as well).

- L78: I suggest 'Thus,' be removed (or replaced with 'Indeed,'), as this sentence is not a necessary consequence of the preceding point.

- L82: Should re-word 'translates their solution to us' to be something like ', from which we can potentially gain medical insights'.

- L107: I would re-word 'ion channels have been estimated' to 'putative ion channels have been identified'.

- L147: Capitalize 'E-value'.

- L149: It would be good to be more explicit by what you mean by 'We then compared'. Figure 1 implies that you limited hits to those that intersected both, but it would be good to be clear about that in the text.

- L170-171: Please briefly explain the difference between the two sets of models that are compared (i.e., explain why there isn't just one set of nested models compared).

- L172: 'null model (M1a and M7)' should be 'null models (M1a and M7), which'.

- L181: 'other' should be 'another'.

- L216: I suggest the comma after 'literature' be changed to a colon.

- L227-229: The orthologous groups are gene families encoded by specified species, so they are present in the species' genomes rather than the species present in the groups. This should be re-worded to reflect this distinction.

- L247-249: Should clarify what you mean by 'studies where groups of genes related to specific characteristics are studied', as it's not clear whether you are referring to the same actual overlapping functions/genes, or just similar functions, or whether you just mean any genotype/phenotype comparison more generally.

- L299-301: Should specify that these known polymorphisms and observations have been in humans.

- L447-448: I suggest you remove 'In fact, the cetacean hearing has evolved to be remarkable' (while clarifying this section as suggested by Reviewer 1).

- L524: 'fasta' should be 'FASTA'.