Dear Gavin, here, you will find a revised version of our manuscript incorporating most of the changes you suggested. We appreciate all the time you have invested in our work; it has been a nice peer review process. We hope you consider that the revised version is now in good shape to be recommended in PCI genomics.

Hi Dr. Opazo and colleagues,

I have read through your updated manuscript and responses. I find it has greatly improved, and the reviewers' comments have been adequately addressed. However, prior to recommending your work I do have a few final comments that I would like to see addressed.

My major comment is regarding the functional enrichment analysis. You clarified that only ion channel genes were used for this analysis, but readers will likely find this somewhat confusing as they will likely expect the analysis to have been conducted against the background of all genes, rather than just ion channel genes. This distinction means that although the ion channel genes/functions identified through this analysis are indeed the best candidates for positive selection among ion channel genes, they are not necessarily the strongest candidates of positive selection across all cetacean genes. Explicitly clarifying this point would be important, as otherwise it could be misleading about how important ion channel genes may have been for the aquatic transition.

We appreciate this comment. This concern originated from our lack of enough details regarding the analysis performed. The analysis we conducted using the mammalian phenotype database (Blake et al. 2003, Eppig et al. 2015), as implemented in the enrich web platform, tested our gene list of positively selected ion channels (query list) against the gene list as implemented in their database (subject list), not against a list of just ion channels. To solve this problem, we reworded the methods to make them clearer.

Blake et al. (2003). Nucleic Acids Res. PMID: 12519980
Eppig et al. (2015). Nucleic Acids Res. PMID: 25348401

The new section reads as follows:

Physiological phenotypes associated with the positively selected genes

To understand the physiological phenotypes to which our list of ion channels with the signature of positive selection are associated, we performed an analysis using the mammalian phenotype ontology database (Blake et al. 2003; Eppig et al. 2015) as implemented in the Enrichr platform (https://maayanlab.cloud/Enrichr/) (Xie et al. 2021). In this analysis, we compared our gene list of positively selected ion channels (query list) against the gene list of the platform (subject list), which has 9767 genes grouped in 4601 phenotypical categories. The platform calculates the adjusted p-values from all raw p-values using the FDR procedure, and we considered only the phenotype categories with adjusted p-values equal to or below 0.01.

Blake et al. (2003). Nucleic Acids Res. PMID: 12519980

Eppig et al. (2015). Nucleic Acids Res. PMID: 25348401
Xie et al. (2021). Curr. Protoc. PMID: 33780170

On a related point, although Figure 3 provides a clear high-level overview, there are insufficient details given to help the readers evaluate the overall results. Readers need to dig into the supplementary tables to see specific details, which are tables of P-values for significant categories.

We understand this concern but would like to continue using this figure instead of the tables in the main text. The actual figure delivers the main message, primarily for general readers. Readers who are especially interested in the details will find them downloading an Excel file. Checking supplementary material is common today, so we should be fine. Additionally, the actual figure will call the readers' attention because it is a diagram that can be easily included in a presentation. We respectfully request not to make the change.

Although significance tests are important, a report on the effect sizes of the tested categories would be more relevant, and would help readers evaluate the work. I strongly recommend that a figure or small table be (re-)added to the main text to provide more specific details to help readers evaluate this analysis (ideally that would contain effect sizes, e.g., of odd's ratios of how enriched a certain gene / function is among positively selected genes, or the actual count breakdown). The original Table 1 was better, but without effect sizes (or at least the numbers of genes in each category vs not in that category) it was not very informative.

Thank you for the suggestion. We agree with the reviewer. We have added additional columns in Tables S2 to S5 with all the output information from the Enrich server. This includes the number of positively selected ion channels in each category and the number of genes per category, among others (Supplementary Tables S2 to S5). Also, in the methods section, we have added the total number of phenotypical categories (4601) and the total number of genes in all categories (9767) of the Phenotype Mammalian Ontology Database Level 4 2021, the one used for the Enrich analysis.

Last, in the enrichment analysis methods section, the authors describe identifying significant hits with an "adjusted probability value of less than 0.01" and where "The adjusted probability is calculated from the resulting list of categories with raw p-values equal to or lower than 0.05 through the procedure of [FDR]". This makes it sound like FDR correction was run based on the set of raw P-values below 0.05, which would be incorrect (this correction would need to be run on the set of raw P-values for all tests, not just those considered significant). If I am interpreting this description correctly, then the authors would need to change their correction approach and the enrichment results would need to be re-analyzed after appropriately applying multiple-test correction. Please clarify what was done exactly and whether I am misunderstanding what correction was performed.

We understand this concern. Similar to our first response, this confusion arose because we needed to include more details. The confusion stems from our description of the FDR procedure as if we had performed it when the one who applies the FDR procedure to the raw p-values is the Enrich platform. We took the platform results and considered them

valid only with adjusted p-values equal to or below 0.01. To solve this confusion, we reworded the text. It reads: "*The platform calculates the adjusted p-values from all raw p-values using the FDR procedure, and we considered only the phenotype categories with adjusted p-values equal or below 0.01.*"

See our first response for further details.

Other comments

The authors use the phrase "conquest of the aquatic environment" in numerous places. I strongly suggest that they use phrasing like "aquatic transition" instead, at least in most cases.

We agree with this comment. In the new version of the manuscript, we leave the expression "conquest of the aquatic environment" only once.

It would help readers who are non-experts in the cetacean field if the crown and stem clades were indicated on the phylogenetic tree. Please add this to Figure 2.

We agree with this comment. In Figure 2, we differentiate the stem and crown with different colors, and in the legend, we added a sentence explaining that.

I suggest you add a README file to your GitHub repository to describe the scripts (or at least the subfolders), and you could unzip the scripts so they can be explored online as well.

As requested, we added a readme file and unzipped the script folder.

Also, please add the supplementary tables to a more permanent location, since GitHub repos can be taken down over time. For example, a Zenodo repository would work better for this purpose.

We understand the concern regarding Github. We initially had the supplementary information in Zenodo, but when we tried to generate a new version, the platform did not work. We sent emails via a web page and messages via X (ex Twitter), but we have not received a response. Thus, based on our experience, we moved to Github.

In the abstract, ion channels are first brought up in an abrupt way: "Ion channels are a crucial component of the cellular machinery for the proper physiological functioning of all living species." My understanding is that the authors specifically are interested in ion channels because they hypothesize that they might have particularly been under selection for physiological changes during the aquatic transition. Bringing up ion channels in this context would be a lot clearer, as at the moment it is not clear why you focused on them specifically (in the abstract).

We agree with this comment. To make the presentation of the ion channels smoother, we reworded this passage. The new text reads: "*Among the different molecular systems that*

***maintain cellular homeostasis, ion channels are crucial for the proper physiological functioning of all living species.*"**

More details on where the phylogenetic tree was taken from are needed (this is reported in a figure legend, but should be in the methods itself). If the tree generation required some manual steps, it would be good to include these in your GitHub repo. Otherwise please make it clear where the file was downloaded from.

We think that we do not need a whole material and methods section to justify the tree topology and divergence times used. Since it can be explained in two lines, we believe that including it in the legend of the figure is enough. We respectfully request you to allow us to maintain this information where it is.

In the methods, I found this sentence a little vague: "Both analyses take into account the sister group relationships of the included species". It would be good to clarify what exactly is meant, e.g., do you mean that the phylogenetic similarity of species was taken into account for these analyses? Or something specifically about sister species?

With this sentence, we are trying to communicate that the method used to infer OGs and HOGs considers the phylogenetic relationships of the species involved in the study. To clarify this statement, we reworded it, and now the new text reads: "*Both analyses take into account the phylogenetic relationships of the included species.*"

The dN/dS null models are described as only permitting sites with omega < 1, but to my understanding these models permit sites with omega <= 1 (e.g., see Table 1 of this paper: https://doi.org/10.1093/molbev/msad041).

This is right. We modified the text accordingly.

Please include a mention of the Creative Common license that the silhouette images downloaded from PhyloPic are under (and include a link to the license if that is required under the re-use terms).

Done

The authors mention: "Among the associated genes related to heart physiology in cetaceans … the Sodium Voltage-Gated Channel Alpha Subunit 5 (SCN5A) is the most frequent." I think it would be more correct to state that homologs to this gene were identified, or use genes (plural), rather than "the gene". My thinking is that there are presumably multiple such genes per genome if it was the most frequent hit, so clearing that up grammatically would be good.

We believe there is confusion, and the text needs to be reworded. The mammalian phenotype database assigns genes to different physiological categories, and categories can have exclusive and shared genes. So, if you take all the genes associated with each of the 18 categories related to heart physiology and count them, you will see that the Sodium Voltage-Gated Channel Alpha Subunit 5 (SCN5A) is the most frequent. This is what we are trying to say and why we are writing this sentence this way.

To solve this problem, we reworded this sentence: "*Among the genes associated with the recovered categories related to heart physiology (Fig. 3 and Supplementary Table S2), the sodium voltage-gated channel alpha subunit 5 (SCN5A) gene is the most frequent.*"

Nav1.5 is defined in regards to SCN5A, but it would be good to quickly describe the relationship with Nav1.7 too, as this wasn't clear to me upon re-reading.

The recommender is right. To solve this problem, we modified the text in the first appearance of Nav1.7 in the result and discussion section.

The new text reads: "*Although the NaV1.5 and its paralog, the Sodium Voltage-Gated Channel Alpha Subunit 9 (NaV1.7) sodium channel have similar selectivity filters (Shen et al. 2019; D. Jiang et al. 2020), the NaV1.5 channel.....*"

Related to the above, I found the paragraph starting with "Although the NaV1.5 and NaV1.7 sodium channels have similar selectivity filters" (page 13) confusing, as it began as a comparison between Nav1.5 and Nav1.7 and then changes to contrasting human and cetacean Nav1.5. Please rephrase this sentence. Introducing Nav1.7 vs Nav1.5 would be a good starting point.

We agree with this concern. To solve this problem, we first defined at the beginning of the paragraph that NaV1.5 and NaV1.7 are paralogs. Then, when comparing the affinity for TTX due to a single amino acid difference, we removed the word "human" so the comparison is now species-independent. So, the structure of the text is as follows: 1) presenting both channels, 2) defining their homology relationship, 3) showing the difference in affinity for TTX, and 4) presenting the results of cetaceans.

"Frogs" are mentioned as marine animals at the top of page 13. Please correct this.

Done

Please note -- I have corrected a few typos and made some minor suggestions in the attached Word document (with tracked changes).

Thanks, we followed your suggestions.