From: **Mark Burgman** onbehalfof@manuscriptcentral.com
Subject: Decision on Conservation Biology 21-763
Date: March 29, 2022 at 7:08 AM
To: nahardy.wildlife@gmail.com, nahardy@ualberta.ca
Cc: fjarrad@conbio.org, luzhi@pku.edu.cn

29-Mar-2022

Dear Dr. Hardy,

Thank you for submitting your manuscript "Quantifying wildlife conflicts with metabarcoding and traditional dietary analyses" (21-763) to Conservation Biology. I have received two thorough, constructive reviews and the comments and recommendation of the handling editor, Dr. Zhi Lü. The full set of comments is pasted below.

Reviewer 2 and the handling editor appreciate the effort taken to reviser the analysis and to clarify the assumptions. Unfortunately, however, both reviewers recommend that we reject the revised manuscript, primarily because of their remaining concerns about the methods, and the support for the inferences drawn from the analyses. I know you will be disappointed, given your hard work on this. Their detailed reasoning is below. On the basis of the reviews and recommendation, I must finally decline the manuscript.

However, the topic and the essential idea behind this remain valuable. I encourage you to submit a revised manuscript to a journal with a more specialised focus, or perhaps to our companion journal, Conservation Science and Practice. Thanks again. Best of luck in your work and in applying your insights to improve the practice of conservation.

Sincerely,
Mark Burgman
Editor in Chief, Conservation Biology

HANDLING EDITOR'S COMMENTS
Comments to the Author:

After a careful review to the paper by two reviewers, we find that due to the unclarity of methodology and data treatment, the conclusions of this paper are not solid enough to be published.

REVIEWER'S COMMENTS

Reviewer: 2
Comments to the Author

The manuscript entitled “Quantifying wildlife conflicts with metabarcoding and traditional dietary analyses” compared two parallel methods, i.e., morphology vs. metabarcoding, in identifying dietary components from the scats of the long-nosed fur seal. The analyses focuse on the little penguin and other birds, for conservation concerns. A small fraction of the mitochondrial 12S gene was used to amplify the target bird species from scats. Of 99 scat samples analyzed here, 32 showed positive results in PCR, of which 21 passed the sequencing quality screening and were used for the final analyses. These 21 samples were examined using both morphology (looking for tissue remains, esp. hard-parts) and metabarcoding. Each method resulted in a set of positive results (i.e. bird detected), with partial overlaps. The authors then concluded that the metabarcoding method was more sensitive as it increased the overall detection ratio, with additional advances in being quantitative (it could provide relative abundance information based on number of sequences assigned to various prey species). Based on these results, the authors interpreted the data in a conservation perspective, that the predation from the fur seal might present a threat to the little penguin, which was already challenged by shrinking population.

Overall, the manuscript presented an interesting research, where new technologies were applied in the context of conservation biology, with the hope to construct the food web structure that might serve as a critical basis for conservation management.

However, I have some major concerns in experimental design, description of methodology, and the interpretation of data. I will specify these in below.

My major concern is that the authors showed some very different results from the two methods, without proper validation of the credibility of the DNA results.

1. From what I understand, the whole point of the paper is to come up with a best practice that helps to identify a reliable method in identifying prey items of the fur seal, with a special concern on the little penguin. The results clearly showed that the morphology and metabarcoding methods were not compatible, with large disagreements on species discovery and species identity (the methods were “complementary and quasi-independent”, quoting the authors). As a methodology paper, it is important to identify the source of this discrepancy so that future pipeline could either remove such differences or to develop a reasonable criterion that could exclude false results from the combined data. Although explanations were provided in the discussion about the possible causes of the aforementioned discrepancies, there was no evidence to back up the theory. In particular, it would seem that the hard-parts of the birds should have been sufficient to provide DNA good enough for the 12S amplification, which was, according to the paper, a conservative marker. Nevertheless, assuming that the low DNA success is indeed an intrinsic issue with the pinniped system, it would mean that a DNA amplification based approach is NOT appropriate to reveal the whole picture of their diets. If the PCR results may or

mean that a DNA amplification based approach is NOT appropriate to reveal the whole picture of their diets. If the PCR results may or may not indicate the actual diet, then why bother? Without stringent validation in tissue-DNA quality, it would be impossible to exclude false negatives and false positives from the DNA results.

2. In fact, other than tissue remains not containing sufficient DNA, the DNA failure (specifically, PCR failure) might have been caused by alternative reasons, e.g., extraction method, PCR primer choice, PCR conditions, to name a few. For instance, the study used a DNA extraction kit designed to optimize bacterial success in soil samples. The reason of such a choice was not provided in the paper. But I may assume a relevance to the nature of the scat samples. But wouldn't it be better to use an animal DNA kit instead, given that the targets are birds? Similarly, the rationale in the choice of gene fragment was not provided. (In fact, some details were missing from the main text. It was not until later when I went through the discussion, I could confirm that the target gene was 12S. I assume that the lack of information was likely caused by restructuring of the text during revision. But key information, such as gene choice, should be provided in the Methods.) My point here, is that more stringent test should be carried out to identify the cause of the discrepant results. As the authors pointed out, the metabarcoding method is often more efficient in diversity discovery, therefore providing a generally more desired solution. The results in this study are very unique in this sense. I would like to see some better effort trying to remedy the difference, instead of a simple presentation of the results.

3. Additional issues in DNA methodology

a) It is generally recommended that a host-inhibitor should be applied when amplifying food contents from fecal samples (Shehzad et al. 2012. Plos One. doi: 10.1371/journal.pone.0032104). This is to prevent the host DNA from over amplifying, which will compete with the target (prey). Given that the 12S marker used here is a conservative gene across animals (especially the primer region), it may not be surprising that PCR will pick up a large amount of host homologs. Unfortunately, the paper did not provide a full list of amplified 12S and taxonomic assignments associated with them. It would be good to get a whole picture of the amplicon composition. As a conservative marker, the 12S probably will amply taxa other than birds. So what else are the seals catching for food? This information is very relevant to the conservation of the little penguin. On a relevant note, if 67/99 of the scat samples failed in delivering a bird PCR band, wouldn't that mean that the little penguin is not the main prey item to the seal? Shouldn't this be considered in the discussion about the predation threat?

b) A total of 64，700 DNA sequences were generated in the study, how do we know if these were sufficient? Apparently, deeper sequencing will produce more sequences, which may in turn result in more scat samples passing the quality criteria. Rarefication curves are often adopted in such situations to demonstrate sufficient sequencing. On a relevant issue, the NGS approach is typically more sensitive to amplicons at low abundance. Our own experience is that it often managed to successfully sequence PCR products that did not show bands on agarose gel at all. I will not be surprised if some of the scat samples failed in delivering PCR bands here, could be successful in metabarcoding, if sequencing depth was deep enough.

c) The choice of 12S marker: Is this full-length 12S? What is the length of the marker? Why using this specific marker? Is it the only marker sequenced in this study? The part about COI was completely missing in the methods. Much of these details need to be clarified in the main text.

d) Quantification (relative abundance): I have a major objection on the interpretation of sequence abundance and its relevance to food priority. Shortly, multiple reasons could have affected sequence abundances derived from varied food items: primer bias, mitochondrial copy number variation among prey species, timing of digestion and the production of feces, among others. In fact, any attempt trying to link sequence abundance with the original food composition should be treated with extreme cautions. For some reason, the results related to sequence abundance were mostly moved to the appendix, yet the conclusion and discussion on the relative abundance were kept in the main text. If the authors really intend to leverage on this proposed advantage of the metabarcoding approach, some significant effort will have to be made to make this convincing. Basically, if one cannot judge food composition based on tissue remains, DNA won't work either, with even added technical constrains.

e) Of the 21 scat samples detected with bird amplicons, 5 were identified by the DNA method to the species that were different from morphology. Additionally, 13/21 samples containing tissues failed in metabarcoding. This result was very disturbing. It could mean either that metabarcoding failed in detecting birds represented with tissues, or the 12S marker did not have the taxonomic resolution to differentiate bird species. Alternatively, the morphological characters were not reliable. Either way, it would suggest that much needs to be done to clarify the validity of the two methods. Before that is solved, hardly anything could be trusted from the current results.

f) Haplotype diversity and number of individuals: the thoughts in using mitochondrial haplotype to refer individuals was rather clever. However, the ribosomal DNA is known to be polymorphic, especially in the loop region, even within the same individual (Gong L, Luo H, Shi W, Yang M. Intra-individual variation and transcribed pseudogenes in the ribosomal ITS1-5.8S-ITS2 rDNA of Paraplagusia japonica (Pleuronectiformes: Cynoglossidae). Biochem Biophys Res Commun. 2019 Jun 4;513(3):726-731). Although this reference was a study on nuclear ribosomal DNA, the general idea should also be applicable here. Please provide evidence to justify that sequence variation of this gene (or a fragment of it) is valid in differentiating individuals. In general, I would think that protein-coding genes would be less prone to this issue. Leaving the theoretical arguments aside, it is simply difficult to digest the result where a single seal scat could actually contain DNA from 4 individual penguins!

Reviewer: 1
Comments to the Author

The authors have made efforts to improve data stringency by applying a number of arbitrary thresholds to filter sequences. However, the method regard sequence processing is still unclear at several places. For instance, following a sequence clustering step, 7370 unique seabird sequences and 64700 disaggregated sequences remained, which were further filtered down to 47478 sequences that corresponded to 5 OTUs (L 157-161). However, in the supplemental file, it said that 64700 sequences were filtered to yield 35353

corresponded to 5 OTUs (L.157-161). However, in the supplemental file, it said that 47478 sequences were filtered to yield 35353 sequences (L.105). What happened to those ~12,000 (47478–35353) sequences? How many unique sequences remained after filtering? How were these (a few thousand?) sequences clustered to 5 OTUs? I assume that the authors used yet another similarity score for this step but could not find it anywhere in the manuscript.

Then, the authors picked six out of the 99 total samples that had the most abundant penguin DNA and selected the nine most abundant unique sequences because they each accounted more than 7.5% of the sequences of the sample (L. 177-179). How was this threshold determined? What was the minimum number of reads for a sample? What are the read counts for the other possibly thousands of unique penguin sequences? Also, in Table S4, three of the 10 samples that showed positive detections of penguin DNA only had 2-5 reads each. Those could hardly be considered “positive detections”. To improve data quality, a read number threshold for a PCR should be determined (e.g. 1000) and only those PCRs with sufficient read counts will be included in further analysis.

The authors have carefully addressed some of the questions I raised in last review, but unfortunately there are still a number of serious flaws with data treatments and result reporting. All in all, conclusions about fur seals’ predation on the penguin came from six samples, a sample size too small to make any sound inferences.