Dear Dr. Westbury and colleagues,

We have now received the reviewer's report for your manuscript. While like me he found that this new version has been improved a lot, and the new analyses you added have brought some clarity, the text is still in need of some important clarifications. The reviewer has raised six points to address raised, and I agree with him. Thus, I am inviting you to revise your manuscript considering carefully the reviewer's comments, and resubmit a revised version.

At this point, I consider this as a minor, but important revision.

I look forward to receiving your revised manuscript.

Best Wishes,

Michael C FONTAINE

Answer: Thank you for your consideration of the. We have addressed all comments by the reviewer below. We hope that not your find the manuscript more suitable for recommendation.

Reviews
Reviewed by Simon Henry Martin, 09 Dec 2021 22:21
The authors have revised their manuscript, addressing my requests for more interpretation of the introgression analyses, and to add the f-branch analysis. I think brings some clarity to the results, suggesting that ancestral gene flow might be responsible for many of the observed signals. Having said that, there is no easy way to interpret all of the results, and I think this paper might highlight how at some level of complexity, different introgression scenarios become impossible to accurately reconstruct. I would not suggest any further analyses for this paper, but I do have some more suggestions to clarify the message/reasoning.

1. Despite the extensive evidence for gene flow on a per species level, it is important to note that the estimated proportions of the genome affected are small. Looking at the fb estimates of Malinsky et al. 2018, many are over 5%, whereas none of the estimates in the present study reach that level, with most below 1%. This suggests widespread gene flow but in small quantities, though it is also worth noting that f estimates tend to be lower the older the introgression event (Martin et al. 2015 MBE).
Answer: We have now added a comment onto these values but also a cautionary sentence about other factors that could make these values lower than they truly were (i.e. they represent relative values and decrease the older the introgression event is).

2. I generally agree with the interpretation of the QuIBLE results in tables S3 and S4, described in lines 149-168. However, both the text and table seem to use the term "outgroup" incorrectly. For example, in the triplet ((pilot, bot), orca), it is not correct to refer to pilot as the outgroup. I don't know what the ideal term would be. In a hypothesis-based framework

such as ((African, European), Neanderthal), it may be suitable to use the term “control taxon”, but since there is no prediction that one ingroup would be more likely to be subject to gene flow than the other, this seems wrong. Perhaps just the term “other ingroup” would be acceptable.
Answer: The default output from the software lists this as “outgroup” which is why we had it this way. However we now see how this could be confusing. We decided to go with the term control taxon.

3. Line 312:
“When considering the uppermost limit of when two target genomes coalesce (equating the oldest date), and the lower confidence interval of each divergence date (equating the most recent date) (McGowen et al., 2020)”
I think this statement is talking about comparing the lower bound of the divergence time with the upper bound of the estimated cessation of gene flow, but the way it’s worded doesn’t sound like that to me.
Answer: Your interpretation was correct. Since gene flow ceased after the initial divergence and we wanted to be able to compare the values where the timing of these events are closest, i.e the oldest date at which gene flow may have ceased with the most recent data of divergence. We have now changed this to “When considering the upper bound of when two target genomes coalesce (equating the oldest date), and the lower bound of each divergence date (equating the most recent date)” and hope that has made it clearer.

4. hPSMC and ILS
Various lines seem to be suggesting that the hPSMC estimate of the date of cessation of gene flow could also indicate ILS of ancestral alleles in the absence of gene flow (line 315, 321, 324, 336, 340, 342, 345). My understanding is that the hPSMC approach used gives an estimate of when gene flow ceased, not when lineage sorting ceased. Incompletely sorted alleles could persist after the cessation of gene flow, and their coalescence times will necessarily all be older than the cessation of gene flow, so I don’t think the hPSMC result tells us anything about lineage sorting.
Answer: This could be an issue but it was discussed in some previous manuscript iterations with a reviewer which is why we added the two potential options. The thought behind this is that if the Ne of a population is very large, even though ILS is a more ancient event, it would take more time for the ILS genomic region to break down through recombination. As hPSMC uses the length of runs of heterozygosity/homozygosity (in our case caused by allele differences between species), if ancestral Ne was high, regions of similarity due to ILS could in theory stay intact for longer, delaying the exponential increase in Ne (indicating the point of coalescence) we see in the hPSMC results. The pairwise comparisons with the biggest discrepancies between divergence time and end of lineage sorting/gene flow do indeed have the biggest pre-exponential growth Ne (Supplementary table S7) which we now mention in the results. However it is difficult to truly estimate the Ne at the times of divergence and its impact on the retention of regions displaying ILS therefore we opted to keep both options available for interpretation.

Could the hPSMC cessation of gene flow estimates instead be biased downward by some artefact? One idea that comes to mind is that mapping to the baji genome could bias the analysis to conserved regions, leading to underestimated coalescence time estimates.
Answer: We have tested whether there are reference biases in previous studies and this seems to not be the case. We mention this and provide the citations in the Cessation of lineage sorting and/or gene flow section in the methods.

5. Node age analysis and ILS
In their analysis of node ages as additional evidence for recent gene flow, the authors again seem to suggest that lineage sorting can result in recent node estimates (line 396). I don't think this is right. If a gene tree groups killer whale with white-sided dolphin (Figure S6 B) due to ancestral ILS alone, the node would be older than the speciation date, not younger. The authors seem to agree with this in line 407. So I think the authors can confidently attribute this younger date to gene flow.
Answer: Yes that is correct. However since we use all windows showing a single topology this means we include the alternative topologies that arose both due to ILS and gene flow. Due to these ILS driven topologies our overall estimates will be somewhat older if we only used windows with alternative topologies that arose exclusively due to gene flow. We now see that the last sentences of that paragraph could be a little confusing in regards to this so decided to remove them.

Line 408-409 seems to then show some misunderstanding. A gene tree that is discordant with the species tree can arise through ILS in the ancestral population, and does not require that ILS persisted after the speciation event. The gene tree can forever remain discordant with the species tree after lineage sorting is complete.
Answer: We have now removed this section from the paragraph.

6. Finally, just a comment:
The hypothesis that many of the apparent species-specific gene flow signals could indicate differential retention of ancestrally introgressed tracts in descendent lineages is interesting, and should be considered in future studies. Assuming the gene flow occurred more than just a few generations prior to the diversification of each family, we would expect there to be many small introgressed tracts in the genome at the time of speciation, with a similar number of small introgressed tracts in each individual. Therefore, under neutrality, we would not expect the variance in introgressed proportions to change very much after the diversification of each family. To me, this leads to two conclusions. First, much of the introgressed alleles had not reached fixation or loss by the time the families diversified (otherwise in would be impossible to purge the introgressed tracts). Second, selection must have been involved in purging introgressed variation from some lineages more than others to create the significant differences observed.
Answer: That is indeed a very interesting comment and something to consider in the future. Once more population level datasets from some of these species become available it would be very interesting to see how selection has acted across the genomes of the different species that likely both inherited introgressed alleles from an ancestral gene flow event.