Palaeoproteomic identification of a whale bone tool from Bronze Age Heiloo, the Netherlands, reviewer response

We like to thank both reviewers for their kind words and constructive comments. Below you will find a point-by-point response to each of the comments. Additionally, we have also modified the manuscript in “track changes” mode and attached it to this submission. We hope that the revised manuscript will be to your satisfaction.

Review by anonymous reviewer 1:

Introduction.

Sentence : “Ancient DNA analysis is understood……” you can indicate that the quantity of material required for analysis is too great for these archaeological objects.

We have added a sentence regarding the used sample sizes in aDNA analysis and their citations. We also added citations to studies using minimally invasive protocols for aDNA analysis of bone/antler tools.

In the next sentence, for paleproteomics, you could indicate that there are minimally invasive methods available.

We have written a few sentences to discuss sample sizes in palaeoproteomic studies and the available minimally invasive methods.

For SPIN analysis, which is a shotgun analysis, it is possible to obtain phylogeny information.

Although generally speaking SPIN is able to provide taxonomic identifications to a more precise level than ZooMS, its ability to identify taxa is determined by the used database. As such it is not able to identify taxa not included in the database nor can it identify new sequence variants. As a result, it would be incorrect to describe SPIN as a phylogenetic analysis.

It would be interesting to add a paragraph on the use of proteomics on archaeological tools. There are examples in the literature.

We agree with the reviewer that there are numerous interesting case studies on palaeoproteomics for the taxonomic identification of bone tools. We have added a few lines highlighting the main research aims of these studies as well exemplary citations.

Materials & methods

Page 3, Figure 1 c. can you indicate the position of the bone

We have added a circle indicating the position of the specimen.

Page 3, Why 2 ZooMS analyses? Why take 3 samples of 38 mg? Current methods show that 5-1 mg is sufficient. please clarify these points.

The text now contains an explanation of the double ZooMS analysis. As for the sample size, we discovered that the average size of 38 mg was a left over from a previous draft that contained more samples. During sampling of one of these other objects, a flake with a mass of 85 mg detached, inflating the average sample size. We have now adjusted the text to contain the measured masses of the three analysed samples.

Page 4, ZooMS

NH4HCO3 ----- > incorrect number format, NH_4HCO_3

We have adjusted the notation of ammonium bicarbonate.

Page 5, 50µL ----- > 50 µL a space is missing

We have added appropriate spacing.

Page 5, Missing MALDI information: mass range, positive mode, calibration....

The missing technical details have been added.

Page 5, a space is missing “50 mM”

We have amended this.

Page 5, SPIN. did you measure the quantity of peptides injected? please indicate

The amount of protein extracted from the samples was not quantified so unfortunately we have no measurement of the amount of peptides injected.

Page 5, Maxquant:

Custom database of COL 1 form NCBI? Or Swissprot? These are the complete or mature sequences (with ou without signal peptide)?

We have edited the text to clarify that we included the signal peptides and the source of the proteins.

Why you used 2 version of Maxquant?

There was a substantial period of time between the initial MaxQuant search with v1.6.7.0 and the second with v2.4.10.0. As the first search was intended more as a broad screening search to establish what proteins could be recovered and which might be most useful for taxonomic identification, we did not think it would be necessary to redo the search with a newer version of MaxQuant. Although differences in output between versions cannot be discounted, for the purpose of identifying proteins to further investigate the effects of different versions seem marginal.

Why you used specific tryptic cleavage? Using the semi-trypsin mode, the number of peptides identified increases. Did you test it?

We have added some text clarifying our choice for trypsin-specific digestion. In short, although semi-tryptic cleavage can increase the number of peptides identified, this is not always the case.

Bioinformatics data analysis, Mass spectrometric data. “default settings and protein sequence database” can you explain. Database of which proteins? (collagen? Others?)

The database referred to here is the database produced by Rüther *et al.* 2022 consisting of the 20 most intensely expressed bone proteins, with species entries for 156 mammalian

species. With default settings we mean that we did not change the functions in the R script used for SPIN analysis.

Results

ZooMS, page 8. Why didn't you watch the deamidation?

We thank the reviewer for raising attention to deamidation and we have now calculated the deamidation ratio of the proteins observed in the targeted MaxQuant search using deamidation, following Mackie et al. 2018, and reported the values in the manuscript.

SPIN page 10. You have a database of cetacean collagen and you identify bos. There must be information missing somewhere.

The SPIN analysis was done with the SPIN database, which has a much wider coverage of mammalian species, including *Bos taurus*. We've clarified the text to indicate which database was used for the SPIN analysis.

Page 11, "For each of the five SAPs the number of peptide spectrum matches (PSMs)" have you done a blast analysis to check the specificity of the sequence?

As requested we have confirmed the specificity of the peptides covering the SAPs using BLAST and added the results to the manuscript.

Discussion

Page 15 "Three proteomic workflows". You have used 2 different workflows ZooMS and SPIN.

The three proteomic workflows referred to in this sentence are ZooMS, SPIN and the custom-database MaxQuant search. We count SPIN and the custom-database MaxQuant search as different workflows, because the peptides and protein identifications they use derive from different database searches and they analyse the output in different ways.

You don't mention de novo mass spectrometry analysis. Do you think these could help when analyzing samples without protein sequences? not all laboratories can afford genome sequencing.

We thank the reviewer for raising the subject of *de novo* protein analysis, as we think it is an exciting and very promising technique. A variety of new tools using *de novo* analysis have been published recently, including some using large language transformer models, and they seem very promising. However, *de novo* tools have been known to produce erroneous peptide sequences, some of which can be easily identified, while others are more difficult to distinguish. Before they can be applied to archaeological data these tools still require rigorous validation, especially if they are to be used to establish taxonomic biomarkers which may be used frequently in subsequent studies. Due to their infancy we do not think it falls within the scope of this manuscript to discuss *de novo sequencing* at length and a passing mention would not do justice to the complexity of understanding and validating *de novo* output.

Reviewer 2:

The aim of this study is indeed the proteomic study of the specimen under investigation, but I think some more information on the context can be useful to better understand and contextualise. For example, were any animal bones found in this site? of which animals? other bone artefacts?

As requested we have added some additional information regarding the faunal assemblage at the site as well as any other worked bone objects.

Figure 2: Consider adding markers on the figure to indicate where the samples were taken for proteomic analysis. Specify how the samples were collected and what was analysed (powder or fragment).

We have specified in the text that samples were collected by scraping with a scalpel. Additionally, we have indicated on Figure 2 where the samples were taken from.

ZooMS Protocol: Specify how long the samples were left in acid.

We have now specified how long the samples were incubated in acid.

Since the samples are from a single artefact, it might be useful to include the ZooMS spectra obtained from the analyses (even with both protocols, cold acid and AMBIC), to show the spectrum quality, the preservation of the sample, and how this changes based on the protocol in terms of obtaining taxonomic information.

We have added a figure (Figure 4), displaying the ZooMS spectra.

Why was a destructive sampling method chosen over a less invasive one?

We have added a few lines to the introduction to explain our choice for destructive sampling.

SPIN: Are there differences in the specificity of the determination within the SPIN framework considering the different protocols used?

We were not able to compare for differences in specificity between the AmBic and cold acid protocol, as extracts from both protocols were loaded on the same StageTip. The LC-MS/MS acquisition combined the peptides extracted with both protocols and provided a single file as output.

“The extraction blank was matched to Bos sp. by the SPIN script, but it has been excluded from Table 2 as 84% of the peptide spectral matches in the blank were derived from the trypsin used for digestion. The blank intensity signal can therefore be regarded as background noise.” How do you explain this? Is it normal for the blank to be matched with Bos? Elaborate on this point to better support and validate the other results obtained.

The relative protease intensity is a good measure of the protein richness of a sample, as the same amount of trypsin was added to each sample. The fact that the relative protease intensity in the blank is 84% vs ~1% in the samples indicates that there is a lot less protein in the blank. This small amount of protein is the background contamination. It is common for any sample to contain a small degree of background contamination in shotgun proteomic data and the high protease intensity in the blank suggests that it indeed concerns a small amount in our study.
Common lab contaminants can derive from a variety of sources, such as human skin, bovine or sheep keratins and other proteins. SPIN matches all PSMs from a sample to its database and will give the best match as taxonomic identification, but the best match is not necessarily a good match. Other metrics, such as the site count and relative protease intensity provide information on the robustness of the taxonomic identification. In the case of the blank, they indicate that there is indeed no robust support for the taxonomic identification. It merely seems that in our case most of the background noise is consistent with *Bos* sp., but this does not raise doubts regarding the other results. In order to obtain a relative protease intensity all identified PSMs, including contaminants, were processed through SPIN. However, for the targeted database analysis all contaminants were excluded.