12-Aug-2021  
  
Dear Dr. Hardy,  
  
Thank you for submitting your manuscript "Quantifying wildlife conflicts by combining eDNA metabarcoding and traditional diet analysis" (21-372) to Conservation Biology. I have received a thorough, constructive review. The review is pasted below.  
  
Normally, we require 2 or 3 reviews. However, the recruitment of reviewers at this time of year has been slow. The review in hand is thoughtful and recommends 'reject-and-resubmit', raising significant concerns about the methods. Rather than delay things further, on the basis of the review and recommendation, I will decline the manuscript at this time, to give you time to consider these comments and, potentially, to revise accordingly. I will consider a resubmission that fully addresses the concerns described below.   
  
Resubmitting your manuscript does not guarantee eventual acceptance, and your resubmission will be subject to the full peer-review process. You have 90 days from the date of this letter to resubmit. If you anticipate resubmitting beyond that date, please contact me.  
  
To resubmit the manuscript, log into your Author Center at <https://mc.manuscriptcentral.com/conbio>. Click on "Manuscripts with Decisions" and then "Create a Resubmission," which is located next to the manuscript number. Then, follow the steps indicated. In the space provided or as a separate Word document, please include a detailed, point-by-point response to the comments of the handling editor and reviewers. Describe the changes you made to the original manuscript and, if applicable, explain why you did not address certain comments. If you have a Cover Letter, upload it as a separate document.  
  
Thank you.  
  
Sincerely,  
Mark Burgman  
Conservation Biology

REVIEWER COMMENTS  
  
Reviewer: 1  
Comments to the Author  
  
Interactions and conflicts between protected/valued species are interesting research topics and have implications for conservation planning. This study provides empirical evidence for predation between two such species, and the use of fecal DNA for quantitative estimation of the predation impact is relatively novel. However, I have major concerns about the technical treatments in this study (see detailed comments below), which should be carefully addressed for the results and conclusions to be credible.    
  
1. Although both morphological analysis and molecular diet analysis showed similar detection rates for seabirds, they detected seabirds in largely different samples (only 1/2–1/3 overlap). This low consistency between the results is alarming. For samples with visible prey parts (very good DNA source), it is baffling that the DNA analysis failed to pick the signal up. There could be important improvement to the DNA workflow to increase the detection sensitivity hence its reliability. For example, were the fecal samples thoroughly mixed and subsampled from multiple locations per pellet? Multiple independent PCR replicates are also necessary to increase detection probabilities. It was mentioned that duplicate PCRs were conducted, but I saw no further report in data processing and the results regarding the duplicates.  
  
2. During dietary sequence processing, sequences were clustered based on >97% similarity (L.180). One would expect haplotypes with <3% differences (= 7 nucleotides for a ~230 bp fragment) from each other to be collapsed into one sequence following this step. It would then be impossible to analyze haplotype diversity of the penguins in the samples. However, Fig. 5 showed 7 haplotypes, 6 of which (haplotypes 1-6) had < 5 nucleotides differences and should have been collapsed into one during sequencing processing. This is quite confusing.  
  
3. Another issue with haplotype assignments. Were all haplotypes recovered from fecal DNA also found in tissue samples from the local penguin populations, i.e., were they all true haplotypes or possibly errors generated during PCR and Illumina sequencing? Artificial errors are common in DNA metabarcoding, and there should be measures to specifically identify and eliminate them to ensure data credibility. Multiple PCR replicates, sequence filtering procedures, and statistical methods can all help to reduce errors (see Tsuji et al., 2020 Mol Ecol Resour 20:1248-1258 and Tsuji et al., 2020 Environ DNA 2:42-52). It would also be greatly helpful to use the penguin remains in fecal samples to generate haplotype sequences and frequency distribution, and compare the data with fecal DNA-derived penguin haplotype diversity.

RESUBMIT BY NOVEMBER 9TH!!!!

RESPONSE TO EDITOR & REVIEWER

(Aim for it to be the same)

RESPONSE TO REVIEWER

We sincerely thank our anonymous reviewer and Mark Burgman, editor, for reviewing and considering our manuscript for publication with Conservation Biology. We are pleased to resubmit a thoroughly revised manuscript and a detailed response to the reviewer’s concerns. We have been thorough in weaving clarifying text throughout the manuscript and in communicating these changes in our response, and we hope that this effort provides a high level of confidence and clarity in our methods and work.

The concerns raised typically reflected details that we were too brief in or failed at providing in the original submission, but that were indeed done in a way that the reviewer indicated. We thank the reviewer for this feedback, as this served to ensure that our manuscript is understandable and transferable to the wide readership of Conservation Biology.

EXTRA GENERAL RESPONSE CONTENT

Several of the reviewer’s points contained within the first concern highlight valid issues that are current to the field of metabarcoding and that (maybe elaborate or not about how they do not negate this study, rather this study applies tested methods that share these issues, such as overlap in hard parts vs DNA).

Additionally, several points raised in their third concern, are extremely salient but currently not feasible in this area, whilst not precluding high impact and quality papers from publication.

Ultimately, we were conservative in all our sampling, DNA processing and sequencing, as well as sequence quality filtering and management of erorr. In addition to conservative thresholds for inclusion of DNA at every step of the way, we further exclude any sample that made it through all these steps and yet contained less than 2 sequences altogether (they probably contained more, if we were not so conservative cite Thomas or Deagle paper who discuss this), even though they present potentially real penguin consumption, they also potentially represent sequencing error. We then apply a novel, but tested, polymorphism approach to a novel question and system, whilst also remaining conservative, we only selected the most abundant haplotypes within samples, with defensible sequence abundances (> 10% of sequences contained within each sample). This approach had to be done on a sample by sample basis BECAUSE – ASK TINA TO CLARIFY? Thus we hope that the reviewer, editor and readership will appreciate the balance we have aimed to achieve in the level of care taken to control for any sources of error while presenting novel and impactful research.

HOW DO WE FULLY ADDRESS REVIEWER CONCERNS?

Concern #1 re overlap in DNA and hard-parts detected in samples:

The reviewer brings up several key details that needed further clarification or better wording and we document here how we address the reviewer’s concern in the body of the manuscript.

* Firstly, we initially used language such as “comparing” or “combining” methods, and to a certain degree we do as we are aiming to publish the results of both methods together. However, this language may have caused confusion, as we ultimately posit that these assay methods represent two nearly independent environmental surveillance and assaying techniques. We therefore changed the language to “apply” and we clarify on the scope being “overall” as opposed to a sample-by-sample basis (Abstract line 12, Introduction lines 114 and 123).
* We add the following clarifying text to better set up reader expectations of our study and assaying techniques (Introduction lines 118–123): “Due to known biological and methodological differences in dietary information obtained from hard-parts compared soft tissues (Casper et al., 2007a, 2007b; Tollit et al., 2009), and differences in the quantities of DNA available from hard parts compared to soft tissues (McDonald & Griffith, 2011, Rothe & Nagy, 2016), we apply and present the results of the morphological and DNA metabarcoding techniques as simultaneously informative and complementary assaying techniques.”
* Secondly, to ensure the homogenisation step is clear and emphasized, we replaced the former Methods text from lines 135–138 “Subsamples (2 mL) were taken from whole scats directly at the point of collection in the field for DNA-based analyses of prey remains, by homogenising whole scats with individual disposable spatulas and storing in 2 mL in Eppendorf tubes.” With text in lines 143–145 “Whole scats were thoroughly mixed with individual disposable spatulas at the point of collection and as many subsamples were taken from different lobes of the same scat as would fit in a 2 mL Eppendorf tube for DNA-based analyses of prey remains.”
* Samples were indeed screened in duplicate using diagnostic endpoint PCRs which we noted formerly in lines 162–63, however we were too brief in describing what happened afterwards. We have added some clarifying text in Methods line 190 and the following detail from line 190–196: “A total of 32 samples showed target amplicons either in both or one duplicate, whilst none of the extraction blanks and PCR blanks (negative controls) tested positive for target amplicons. We proceeded with sequencing using a single sample of DNA extracts (give concentration) for each of the 32 samples that tested positive for birds, and we include two of the extraction blanks, and one positive control.”
* Note to Tina – I’m wondering if there’s any need to say anything about not performing sequencing in duplicates? I’m wondering if this reviewer is familiar with single-step fusion tagging? I can’t tell if they’re expecting us to comment on why sequencing wasn’t done in duplicate (because it would have cost $2000 to repeat the sequencing on Illumina Miseq? This would be unreasonable and I’m not familiar with papers that did or do this routinely) or if they’re expecting to see us dealing with MID tags from duplicate PCR products and from the first step of a two-step PCR??
* To clarify on sequence quality control and filtering, a concern of the reviewers both in their first and third issues raised, we added text in Methods lines 203–207: “Our bioinformatics and sequence quality filtering procedures are described in reproducible detail in Appendix S1.3. We used Geneious R8.1.5 (Kearse et al., 2012) to process the paired-end sequences and to identify only those sequences with exact matched genetic tags and primers for each sample. Primers and tags were removed after this initial filtering step to leave the target sequences. These target sequences….”
* The reviewer was also concerned with an overlap of only half the samples where both methods detected the same taxon. However, this level of overlap and even less is very common in the literature, especially in our study system. For the reasons clarified in this response and in text, we would be concerned if the overlap were higher and it would make us suspicious of the results due to gut passage times and the different tissues, DNA content and morphological information, that are being targeted by both methods. We added a significant discussion of this issue and content in the Discussion section from lines 424–484.
* For example, we quote Tollit et al. (2009) finding “hard parts identifying the same prey in 65% of 213 DNA prey occurrences. In both regions, rankings between the two identification methods were similar, but the relative proportions of prey species occurrences in the diet were very different. We believe our results provide further evidence that hard parts found in scats are from a composite of many past meals (shown to be up to 7 d when eating Gadidae and even longer if cephalopods are consumed; Tollit et al. 2003), whereas prey present in scat soft-part matrix represent only the most recent feeding events (estimated to be diet over one to two days by both Deagle et al. [2005b] and Casper et al. [2007a]).” Additionally, Tollit et al. (2009) also found that only 20% of scat samples contained identical prey composition using genetic vs. hard-part methods of diet analysis.
* Certain tissues, such as feathers and fur, may also be over-represented in scats as we note in the expanded in discussion (lines 419–423), and likely contribute to a greater number of samples containing those morphological remains than DNA.
* Lastly, the reviewer questions why we did not attempt to extract DNA from hard-parts. Depending on the study objective, it is possible to formally combine morphological and genetic techniques in this way and by extracting DNA from specific pieces of non-identifiable prey tissues found in predator stomachs or hard-parts remaining in scats. However, this would represent a third and only partial assaying technique. The usefulness of this approach must be balanced with the increased costs and time involved. In this proof-of-concept study, it was determined that it was more informative to assay simultaneously all samples with both a bulk morphological and DNA-based approach as these would be more comparable to past assays that only used one and no combination of these techniques for the study species and prey taxa. This may indeed be an area for future investigation. However, we do not include this in our discussion as we believe this is outside the scope of this paper and an area for further investigation, possibly the body of work for a thesis.
* Additionally, we are only aware of a handful of papers that have gone to the trouble of extracting DNA from hard parts identified within scat samples, these keratinised and/or chitinised tissues contain considerably less DNA than soft tissues. We do not recommend this approach for this study system, or indeed many others, as it is labour intensive and expensive for large quantities of samples (> 30), and involves different protocols required for processing and extracting DNA from amorphous material compared to DNA from hard parts. We do add the following sentence to our Discussion lines 456–461: “In addition to this, DNA found in the amorphous, soft parts of an animal is readily available and easily amplified even in a degraded, digested sample matrix compared to DNA available from hard-parts, feather, fur or bone (Mumma et al., 2016; Granquist et al., 2018), the latter chitinised or keratinized tissues often require a different process than that was used for the amorphous scat analysis and the DNA is often of poor quality (McDonald & Griffith, 2011, Zarzoso-Lacoste et al., 2013; Rothe & Nagy, 2016).”
* For duplicate PCRs. Maybe take photos of lab book and include in response??? Is this necessary???

Concern #2, regarding differentiation between species identified in samples and for haplotypes

* Text added by Tina Methods lines 233–onwards: “To enable effective haplotype analysis, we formed the quality filtered sequences into clusters of unique sequences instead of OTUs and selected only the most abundant representative sequences from each of the 10 samples that tested positive for penguins, excluding samples containing trace amounts of DNA. This process excluded thousands of sequences that could be attributed to sequencing error.”
* Make a table as suggested by Tina + include in SI.
* Anything else we want to say about this??

Concern #3, regarding differentiation between species identified in samples and for haplotypes

* We believe that part of this concern may have first originated from confusion and lack of confidence in our methods and which we have thoroughly addressed in our responses to concerns #1 and #2 above, regarding conservative choices made at every step of the sampling and data processing steps of our genetic and morphological workflows. Regarding - Multiple PCR replicates, sequence filtering procedures, and statistical methods can all help to reduce errors (see Tsuji et al., 2020 Mol Ecol Resour 20:1248-1258 and Tsuji et al., 2020 Environ DNA 2:42-52).
* We are therefore confident that the haplotypes selected do not represent sequencing platform or other sources of error. We did sample and sequence the 12s gene for one individual penguin carcass from Victoria, southern Australia, and this haplotype was also identified in the Victorian samples. This is however the only haplotype for this gene currently available to us as previous studies sampled different genes and the whole genome for this species has not beens sequenced
* What studies have published and accessible little penguin haplotypes?
* Check paper / ask Tina to edit/comment 🡪 one of the papers this reviewer highlights actually says that what they are requesting is not commonly possible at this time. Data on haplotypes is not commonly available for all study species. CITE THESE TWO PAPERS IN DISCUSSION.
* Last concern by the reviewer “It would also be greatly helpful to use the penguin remains in fecal samples to generate haplotype sequences and frequency distribution, and compare the data with fecal DNA-derived penguin haplotype diversity.” The reviewer is potentially suggesting that we or others in a subsequent study perform a third assaying technique in order to extract DNA from the most degraded and DNA-poor parts, the seabird hard-parts within faecal samples, and also compare the haplotypes generated from that technique with that which we detected from the amorphous parts of the scat. We believe this is outside the scope of this paper, and propose that this would make a great thesis study or further paper. Quote papers that haven’t done this – this a valid recommendation not a reason to disregard this paper for publication. We posit that there could be severe issues with the idea proposed however. How could one be sure that the DNA extracted from hard-parts would not be contaminated with DNA from the soft parts of the scat?? (Remember these contain the least amount of DNA and are the most difficult to amplify that DNA, preferentially amplifying fresher DNA from the soft tissues / amorphous parts of a scat sample).

FINAL EXTRA CONTENT

* Note on best avenues moving forward and add one or two of these thoughts to the MS – for example in future work comparing DNA / morphs, we don’t feel that efforts should be directed at understanding too deeply the differences in results we get from these methods. We believe that the biological processes and constraints that result in us getting different results from hard parts and DNA are beyond our control, for example, we can’t ask seals to stop regurgitating prey in the field or passing hard parts up to a week after the soft tissues passed. Even if we can control for this in captive studies, the wild samples will still have this problem.
* We CAN engineer solutions to the ultimate problems of diet analyses which are to obtain quantitative (numerical abundance / biomass) of prey consumed, particularly in sensitive prey or predator species. And this study does that, save for the lack of baseline data on prey species’ haplotypes currently.
* Good idea maybe inadvertently by the reviewer or at least the papers they cite to us to highlight that issue in the paper.