nature portfolio

Peer Review File

Phylogenetically-informative proteins from an Early Miocene rhinocerotid

Corresponding Author: Dr Ryan Paterson

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Referee #1

(Remarks to the Author)
It is a beneficial attempt to analyze the phylogenetic relationship of rhinocerotids by the enamel proteome which pushes forward the earliest record of ancient protein sequences to the Early Miocene. But in specific aspects, there are still some areas that need improvement and enhancement.

Line 3: The IUGS (International Union of Geological Sciences) Executive Committee has voted unanimously to ratify the proposal for formal adoption of the chronostratigraphical/geochronological unit divisions subseries/subepoch within the International Stratigraphic Guide as approved by the International Commission on Stratigraphy and forwarded to the IUGS EC on 24 March 2021. The subseries/subepochs are now incorporated in a six-tiered chronostratigraphic hierarchy of units that are formally defined by a designated GSSP (Global Stage Stratotype and Point) at the base of designated type stages. Henceforth, subseries/subepochs of the Cenozoic are to be denominated by capitalised positional adjectives - Lower/Early, Middle, and Upper/Late - added to the names of the relevant series/epochs (Aubry et al., 2022, Episodes).

Line 37: The recovery of protein sequences providing novel phylogenetic insights exceeds 6.5 Ma (Late Miocene) based on Demarchi et al. (2022, eLife: This discovery has fuelled the analysis of ancient proteins from other mineral matrices, namely tooth enamel in order to reconstruct phylogenies).

Line 38: About the recovered enamel proteome dates back to 20 Ma, have authors analyzed any other materials from the same site? Do they show comparable proteome?

Lines 46 and 47: There is no any fossil record to support these two divergence times, and the existing fossil records are much later than these two times.

Line 58: Should “beyond” be “before”? On the other hand, the total-evidence phylogenetic analysis of deep-time fossils widely used in paleontology incorporates genetic data rather than just morphological observations.

Line 71: Here, the authors ignore another group of Rhinocerotidae, namely Aceratheriinae, which is more basal than the sister group composed of the Rhinocerotinae and elasmotheriinae. This deep ‘basal split’ at least occurred between the Aceratheriinae and the Rhinocerotinae-Elasmotheriinae.

Line 72: Although the Rhinocerotidae appeared in the Late Eocene, they cannot be described as “bursts” because fossil records are quite rare.

Line 78: Should “relatively-young” be “relatively-old”? However, these two genera, Elasmotherium and Coelodonta, did not appear until the Late Miocene and Pliocene. When using genetic data to calculate, there must be at least some clues from fossils. If this split age has not been confirmed by fossils, it is not appropriate to use the age as evidence to cite other analysis results.

Line 89: In the Miocene and Pliocene, the Arctic region was not in the same frozen environment as today. Although the current permafrost environment of this locality favors protein preservation, the temperatures varies from 20 Ma to the present. The temperature here was high until the Pliocene (Brigham-Grette et al., 2013, Science; Ballantyne et al., 2010, Geology;

Csank et al., 2011, Earth Planet. Sci. Lett.). In this manuscript (Lines 365-367), the authors also indicate “In the early Miocene, the Haughton crater lake and its surrounding environs experienced a significantly warmer annual temperature (8-12°C) than the present day). How can authors ensure the biomolecular preservation of Early Miocene fossils in a favourable temperature? Could authors do more evaluation based on the climate change? I notice authors have prepared some experimentally heating samples, have they conducted protein extraction on these samples? It may also help validate the reliability of the results. Thus I suggest the authors discuss more about the possibility of enamel protein preservation to this deep time.

Line 94: Why does this tooth belong to the genus Epiacaratherium? There is no description or comparison.

Line 97: In the paper "Phylogenetic signal in primate tooth enamel proteins and its relevance for paleoproteomics" of this research group, it mentions that for 1-2Ma samples with simulated 1139 amino acids (aa), the phylogenetic placements of most nodes at family level are consistent with the reference species tree. But for older and more degraded samples ranging from 5-10Ma (simulated enamel proteome of 593aa-98aa), the resolution and support of the phylogenetic tree are relatively low. This old Rhinocerotidae proteome reveals only 251 aa, maybe its application on phylogenetic analysis need more evaluation. As observed in their constructed phylogenetic tree, the posterior probabilities for some nodes are not high. Could they perform a similar analysis of sequence conservation and phylogenetic signal of all available proteins of the relevant rhinoceros species?

Line 104: In addition to comparing the enamel proteomes of the two rhinocerotids, it is also important to compare the proteomes of other fossil taxa.

Line 115: While the authors have employed some strategies to manually inspect the reliability of all the spectra, why they do not apply the criteria used in previous studies to validate the SAPs, which requires two spectra, and at least 1 PSM identifies the SAP using both y and b fragment ions. That is a stricter strategy.

Line 162: In recent years, many phylogenetic analyses in paleontology have attempted to use total-evidence analyses. The rhinoceros at this location has a nearly complete skeleton, why not combine morphological and molecular data for a total-evidence phylogenetic analysis?

Line 170: The Elasmotheriinae and Rhinocerotinae form a sisters group, so Elasmotherium sibiricum, a species in the subfamily Elasmotheriinae, cannot have a direct divergence with the Rhinocerotina, a tribe of the subfamily Rhinocerotinae.

Line 316: Should “after destructive palaeoproteomic analysis” be “before destructive palaeoproteomic analysis”?

Line 721: Extended Figure 1 showed the different PTM profiles for CMNF-59632, compared to enamel proteomes from other ancient rhinos and a mediaeval ovicaprine. While the CMNF-59632 showed obvious higher incidence of PTMs related to oxidative degradation, the proportion of different type of PTMs of tryptophan or histidine in these samples also varies a lot. Do the authors have any interpretation for this phenomenon?

Referee #2

(Remarks to the Author)
The authors report palaeoprotoemics data on ancient enamel samples and they have done thorough analysis for assessing the evolutionary history of rhinocerotid specimens. The significance of the paper are the ancient samples (which are around 21-24 Ma) and good preservation of proteins within, that allowed to analyse evolutionary history of rhinocerotids. It is a well written paper and in general, of good quality (in terms of approach and presentation). The proteomic analyses are performed well, with thorough data analysis, peptide assignment and filtration (I cannot evaluate the phylogenetic analyses well as this falls out of my scope of expertise). The results are robust and will be of great interest to the scientific community.

I have some questions, thoughts and suggestions for possible improvements of the paper.

I would be keen to know more about what future implications the work gives. I understand that proteomic analysis of such ancient samples will allow to dig the evolutionary history of different samples even further, but is there more? Or maybe, I have missed it in the paper (given that I’m not an expert in this?). Perhaps it would be good to highlight main (or most interesting) message more clearly in the paper.

For palaeoproteomic analysis:
I wonder if all the enamel proteins are already cleaved in-vivo (or due to diagenesis)? Would there be a point in doing a workflow where enamel peptides are collected in parallel to digesting remaining proteins with trypsin? Perhaps authors have tried this?

Also, what is the possibility of differences when comparing diagenesis related ptms, recovered peptides with different amino acid lengths and sequence coverages between samples (Figure 2) given that samples were analysed with different instruments (exploris 480 and q-exactive). Perhaps it would be good to discuss this (using the same instrument for all the samples would be an ideal case scenario, but often its not reality. But I think it would be useful to briefly mention possible differences).

Minor suggestion: when referring to citation would be to include the name of the main author, followed by the number of the

publication in the bibliography (this is done in many cases, but not everywhere). For example, sentence “The remaining steps generally follow that outlined by 9”; would read much better : “The remaining steps generally follow that outlined by Cappellini and co-authors (2019) 9. (such case was repeated several times in the manuscript, I suggest checking).

Referee #3

(Remarks to the Author)
In this manuscript, the authors extract proteins from the enamel of a now-extinct Rhinocerotidae specimen estimated to be more than 20 million years old. Seven proteins were recovered and partially sequenced from the specimen CMNF-59632. These were used to reconstruct the phylogeny for this specimen, providing new insights into the split between Elasmotheriinae and Rhinocerotinae.
This study represents a novel contribution to the field and is the oldest protein extraction from a specimen with sufficient preservation for phylogenetic placement. This is an exciting result for the field as a whole and represents a significant discovery. I look forward to it being published.
The manuscript is well-written and appropriately cited. Sections of the manuscript could be improved with greater specificity or with a supplemental (as discussed below) however the methods and data analysis appear very sound.

My major concern surrounds the discussion of observed PTMs in the specimen. Given the reliance on PTMs to establish that the proteins are endogenous and that the enamel matrix has preserved its closed system, their reporting lacks specificity. Peptide lengths are reported as being “shorter” and PTMs related to diagenesis are reported quite vaguely. In the figures (especially Fig 2) it is very difficult to estimate counts or percentages and the raw numbers are not reported in the figure captions or the additional information. The reader would be required to look through the raw data and reconstruct this information themselves or guess based on the figures provided – both unnecessary as the authors must have collated this information to produce the figures and to make the statements in the text.
Given the surprising age of the fossil and the use of PTMs to support the endogeneity of the proteins, a greater degree of information should be included in the manuscript. A table or supplement with these results would be an invaluable addition. More information should also be added to the figure captions.

Additional comments:
Lines 84-92: does this section imply that other fauna were also analysed? The paper presents only a single Rhinocerotidae specimen (tooth and tusk). If attempts to extract proteins from other specimens were unsuccessful this would be an important inclusion to the paper or the supplemental. As the authors discuss in the Introduction, there is no systematic understanding of protein preservation in deep time. It remains unclear if this Rhino tooth is an exceptional find or if we should expect biomolecular preservation in other enamel samples from similar regions and time periods.

The image of the specimen CMNF-59632 (Figure 1) makes it difficult (impossible) to do any kind of visual inspection of the fossil. High-quality images of the fossil should be supplied either by enlarging the image in this figure or by providing better quality images in the SI. Additionally, no information on the attributes, taphonomy, or morphology of the fossil are included in the manuscript. The authors state that the “size and morphology” are characteristic of rhinocerotids but this information is not included with the exception of the banding on the enamel surface.

Congratulations to the authors on this very successful manuscript.

Referee #5

(Remarks to the Author)
Referee Report: Nature

A 20+ Ma old enamel proteome from Canada’s High 1 Arctic reveals diversification of Rhinocerotidae in the 2 middle Eocene-Oligocene (Ryan Paterson et al.)

Summary of the key results
This study has recovered protein sequences from an exceptionally preserved Miocene Rhinocerotid enamel from the polar Canadaian High Arctic. The authors satisfactorily demonstrate the authenticity of the proteins and present some discussion on their preservation state and diagenetic alterations. The authors then proceeded to utilize this preoteomic data in a phylogenetic analysis to place the divergence of Epiaceratherium sp. (the current sample) to the middle Eocene-Oligocene. This finding contradicts previous estimates of a basal split between Elasmotheriinae and Rhinocerotinae and also supports a Eurasian origin of this clade.

Originality and significance: if not novel, please include reference
The study is original in successfully applying robust paleoproteomic data to phylogenetic analysis on Miocene specimens. The significance of this study is in providing other paleoproteomic and molecular taphonomic researchers a targeted environment to look for exceptionally preserved fossils capable of yielding such data and as a model in application of paleoproteomic data in phylogenetic analyses beyond the Pleistocene.

Data & methodology: validity of approach, quality of data, quality of presentation

The methodology is sound, and the data appear robust. I would have liked to see an additional method for the determination of degree of diagenesis such FTIR or Py-GC-MS, however, given the totatlity of the data presented, it is my opinion that the evidence is sufficient for protein authenticity. The figures are sufficient.

Appropriate use of statistics and treatment of uncertainties
Appears to be appropriate

Conclusions: robustness, validity, reliability
Based on the data presented, the conclusions seem valid. I am sufficiently satisfied regarding the reliability of the methodology with regards to protein recovery and authentication.

Suggested improvements: experiments, data for possible revision
While the authors have discussed the previous reports of protein sequencing, I also suggest adding a section on the general nature of protein preservation, particularly when it comes to the topic of deep time (although, Miocene may generally be considered rather young to employ the term “deep time”). The traditional conventions for the preservation of proteinaceous material in fossils have continued to be revised. Certainly, they are often altered to the point that sequence information is impossible to retrieve, however the vestiges of original proteinaceous character may remain. I suggest some references below. Similarly, the authors do not acknowledge the work of Stolarski et al., 2023, who have also recovered protein sequence data from Miocene otoliths.

As mentioned, I would also have liked to have seen more discussion about the nature of the diagenetic alteration of the proteins, and its relation to the depositional environment and type of bone. Could they expand, for example, a little more on Figure 2a? Could they also explain in more detail why dental enamel would work better for such a study than say, femur bone? Similarly, I would have liked to have seen a reporting of thermal age calculations (cf. Demarchi et al., 2016, eLife). To what point was the diegenetic alteration retarded? Apart from this, I have raised a few minor comments below.

References: appropriate credit to previous work?
While the authors have used appropriate citations for the most part, they have omitted a reference to Stolarski et al., 2023, which I believe is important to discuss. That was also a paper that showed compelling evidence for the recovery of protein sequences from Miocene fish otoliths (~ 14 Ma). They have also shown how the mineral phase relates to the peptide preservation. That paper does not use it for phylogenetic analysis but in its context, while the current study does indeed push back the age for confirmed protein sequence recovery, there is a difference between pushing it back from 3.7 Ma, as written by the authors, to pushing it back from 14 Ma. Similarly, please see some references related to the recent advances on the knowledge relating to diagenetic alteration of proteins in deep time.

Stolarski et al., 2023. First paleoproteome study of fossil fish otoliths and the pristine preservation of the biomineral crystal host. Scientific Reports 13:3822

Dutta et al., 2020. Chemical evidence of preserved collagen in 54-million-year-old fish vertebrae. Palaeontology 63, 195 – 202.

Umamaheswaran et al., 2023. The diagenetic fate of collagen as revealed by analytical pyrolysis of fossil fish scales from deep time. Geobiology 21, 378 – 389.

Boskovic et al., 2021. Structural and protein preservation in fossil whale bones from the Pisco formation (Middle-Upper Miocene), Peru. PALAIOS 36, 155 – 164.

Alfonso-Rojas and Cadena, 2020. Exceptionally preserved ‘skin’ in an Early Cretaceous fish from Colombia. PeerJ 8:e9479

Wiemann et al., 2018. Fossilization transforms vertebrate hard tissue proteins into N-heterocyclic polymers. Nature Communications 9:4741.

Clarity and context: lucidity of abstract/summary, appropriateness of abstract, introduction and conclusions

The abstract is clear regarding the content of the manuscript. Introduction and conclusions are appropriate, except for the discussion of protein diagenesis in deep time, which I have mentioned above.

Comments:

Abstract:

Pg 2 line 37: While it may be technically correct, the phrasing makes it seem that protein sequence data has not been confirmed beyond 3.7 Ma. I suggest phrasing that makes it more clear that such data had been recovered (Demarchi et al., 2022, eLife, which you have also cited elsewhere; and Stolarski et al., 2023, Scientific Reports) although perhaps not phylogenetically applicable.

Manuscript:

Pg 2 line 61: Please see Stolarski et al., 2023, Scientific Reports https://doi.org/10.1038/s41598-023-30537-8. They have reported paleoproteomic data from fish otoliths from the Mid-Miocene (14 Ma).

Pg 3 line 67-68: This line is somewhat confusing. Do you mean to say that poor preservation record leads to a bias in inferring evolutionary relationships? Please rephrase

Pg 4 line 88: The geological age is advanced from the perspective of protein preservation potential, while most geologists would consider the fossils relatively young. I would recommend making this distinction.

Pg 4 line 89-90: When did the temperature drop initiate? At what timescale did permafrost occur? Since the depositional environment was at relatively higher temperatures, it is likely that diagenetic alteration would have commenced to some degree, hence I would have liked to have seen some comparison of thermal ages to elucidate it.

Pg 4 line 101: Please avoid the term “surprising”. I would recommend replacing it with a less stronger term, such as “unusual”, especially as the deposits are not so geologically ancient, and proteomic data has previously been reported from Miocene specimens.

Pg 9 line 211: Please change “well beyond” to “beyond”

(Methods section)

Pg 17 line 374-375: I would like to see a discussion about enamel in the main section. It is important, I think that such a preservation was not found in the tusk of the specimen, but the main discussion appears to only focus on the frigidity of the locality. However, a discussion on suitable fossil types (for example, eggshell, or fish otoliths), perhaps even from the same organism, that can be targeted for such analyses will be very useful to future researchers.

Version 1:

Reviewer comments:

Referee #1

(Remarks to the Author)
Although the authors agree that a deep basal split is probably more likely to have characterised the divergence between Aceratheriinae and the Rhinocerotinae-Elasmotheriinae, there is a claim that they are unable to test the relationship between Aceratheriinae and Elasmotheriinae with their palaeoproteomic dataset. In fact, combining morphological and molecular data for a total-evidence phylogenetic analysis is a good way to solve this problem.

Line 539-552: The authors employ two filter strategies to construct two datasets: a minimally-filtered dataset and a strictly-filtered dataset. However, the manually inspection part is not clearly explained, I suggest the authors include the details explained in the response in the method section. Thus, other researches can easily reproduce their results and conduct comparative analysis.

I acknowledge the existence of biased fragmentation patterns for different peptides. However, I still think the confirmation of SAPs should be stricter than other amino acid positions because the SAPs usually influence the phylogeny more. Regarding to the SAP confirmation, Presslee et al. (2019; doi:10.1038/s41559-019-0909-z) have suggested that “a higher confidence was given to spectra that identified the selected amino acid using both b and y ions”. But as the authors mentioned in the response, we still do not know how strict the criteria should be to balance the resolution and accuracy. Maybe additional analysis on the experimentally heated samples can give some hints.

Figure 2d should have two y-axes: one for amino acid mutability, whose approach should be explained in the method section, and another for peptide count, whose y-axis is missing.

Extended Figure 1A: when the authors refer to “Before filtering” and “After filtering”, which filter strategy do they mean, minimally-filtered or strictly-filtered? I suggest including a comparative analysis between these two filter strategies.

Extended Figure 8: Please include the overlapping peptides for the other 9 SAPs.

Referee #2

(Remarks to the Author)
The paper now reads very well. The suggestions and comments were addressed and included in the paper, where needed. Congradulations for the manuscript!

Referee #3

(Remarks to the Author)
I thank the authors for their diligence in replying to the reviewers concerns.
I have one final comment on the draft:
While I think the manuscript is technically well prepared, I feel the authors need to be more cautious in their interpretation of these results and their implications for biomolecular preservation outside of this single fossil. The inclusion of thermal age data is interesting but statements like those found in Line 238 (the description of the formation as a palaeomolecular lagerstaette) seem like an overstatement. Especially if there has only been one fossil analysed, as stated by the authors in their reply document. One tooth with exceptional biomolecular preservation does not demonstrate a trend for an entire geological formation when it has been demonstrated in previous research that preservation can be variable between fossils from the same sites and the same geological conditions. At least it does not seem appropriate without the study of additional fossils.
Certainly, this supports a larger trend that high latitude and cold regions are likely to preserve proteins in deep time, and I agree that the authors' findings should encourage further studies (as stated in their conclusion) but with such a small sample size I think more caution should be advised in the wider interpretation.

Referee #5

(Remarks to the Author)
The authors have sufficiently addressed my two main concerns from the previous draft - 1.The lack of thermal ages calculations which, in my view, was necessary in order to substantiate the suggestions of investigating fossils from present cold climates, and 2. An appropriate discussion of the diagenetic context of the preserved proteins. The authors have also included some references that were previously excluded in the initial manuscript, but which I felt were necessary to add historical context to significance of their findings.

I recommend that the manuscript be published.

Response to Referees

Referee #1 (Remarks to the Author):

It is a beneficial attempt to analyze the phylogenetic relationship of rhinocerotids by the enamel proteome which pushes forward the earliest record of ancient protein sequences to the Early Miocene. But in specific aspects, there are still some areas that need improvement and enhancement.

> We thank Referee #1 for their feedback, which was particularly helpful in improving and clarifying our interpretation of the results of our phylogenetic analyses of rhinocerotids.

Line 3: The IUGS (International Union of Geological Sciences) Executive Committee has voted unanimously to ratify the proposal for formal adoption of the chronostratigraphical/geochronological unit divisions subseries/subepoch within the International Stratigraphic Guide as approved by the International Commission on Stratigraphy and forwarded to the IUGS EC on 24 March 2021. The subseries/subepochs are now incorporated in a six-tiered chronostratigraphic hierarchy of units that are formally defined by a designated GSSP (Global Stage Stratotype and Point) at the base of designated type stages. Henceforth, subseries/subepochs of the Cenozoic are to be denominated by capitalised positional adjectives - Lower/Early, Middle, and Upper/Late - added to the names of the relevant series/epochs (Aubry et al., 2022, Episodes).

> We thank the reviewer for pointing this out, and so we capitalised subseries/subepochs of the Cenozoic following Referee #1's comment. We would like to point out though that—despite the cited resolution of the IUGS EC and contrary to what has been ratified for the Neogene and Quaternary (Aubry et al., 2022)—no formalised Palaeogene subseries exists yet, causing inconsistencies of the chronostratigraphic language within the Cenozoic (Aubry et al., 2024) that are difficult to navigate for researchers outside the chronostratigraphic community.
>
> Aubry, M. P., Miller, K. G., Turco, E., Flores, J. A., Gladenkov, A., Grunert, P., ... & Vallejo, F. (2022). Ratification of Neogene subseries as formal units in international chronostratigraphy. Episodes, 32, 152-164.

Aubry, M. P., Piller, W. E., Van Couvering, J. A., Berggren, W. A., Flynn, J. J., Head, M. J., ... & Miller, K. G. (2024). Unifying Cenozoic chronostratigraphy and geochronology: applying the rules. Newsletters on Stratigraphy, 57, 25-36.

Line 37: The recovery of protein sequences providing novel phylogenetic insights exceeds 6.5 Ma (Late Miocene) based on Demarchi et al. (2022, eLife: This discovery has fuelled the analysis of ancient proteins from other mineral matrices, namely tooth enamel in order to reconstruct phylogenies).

Demarchi et al. (2022) is cited in the manuscript (lines 39, 65, 72, 99, 235) as an example of the oldest well-validated shotgun proteomics approach recovering ancient proteins from highly-biomineralised matrices. However, the peptide preservation within eggshell means that the recovered sequences do not provide novel phylogenetic insights, which is an important distinction we would like to stress. In their analysis, Demarchi et al. recover and validate ancient peptides that all derive from a single sequence region from a single protein, providing critical insights into protein preservation and diagenesis, and allowing for their taxonomic assignment to *Struthio*, but unsuitable for more detailed phylogenetic reconstruction. In the current manuscript, we extract peptides spanning multiple sequence regions in multiple different proteins, permitting the testing of hypotheses on the evolutionary relationships through a phylogenetic analysis. Previously, a sufficiently-sized and sufficiently-variable ancient proteome was extracted by Buckley et al. (2019) from a 3.7 Ma site in Canada's High Arctic, and we believe this is the oldest published example of this "Evolutionary Palaeoproteomics" approach in the literature.

So, we have modified this sentence (Line 38 (prev. Line 37)) accordingly: "Still, **while ancient proteins have recently been reported from the Middle-Late Miocene,** the recovery of protein sequences providing sub-ordinal level phylogenetic insights does not exceed 3.7 Ma (Pliocene)."

Line 38: About the recovered enamel proteome dates back to 20 Ma, have authors analyzed any other materials from the same site? Do they show comparable proteome?

We have not analyzed any other materials from the same site, excepting the tusk fragment mentioned in Line 436. We included the tusk's RAW file in our data upload on PRIDE, but we do not discuss the results as proteomic recovery was not high. The low recovery is likely due to the enamel on the tusk being much thinner than on the tooth, and therefore not providing as much protection to the proteins inside. However,

as little comparative work has been done between tusk and tooth enamel, nor for preservation differences due to enamel thickness, this is only a hypothesis at this time.

Lines 46 and 47: There is no any fossil record to support these two divergence times, and the existing fossil records are much later than these two times.

It is true that the existing fossil records are later than these two times, but our estimated divergence times actually align more with the fossil record than previous divergence time estimates. Our divergence estimates are younger than previous estimates using morphology or molecular data. Accordingly, we believe our tree topology is our best hypothesis given all of the data at hand, now that proteomic and genomic data is available for some of these extinct clades, complemented by morphological calibrations.

Line 58: Should “beyond” be “before”? On the other hand, the total-evidence phylogenetic analysis of deep-time fossils widely used in paleontology incorporates genetic data rather than just morphological observations.

This has been changed accordingly from “beyond” to “before”

Line 71: Here, the authors ignore another group of Rhinocerotidae, namely Aceratheriinae, which is more basal than the sister group composed of the Rhinocerotinae and elasmotheriinae. This deep ‘basal split’ at least occurred between the Aceratheriinae and the Rhinocerotinae-Elasmotheriinae.

We agree with the reviewer that a deep basal split is probably more likely to have characterised the divergence between Aceratheriinae and the Rhinocerotinae-Elasmotheriinae, following Lu et al. (2023), contra Tissier et al. (2021). Unfortunately we are unable to test the relationship between Aceratheriinae and Elasmotheriinae with our dataset. In the future, as palaeoproteomic methods improve, sequencing the proteomes of Pliocene *Teleoceras*, a late-surviving aceratheriine from the lower latitudes of North America, or *Shansirhinus*, another late-surviving aceratheriine from the Linxia Basin deposits mentioned in supplementary, should present a solution to this specific problem, and such data could be integrated into our dataset when the time comes. We have added a reference to Aceratheriinae now on Line 86, referencing a major clade for which genetic sequence data is currently unavailable, and then we follow that up with additional text in the discussion/conclusion on the prospect of retrieving protein sequences from the clade (Line 233-235).

Line 72: Although the Rhinocerotidae appeared in the Late Eocene, they cannot be described as “bursts” because fossil records are quite rare.

This has been changed to “episodes of diversification”

Line 78: Should “relatively-young” be “relatively-old”? However, these two genera, Elasmotherium and Coelodonta, did not appear until the Late Miocene and Pliocene. When using genetic data to calculate, there must be at least some clues from fossils. If this split age has not been confirmed by fossils, it is not appropriate to use the age as evidence to cite other analysis results.

This phrase has been removed to be more accurate.

Line 89: In the Miocene and Pliocene, the Arctic region was not in the same frozen environment as today. Although the current permafrost environment of this locality favors protein preservation, the temperatures varies from 20 Ma to the present. The temperature here was high until the Pliocene (Brigham-Grette et al., 2013, Science; Ballantyne et al., 2010, Geology; Csank et al., 2011, Earth Planet. Sci. Lett.). In this manuscript (Lines 365-367), the authors also indicate “In the early Miocene, the Haughton crater lake and its surrounding environs experienced a significantly warmer annual temperature (8-12°C) than the present day). How can authors ensure the biomolecular preservation of Early Miocene fossils in a favourable temperature? Could authors do more evaluation based on the climate change? I notice authors have prepared some experimentally heating samples, have they conducted protein extraction on these samples? It may also help validate the reliability of the results. Thus I suggest the authors discuss more about the possibility of enamel protein preservation to this deep time.

This region would indeed have been warmer in the past, and so we have reworded the relevant text to make clearer that because of the relatively high latitudes, preservation is likely to be more favourable compared to other regions of the globe of similar age. While not directly requested in this comment, we have also added thermal age calculations (see comments below). The exceptional preservation of tooth enamel (and eggshell) peptides seems to be linked to their composite structure, with certain portions of the peptide sequences stabilised by interaction with the mineral (Demarchi et al. 2016). This is an area of active research.

Protein extraction studies have been done on younger (~80,000 ka) woolly rhino (*Coelodonta antiquitatis*) samples by the University of York team. These samples were heated at various temperatures (200-400˚C) and durations (3, 5, 10, 15 min) to

artificially degrade them. Analysis of the peptide sequences was undertaken on a different LC-MS/MS system to the one used for the Arctic rhino sample (Orbitrap Fusion Tribrid™mass spectrometer (Thermo Fisher Scientific, Germany) with an EasyNano ionisation source (Thermo Fisher Scientific, Germany)). We successfully retrieved very similar regions of two of the major enamel proteins (AMELX and ENAM) from these enamel samples that had been heated at 300˚C for 10 min. This information is now included in the manuscript.

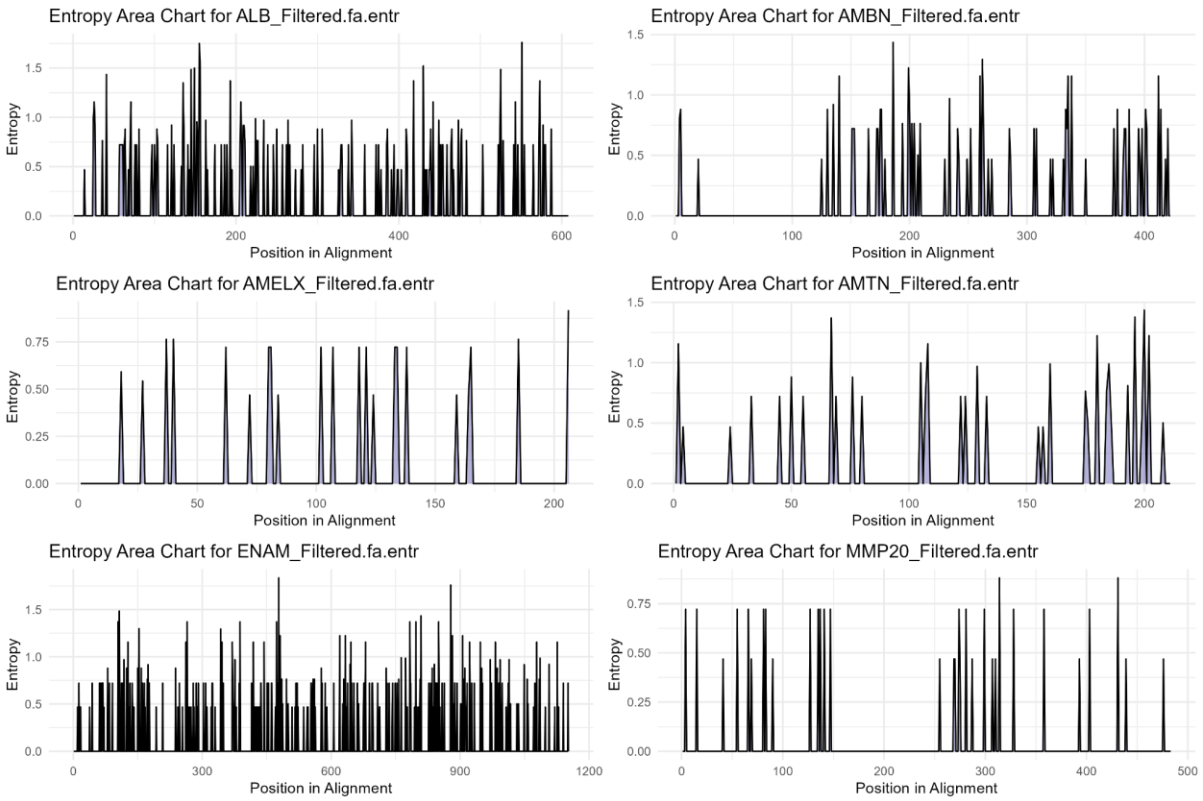
The levels of intra-crystalline protein decomposition, including amino acid racemisation, were found to be similar between the artificially heated woolly rhino samples and the Arctic rhino sample. These additional analyses thus support the reliability of the proteomic results for Arctic Miocene rhino, and we have added this information to the manuscript (Line 151-156, Ext Figure 7, and Supplementary Info).

Line 94: Why does this tooth belong to the genus Epiacaratherium? There is no description or comparison.

The morphology is described in detail in the preprint manuscript "*Post Eocene rhinocerotid dispersals via the North Atlantic"*, which we cite, in the current version of the manuscript, as ref. 36, specifically in its supplementary information appendix (Page 9), where the dentition is figured alongside the rest of the specimen. We realize we probably did not make it so clear in our initial submission that the detailed morphological description and diagnosis is present in "*Post Eocene rhinocerotid dispersals via the North Atlantic"*, so we have added additional citations to reinforce this. Also, we now have higher-resolution photos in the Appendix, including one photo with discernible morphological features, all with legends documenting morphological features of interest.

Line 97: In the paper "Phylogenetic signal in primate tooth enamel proteins and its relevance for paleoproteomics" of this research group, it mentions that for 1-2Ma samples with simulated 1139 amino acids (aa), the phylogenetic placements of most nodes at family level are consistent with the reference species tree. But for older and more degraded samples ranging from 5-10Ma (simulated enamel proteome of 593aa-98aa), the resolution and support of the phylogenetic tree are relatively low. This old Rhinocerotidae proteome reveals only 251 aa, maybe its application on phylogenetic analysis need more evaluation. As observed in their constructed phylogenetic tree, the posterior probabilities for some nodes are not high. Could they perform a similar analysis of sequence conservation and phylogenetic signal of all available proteins of the relevant rhinoceros species?

This was a very useful recommendation for us, and we completed both analyses requested (sequence conservation and phylogenetic signal). These methods and results for the phylogenetic signal analysis are now in the supplementary document. The sequence conservation analysis is not particularly informative due to the relatively-low number of taxonomic units in our analysis compared to that of Zazueta et al. (2024; doi:10.1101/2024.02.28.580462), due to the relative lack of whole genome assemblies available for ceratomorphs compared to primates. As such, we include the plots here in this response, and it is evident that the high number of sequence positions without polymorphisms in our alignment precludes the usefulness of this analysis for our dataset, and in fact, produces plots that are misleading.



On the other hand, our phylogenetic analyses of abridged datasets, trimmed to only those regions recovered in the arctic rhino, confirm the phylogenetic utility of these experimentally-recovered protein sequences, particularly with regards to the placement of the arctic rhino. The ‘minimally-filtered’ tree is very similar in topology and divergence time estimates to the original phylogenetic tree in our manuscript, while even the ‘strictly-filtered’ tree retains the same placement of the ancient arctic rhino, and remains informative at higher taxonomic levels (excepting the poorly-supported positions of *Elasmotherium* and the Dmanisi *Stephanorhinus*), though with lower posterior probability values supporting most nodes. Generally, the subsetting of

data appears to obscure the placement of taxa/clades that display overlapping divergence time estimates, and collapse or reposition nodes with lower posterior probability support values. The results from these analyses are now available in the supplementary.

Overall, we remain confident in the topologies and divergence time estimates produced with our dataset, in spite of inconsistencies that arise in the tree based on subsetting of the 'strictly-filtered' dataset. This follows the recommendations of Zazueta et al. (2024), who acknowledge that phylogenetic topologies remain accurate, even at high 'fragmentation'/degradation stages, if only a limited number of taxa in the tree are represented by these 'fragmentation stages'. Accordingly, we have decided to keep the original phylogenetic tree in the main portion of the manuscript. Nevertheless, this does raise concern if we envision a scenario in which the majority of the taxa in a protein-sequence-based tree are represented only by ancient protein sequences, as could be expected in extinct clades for which aDNA is unavailable. Moving forward, this is a major goal of palaeoproteomics - to resolve phylogenetic trees of extinct clades using ancient protein sequences - though caution clearly must be taken in interpreting trees based only on ancient protein sequences, when that scenario arises.

Line 104: In addition to comparing the enamel proteomes of the two rhinocerotids, it is also important to compare the proteomes of other fossil taxa.

In order to limit confounding variables, we focussed these comparisons on rhinocerotids, as they are more similar in enamel thickness, general tooth morphology, and amino acid sequence, which are all factors that affect degradation to a certain degree. In terms of previously-published enamel proteomes, it is difficult to directly compare the proteomes of our analysis with other fossil taxa within a quantitative or statistical framework, as the identification of ancient peptides is at least somewhat influenced by search strategy, choice of peptide identification software, and availability of reference sequences. The latter is particularly relevant in our decision to focus on rhinocerotids here, as it is not clear how precise such comparisons can be at the current moment. Nevertheless, the point to compare proteomes across disparately-related fossil taxa is very important, and so we do qualitatively add a line of discussion concerning comparison of deamidation levels (Line 139), and we hope this will become an important discussion point in the future once palaeoproteomic analysis of ancient enamel becomes more standardized.

Line 115: While the authors have employed some strategies to manually inspect the reliability of all the spectra, why they do not apply the criteria used in previous studies to validate the SAPs, which requires two spectra, and at least 1 PSM identifies the SAP using both y and b fragment ions. That is a stricter strategy.

> This sentence on albumin just refers to additional evidence supporting their ancientness/authenticity. More generally, the methods section documents the stricter strategy that we employed, specifically for the filtered dataset (lines 539-552), which is essentially identical to that outlined by Coutu et al. (2021), and following principles set out in Cappellini et al. (2019) and described by Taurozzi et al. (2024). The latter recommends a complete or near-complete b- or y-ion series surrounding the SAP, ideally covered by multiple peptides. In fact, we take this a step further, and mark amino acids (not just SAPs) as missing (X) if that position is not covered by at least a y- or b-ion, in at least one spectrum, following Coutu et al. (2021). To our knowledge, no peer-reviewed publication has suggested a SAP be covered by both y- and b-fragment ions.
>
> We must also say that we elected not to recommend an even stricter strategy, as it is becoming less clear that such strict strategies may actually not be beneficial for sequence reconstruction. Biased fragmentation patterns (e.g., the proline effect) can cause a lack of fragmentation in certain regions of peptides (e.g., y- or b- ions at N-term might be unidentified), a problem which may be exacerbated in low-abundance peptides in ancient samples that carry many post-translational modifications that could potentially exacerbate these fragmentation biases. Furthermore, recommendations of a stricter strategy were inspired by the “two peptide rule” (e.g., protein ID requiring two non-overlapping peptides), and this guideline itself refers to protein identification, not peptide identification. Consequently, we adhere to the stricter search strategy for our analysis, due to the potential controversy that may be anticipated with identifying such ancient peptides, but we do not further discuss if it is necessary or not.

Line 162: In recent years, many phylogenetic analyses in paleontology have attempted to use total-evidence analyses. The rhinoceros at this location has a nearly complete skeleton, why not combine morphological and molecular data for a total-evidence phylogenetic analysis?

> This is a fair point, and it is something we considered, but we decided to not do so, following Presslee et al. (2019; doi:10.1038/s41559-019-0909-z), who reported that

the high amount of variance in coded morphological characters swamps out the signal from the quite conserved protein sequences. Generally, such a total-evidence approach will be optimal once the methods underlying phylogenetic analyses of morphological data become more standardized.

Line 170: The Elasmotheriinae and Rhinocerotinae form a sisters group, so Elasmotherium sibiricum, a species in the subfamily Elasmotheriinae, cannot have a direct divergence with the Rhinocerotina, a tribe of the subfamily Rhinocerotinae.

This has been corrected.

Line 316: Should “after destructive palaeoproteomic analysis” be “before destructive palaeoproteomic analysis”?

The highest resolution photos we have of the tooth fragment, used for this figure, are after destructive analysis. We have added an additional photo in the supplementary of the fragment before analysis, in addition to more photos post-analysis.

Line 721: Extended Figure 1 showed the different PTM profiles for CMNF-59632, compared to enamel proteomes from other ancient rhinos and a mediaeval ovicaprine. While the CMNF-59632 showed obvious higher incidence of PTMs related to oxidative degradation, the proportion of different type of PTMs of tryptophan or histidine in these samples also varies a lot. Do the authors have any interpretation for this phenomenon?

At this point, when comparing so few samples, we cannot feel confident in any interpretation on proportions of different oxidative PTMs. There is no doubt that differences in burial conditions between the sites (temperature, pH, humidity, etc), as well as different conditions after recovery (such as exposure to UV), could have a role in these ratios. However, we would not want to speculate on the exact causes of these trends without further research on more samples. This figure (now Ext. Figure 3) was provided to illustrate that, despite the kinds of degradation, CMNF-59632 has more evidence of oxidative damage than the younger specimens (as you acknowledge), which helps to ensure that we are basing our interpretations on endogenous proteins. We do acknowledge that a better understanding of the differences in the proportions of damage products in these comparatively very ancient samples is important to understand the degradation of enamel proteins and encourage more work on this in the future.

Referee #2 (Remarks to the Author):

The authors report palaeoprotoemics data on ancient enamel samples and they have done thorough analysis for assessing the evolutionary history of rhinocerotid specimens. The significance of the paper are the ancient samples (which are around 21-24 Ma) and good preservation of proteins within, that allowed to analyse evolutionary history of rhinocerotids. It is a well written paper and in general, of good quality (in terms of approach and presentation). The proteomic analyses are performed well, with thorough data analysis, peptide assignment and filtration (I cannot evaluate the phylogenetic analyses well as this falls out of my scope of expertise). The results are robust and will be of great interest to the scientific community.

> We thank Referee #2 for their comments and suggestions. Particularly, they encouraged us to more appropriately contextualise our findings and their potential implications for palaeontologists.

I have some questions, thoughts and suggestions for possible improvements of the paper.

I would be keen to know more about what future implications the work gives. I understand that proteomic analysis of such ancient samples will allow to dig the evolutionary history of different samples even further, but is there more? Or maybe, I have missed it in the paper (given that I'm not an expert in this?). Perhaps it would be good to highlight main (or most interesting) message more clearly in the paper.

> We have expanded the concluding paragraph, to more directly discuss the broader implications, and better contextualise our findings in the light of the emerging field of 'biomolecular palaeontology'. We also have added a supplementary which further adds a broader context to the analyses we undertook, and their broader implications.

For palaeoproteomic analysis:

I wonder if all the enamel proteins are already cleaved in-vivo (or due to diagenesis)? Would there be a point in doing a workflow where enamel peptides are collected in parallel to digesting remaining proteins with trypsin? Perhaps authors have tried this?

> The first sentence is correct - the enamel proteins are cleaved *in vivo* and peptides from ancient enamel tend to be too short for further enzymatic digestion, due to

further cleavage by diagenesis-related degradation, as reported in Cappellini et al. (2019), who attempted ancient protein extraction from enamel (and bone and dentine) with both tryptic digestion and no digestion. Protocols with enzymatic digestions only allowed for the sporadic recovery of collagen fragments in a small proportion of samples (6/19), while the digestion-free approach permitted the recovery of enamel proteomes in 13/14 specimens. Nevertheless, this is an important point, and while we have no further discussion to add to the manuscript, enzymatic digestion remains a suitable strategy for even very ancient specimens, though more work into which enzymes would be necessary for such short, highly-degraded peptides is necessary.

Also, what is the possibility of differences when comparing diagenesis related ptms, recovered peptides with different amino acid lengths and sequence coverages between samples (Figure 2) given that samples were analysed with different instruments (exploris 480 and q-exactive). Perhaps it would be good to discuss this (using the same instrument for all the samples would be an ideal case scenario, but often its not reality. But I think it would be useful to briefly mention possible differences).

We agree that using the same instrument for all the samples would be the ideal, but due to upgrades in the facility over time, this was not possible. Between the two new samples analysed here, one with Exploris 480 and one with Q-Exactive HF-X, there should be minimal differences with recovery when operated with the same methods, as the main differences between the two machines are based on the Exploris' improved architecture to enable a more compact and robust platform but with the same analyser set up as its predecessor the HF-X (Bekker-Jensen et al. 2020, doi: 10.1074/mcp.TIR119.001906). Analysis by Thermo Fisher itself suggests that the two models perform comparatively (Arrey et al. 2019, ASMS poster available here: https://assets.thermofisher.com/TFS-Assets/CMD/posters/po-65547-evaluation-modified-quadrupole-orbitrap-ms-asms2019-po65547-en.pdf).

There could be possible differences between these newly analysed samples and the previously published Dmanisi rhinocerotid and mediaeval caprine samples which were analysed with a Q-Exactive HF instead. The HF-X had improvements in scanning speeds and sensitivity that allowed for improved peptide and protein identifications, although this improvement was much more obvious for shorter gradients (Kelstrup et al. 2018, doi: 10.1021/acs.jproteome.7b00602). All samples used in this work were run with the same methodology as those previously published, except for a slight increase in maximum injection time for the HF-X/Exploris due to

the faster scanning capabilities. Because this is not considered a ‘short’ gradient at 77 mins, the improvement would not be as noticeable but we concede it could still have an impact on the comparisons between the two sample sets. That is one of the reasons we use normalised data for the comparison of PTMs, i.e. looking at the percentage of modifications overall, rather than just counts or at specific sites. However, to your point, we have added a line at 508-510 that says: ‘While it is possible that some inter-sample variation between CMNF-59632 and these samples is caused by analysis by different MS models, according to benchmarking studies (Bekker-Jensen et al. 2020; Kelstrup et al. 2018), it is unlikely that the overall trends and interpretations would change.’

Minor suggestion: when referring to citation would be to include the name of the main author, followed by the number of the publication in the bibliography (this is done in many cases, but not everywhere). For example, sentence “The remaining steps generally follow that outlined by 9”; would read much better : “The remaining steps generally follow that outlined by Cappellini and co-authors (2019) 9. (such case was repeated several times in the manuscript, I suggest checking).

The references to citations have changed to match this proposed style.

Referee #3 (Remarks to the Author):

In this manuscript, the authors extract proteins from the enamel of a now-extinct Rhinocerotidae specimen estimated to be more than 20 million years old. Seven proteins were recovered and partially sequenced from the specimen CMNF-59632. These were used to reconstruct the phylogeny for this specimen, providing new insights into the split between Elasmotheriinae and Rhinocerotinae.

This study represents a novel contribution to the field and is the oldest protein extraction from a specimen with sufficient preservation for phylogenetic placement. This is an exciting result for the field as a whole and represents a significant discovery. I look forward to it being published.

The manuscript is well-written and appropriately cited. Sections of the manuscript could be improved with greater specificity or with a supplemental (as discussed below) however the methods and data analysis appear very sound.

My major concern surrounds the discussion of observed PTMs in the specimen. Given the reliance on PTMs to establish that the proteins are endogenous and that the enamel matrix has preserved its closed system, their reporting lacks specificity. Peptide lengths are reported as being “shorter” and PTMs related to diagenesis are reported quite vaguely. In the figures (especially Fig 2) it is very difficult to estimate counts or percentages and the raw numbers are not reported in the figure captions or the additional information. The reader would be required to look through the raw data and reconstruct this information themselves or guess based on the figures provided – both unnecessary as the authors must have collated this information to produce the figures and to make the statements in the text.

Given the surprising age of the fossil and the use of PTMs to support the endogeneity of the proteins, a greater degree of information should be included in the manuscript. A table or supplement with these results would be an invaluable addition. More information should also be added to the figure captions.

> We thank the reviewer for pointing out this oversight on our part. We have added a supplement that includes the requested tables with the raw data on specific PTM accumulation by sample. Here, we have also added a plot for site-specific deamidation, and additional details on these post-translational modifications more generally. We have also amended the figure captions in the main text. Overall, while the PTMs support endogeneity, we also would like to highlight that the amino acid analysis also supports the endogeneity, and we have also added additional details for that analysis to the supplement.

Additional comments:

Lines 84-92: does this section imply that other fauna were also analysed? The paper presents only a single Rhinocerotidae specimen (tooth and tusk). If attempts to extract proteins from other specimens were unsuccessful this would be an important inclusion to the paper or the supplemental. As the authors discuss in the Introduction, there is no systematic understanding of protein preservation in deep time. It remains unclear if this Rhino tooth is an exceptional find or if we should expect biomolecular preservation in other enamel samples from similar regions and time periods.

> The only specimen we have sampled for ancient proteins from this site was the rhinocerotid, so we do not have any more data to make firmer conclusions on these points. However, we have also just added estimations of thermal age of the

> Haughton Crater Rhinocerotidae specimen, and added interpretation and discussion of these results. While the preservation of the sample could be exceptional, we suspect fossil enamel from similar regions and time periods may also provide similar results, in the light of our thermal age calculations. We have added two sentences (Line 97-102) to better describe why we targeted dental enamel.

The image of the specimen CMNF-59632 (Figure 1) makes it difficult (impossible) to do any kind of visual inspection of the fossil. High-quality images of the fossil should be supplied either by enlarging the image in this figure or by providing better quality images in the SI. Additionally, no information on the attributes, taphonomy, or morphology of the fossil are included in the manuscript. The authors state that the “size and morphology” are characteristic of rhinocerotids but this information is not included with the exception of the banding on the enamel surface.

> We have taken additional images of the specimen, and included them in the supplementary document. We have also added additional information on the specific morphological attributes of the fossil that permit its assignment to Rhinocerotidae, included in the photo legends in the appendix (particularly Photo 2), complementing the information related to the fossil and the context it derives from, that are present in Fraser et al. (2024).

Congratulations to the authors on this very successful manuscript.

No Referee #4

Referee #5 (Remarks to the Author):

Referee Report: Nature

A 20+ Ma old enamel proteome from Canada’s High 1 Arctic reveals diversification of Rhinocerotidae in the 2 middle Eocene-Oligocene (Ryan Paterson et al.)

Summary of the key results

This study has recovered protein sequences from an exceptionally preserved Miocene Rhinocerotid enamel from the polar Canadian High Arctic. The authors satisfactorily demonstrate the authenticity of the proteins and present some discussion on their preservation state and diagenetic alterations. The authors then proceeded to utilize this proteomic data in a phylogenetic analysis to place the divergence of Epiaceratherium sp. (the current sample) to the middle Eocene-Oligocene. This finding contradicts previous estimates of a basal split between Elasmotheriinae and Rhinocerotinae and also supports a Eurasian origin of this clade.

> We thank Referee #5 for their comments and suggestions. We initially neglected to consider how our findings may fit within the broader realm of 'biomolecular palaeontology', but the suggestions here prompted us to recognize the importance of this. So, we now have additional sections, methods, and discussion points considering diagenesis.

Originality and significance: if not novel, please include reference

The study is original in successfully applying robust paleoproteomic data to phylogenetic analysis on Miocene specimens. The significance of this study is in providing other paleoproteomic and molecular taphonomic researchers a targeted environment to look for exceptionally preserved fossils capable of yielding such data and as a model in application of paleoproteomic data in phylogenetic analyses beyond the Pleistocene.

Data & methodology: validity of approach, quality of data, quality of presentation

The methodology is sound, and the data appear robust. I would have liked to see an additional method for the determination of degree of diagenesis such FTIR or Py-GC-MS, however, given the totatlity of the data presented, it is my opinion that the evidence is sufficient for protein authenticity. The figures are sufficient.

This is a very good point. We decided not to include these additional methods because we focused exclusively on supporting the authenticity and endogeneity of the peptide sequences at the base of our phylogenetic analysis. Tracking the degree of ancient biomolecule diagenesis falls instead beyond the goals of this project and would have required the sacrifice of more of the very limited amount of available starting material from such a precious specimen. Overall, we very much appreciate this comment, as, in conjunction with other comments in this and the other reviews, it prompted us to better elaborate on how our findings fit within the existing paradigms of “ancient protein research”. To address this general concern we added the thermal age analysis and recommend future directions that do integrate this holistic approach to understanding the diagenetic pathways of a sample.

Appropriate use of statistics and treatment of uncertainties

Appears to be appropriate

Conclusions: robustness, validity, reliability

Based on the data presented, the conclusions seem valid. I am sufficiently satisfied regarding the reliability of the methodology with regards to protein recovery and authentication.

Suggested improvements: experiments, data for possible revision

While the authors have discussed the previous reports of protein sequencing, I also suggest adding a section on the general nature of protein preservation, particularly when it comes to the topic of deep time (although, Miocene may generally be considered rather young to employ the term “deep time”). The traditional conventions for the preservation of proteinaceous material in fossils have continued to be revised. Certainly, they are often altered to the point that sequence information is impossible to retrieve, however the vestiges of original proteinaceous character may remain. I suggest some references below. Similarly, the authors do not acknowledge the work of Stolarski et al., 2023, who have also recovered protein sequence data from Miocene otoliths.

This was a very important point in framing our manuscript, and we have amended the introduction and the conclusion to put our work in the context of these works on

deep-time proteinaceous moieties. We have included all of the suggested references (bar one for space limitations).

As mentioned, I would also have liked to have seen more discussion about the nature of the diagenetic alteration of the proteins, and its relation to the depositional environment and type of bone. Could they expand, for example, a little more on Figure 2a? Could they also explain in more detail why dental enamel would work better for such a study than say, femur bone? Similarly, I would have liked to have seen a reporting of thermal age calculations (cf. Demarchi et al., 2016, eLife). To what point was the diegenetic alteration retarded? Apart from this, I have raised a few minor comments below.

We have now included thermal age calculations in the manuscript. We improved upon previous thermal age estimates, which were based on global models of climate change, and which were not adequate for timescales beyond the Plio-Pleistocene and for sites at high latitudes. We used the variant HadCM3BL-M2.1aD of the HadCM3 Global Circulation Model (spatial resolution of 3.75 x 2.5 arc-degrees in longitude and latitude respectively, run for boundary conditions describing the earth system every 5 million years for the last 500 million years) and derived monthly paleotemperature data over the last 40 Ma. Wide seasonal fluctuations at the Haughton Crater site result in a high standard deviation of the calculated thermal age, which is however perfectly in line with the hypothesis that protein survival is limited to Pliocene/Late Miocene timescales in temperate environments while it can extend to beyond the Miocene in the High Arctic.

To the other points, mounting evidence suggests that highly-biomineralised tissues, like enamel and eggshell, display closed-system behaviour, which may confer protection of peptides. Cappellini et al. (2019) show that proteins present in bone, of an abundance sufficient for phylogenetic analysis, do not persist as deeply in time as proteins present in dental enamel. We now more directly refer to this closed-system behaviour in Line 112 and Line 152, as well as the unique structure of enamel (Lines 97-102), and also add discussion points specific to the highly-biomineralized nature of enamel to the supplementary.

References: appropriate credit to previous work?

While the authors have used appropriate citations for the most part, they have omitted a reference to Stolarski et al., 2023, which I believe is important to discuss. That was also a paper that showed compelling evidence for the recovery of protein sequences from Miocene

fish otoliths (~ 14 Ma). They have also shown how the mineral phase relates to the peptide preservation. That paper does not use it for phylogenetic analysis but in its context, while the current study does indeed push back the age for confirmed protein sequence recovery, there is a difference between pushing it back from 3.7 Ma, as written by the authors, to pushing it back from 14 Ma. Similarly, please see some references related to the recent advances on the knowledge relating to diagenetic alteration of proteins in deep time.

Stolarski et al., 2023. First paleoproteome study of fossil fish otoliths and the pristine preservation of the biomineral crystal host. Scientific Reports 13:3822

Dutta et al., 2020. Chemical evidence of preserved collagen in 54-million-year-old fish vertebrae. Palaeontology 63, 195 – 202.

Umamaheswaran et al., 2023. The diagenetic fate of collagen as revealed by analytical pyrolysis of fossil fish scales from deep time. Geobiology 21, 378 – 389.

Boskovic et al., 2021. Structural and protein preservation in fossil whale bones from the Pisco

formation (Middle-Upper Miocene), Peru. PALAIOS 36, 155 – 164.

Alfonso-Rojas and Cadena, 2020. Exceptionally preserved ‘skin’ in an Early Cretaceous fish from Colombia. PeerJ 8:e9479

Wiemann et al., 2018. Fossilization transforms vertebrate hard tissue proteins into N-heterocyclic polymers. Nature Communications 9:4741.

> We have added these citations (excepting Alfonso-Rojas and Cadena, 2020; only removed to fulfill citation count limitations), and integrated them into expanded paragraphs (primarily Lines 237-251) related to the study of proteinaceous products

via diagenesis, and the implications of our research within this broader 'biomolecular palaeontology' context.

Clarity and context: lucidity of abstract/summary, appropriateness of abstract, introduction and conclusions

The abstract is clear regarding the content of the manuscript. Introduction and conclusions are appropriate, except for the discussion of protein diagenesis in deep time, which I have mentioned above.

Comments:

Abstract:

Pg 2 line 37: While it may be technically correct, the phrasing makes it seem that protein sequence data has not been confirmed beyond 3.7 Ma. I suggest phrasing that makes it more clear that such data had been recovered (Demarchi et al., 2022, eLife, which you have also cited elsewhere; and Stolarski et al., 2023, Scientific Reports) although perhaps not phylogenetically applicable.

We have modified this sentence accordingly: "Still, **while ancient proteins have recently been reported from the Middle-Late Miocene,** the recovery of protein sequences providing sub-ordinal level phylogenetic insights does not exceed 3.7 Ma (Pliocene)."

Manuscript:

Pg 2 line 61: Please see Stolarski et al., 2023, Scientific Reports https://doi.org/10.1038/s41598-023-30537-8. They have reported paleoproteomic data from fish otoliths from the Mid-Miocene (14 Ma).

We have added in this reference and refer now to "Middle" Miocene survival.

Pg 3 line 67-68: This line is somewhat confusing. Do you mean to say that poor preservation record leads to a bias in inferring evolutionary relationships? Please rephrase

This line has been removed.

Pg 4 line 88: The geological age is advanced from the perspective of protein preservation potential, while most geologists would consider the fossils relatively young. I would recommend making this distinction.

This paragraph has been modified to avoid calling the geological age “advanced”.

Pg 4 line 89-90: When did the temperature drop initiate? At what timescale did permafrost occur? Since the depositional environment was at relatively higher temperatures, it is likely that diagenetic alteration would have commenced to some degree, hence I would have liked to have seen some comparison of thermal ages to elucidate it.

We agree with the Reviewer and have now provided calculations of thermal ages for the Haughton Crater across the past 40 Ma, in comparison to two other sites from which ancient proteins have been recovered from the Pliocene/Miocene, based on location-specific past temperature estimates. These results are in line with the hypothesis that this fossil from >20 Ma Haughton Formation, where mean annual temperatures were below 0°C for all time steps considered in the model (0, 3, 6, 11, 15 and 20 Ma), experienced arrested/slow diagenesis during most of the year throughout its history. While it is difficult to establish the precise onset of permafrost conditions using this model, the calculated thermal age of 2.5 Ma@10°C indicates that the expected extent of protein preservation is similar to that of a sample of Pleistocene age but recovered from environments where the mean annual temperature is 10°C.

Pg 4 line 101: Please avoid the term “surprising”. I would recommend replacing it with a less stronger term, such as “unusual”, especially as the deposits are not so geologically ancient, and proteomic data has previously been reported from Miocene specimens.

This has been changed accordingly.

Pg 9 line 211: Please change “well beyond” to “beyond”

This has been changed accordingly.

(Methods section)

Pg 17 line 374-375: I would like to see a discussion about enamel in the main section. It is important, I think that such a preservation was not found in the tusk of the specimen, but the main discussion appears to only focus on the frigidity of the locality. However, a discussion on suitable fossil types (for example, eggshell, or fish otoliths), perhaps even from the same organism, that can be targeted for such analyses will be very useful to future researchers.

> We have more directly referenced the closed-system behaviour of enamel, which makes it an optimal target for deep-time palaeoproteomics. We have also added two sentences describing why we targeted enamel specifically, alluding to its unique microstructure and high degree of biomineralization (Lines 97-102). We do not go into significantly more detail on this point, because while we can now recognize the potential for dental enamel in preserving ancient proteins from challenging geological contexts, we have not sampled a broad enough diversity of sample types to know if certain tissue types are unsuitable. Specifically, while it is currently clear that most vertebrate bone is not suitable for deep-time palaeoproteomics, except in exceptional preservational contexts, we still don't know how bone density exactly affects ancient protein preservation, or whether there is any bone district that is particularly favourable to ancient protein preservation, similarly to what observed for the petrous part of the temporal bone for aDNA. For this purpose, perhaps the auditory bullae of fossil cetaceans would be a suitable target to test, or the very dense long-bones of some semi-aquatic vertebrates. On a similar note, while ostrich eggshell is a suitable target, ostrich eggshell is particularly thick compared to the eggshell of most other avian extant and extinct taxa. Similar variability in thickness/thinness of enamel exists across mammals, and so we err on the side of caution here.

Response to Referees

Referee #1 (Remarks to the Author):

Although the authors agree that a deep basal split is probably more likely to have characterised the divergence between Aceratheriinae and the Rhinocerotinae-Elasmotheriinae, there is a claim that they are unable to test the relationship between Aceratheriinae and Elasmotheriinae with their palaeoproteomic dataset. In fact, combining morphological and molecular data for a total-evidence phylogenetic analysis is a good way to solve this problem.

Response: We thank the Referee for their constructive comments. Following up on this specific point, it is reassuring for us to read that the Referee suggests us to combine morphological and palaeoproteomic data in a total-evidence phylogenetic analysis, because, in a way, it looks like they agree with us when we claim that we cannot test the relationship between Aceratheriinae and Elasmotheriinae by using palaeoproteomic data alone. At the same time though, the reliability of the total-evidence phylogenetic analysis the Referee suggests us to adopt has not been systematically tested when this approach relies on palaeoproteomic data. More specifically, while we agree that, in theory, a total evidence analysis could be a possible way to solve the relationship between the two taxa, in practice, so far there is no tested and robust workflow to reliably deliver a total-evidence phylogenetic analysis that uses ancient protein sequence data.

To our knowledge, a total-evidence analysis incorporating protein sequence data and morphological characters has never been published in a peer-reviewed journal. Interestingly, Presslee et al. (2019; doi:10.1038/s41559-019-0909-z) specifically caution against such an approach, commenting specifically on ‘homoplastic’ characters that are potentially pervasive in morphological analyses. They remark: “While combined analysis of morphological and molecular data will ultimately be necessary to fully resolve folivoran phylogeny, this exercise suggests that it is premature to consider such simultaneous analyses reliable at this point in time”. Such caution is due to the overwhelming amount of ‘parsimony-informative’ morphological characters swamping out the more conserved palaeoproteomic signal. Curiously, this is opposite of a common informal criticism to total evidence analyses where the molecular component, generally DNA, swamps out the morphological signal, making the morphological characters relatively less impactful. For all these reasons, the moderate variability of our ancient protein sequences would most probably lead a total evidence analysis to conclusions identical to those obtained analysing the morphological characters alone. On the contrary, the publication of our topology, raw data, reconstructed sequences, and phylogenetic scripts will allow for any researchers to explore such approaches. From our side, we prefer to focus on the recovery of phylogenetically-informative protein sequences from Aceratheriinae to attempt a direct molecular comparison.

Advances in establishing best practices for designing and sharing the coding of morphological characters will hopefully lessen the difficulty of selecting the most appropriate character matrices to be integrated into such a total evidence analysis. For all these reasons, we believe that in the future it will be necessary to develop methodologies that can reliably integrate ancient protein sequence data and morphological data into a common phylogenetic analysis. In our opinion though, this effort falls beyond the scope of this manuscript.

Line 539-552: The authors employ two filter strategies to construct two datasets: a minimally-filtered dataset and a strictly-filtered dataset. However, the manually inspection part is not clearly explained, I suggest the authors include the details explained in the response in the method section. Thus, other researches can easily reproduce their results and conduct comparative analysis.

Response: It is true that the 'manual inspection' component provides room for ambiguity. Accordingly, we removed this sentence, and replaced it with "Next, ion series coverage is examined for each PSM". Following this sentence, we leave the description of filtering sequences in a reproducible manner based on ion series coverage. We acknowledge that the modified text, and our 'manual inspection' methods still may leave some room for operator bias, as at least a modicum of expertise and/or experience is recommended.

I acknowledge the existence of biased fragmentation patterns for different peptides. However, I still think the confirmation of SAPs should be stricter than other amino acid positions because the SAPs usually influence the phylogeny more. Regarding the SAP confirmation, Presslee et al. (2019; doi:10.1038/s41559-019-0909-z) have suggested that "a higher confidence was given to spectra that identified the selected amino acid using both b and y ions". But as the authors mentioned in the response, we still do not know how strict the criteria should be to balance the resolution and accuracy.

We would like to thank the Referee for directing us to the supplementary text in Presslee et al (2019) that is directly relevant here. However, the phrasing in there looks to us a bit vague. More specifically, it is not clear to us what a "higher confidence" entails exactly. For example, were the SAPs identified with "higher confidence" only when *complete* y- and b-ion series cover the SAP and the surrounding amino acids? And with what minimum number of overlapping peptides?. To address the best we can this comment, and another one reported below from the same Referee, we now added, as supplementary figures (Figs S1-S5), the alignments of the overlapping peptides that cover each SAP, together with their matched fragment ion patterns. We believe this solution offers a clear and direct way to show that the identification of the accepted SAPs is confidently supported.

Maybe additional analysis on the experimentally heated samples can give some hints.

The analysis of experimentally-heated modern protein samples is useful to broadly mimic in the lab the molecular degradation that ancient proteinaceous material experienced over geological times. In particular, heating can approximate, not perfectly though, the patterns of: (i) peptide bond hydrolysis, both terminal and internal, as well as (ii) spontaneous and irreversible amino acid modification. For this reason, in the manuscript we used the results of modern protein heating, alongside the amino acid racemization results, to confirm the general consistency of the patterns of ancient protein degradation and recovery that we observe experimentally. For example, it was reassuring to observe that very similar regions of the enamel matrix protein sequences were recovered from both heated and ancient samples. On the contrary this approach is not particularly useful to validate the accuracy of SAP calls, because the introduction of amino acid substitutions affects the ionisation efficiency of each peptide carrying these SAPs and the relative intensities of the different fragment ions. Fragment ion signal intensity, or their loss, is also affected by the initial amount of each peptide, i.e. ultimately its original number of retrieved copies.

To rigorously address the concern the Referee is raising, the ideal experimental approach would require the analysis of multiple dilutions of all the synthetic peptides, with all their observed combinations of amino acid modifications, supporting the identification of all the SAPs. In practice, it is easy to imagine how this approach would quickly become prohibitively expensive and very demanding in terms of resources needed. While synthetic peptides have been used a few years ago to validate SAP discovery in palaeoproteomics (Cappellini et al. DOI:10.1038/s41586-019-1555-y), new emerging software tools based on AI-generated models for fragment ion pattern prediction, like those used in our manuscript: Prosit (Gessulat et al. DOI:10.1038/s41592-019-0426-7), or MS2Pip (Degroeve and Martens: DOI:10.1093/bioinformatics/btt544), allow for confident and more efficient matching between experimental MS/MS spectra and the corresponding predicted ones. As an example, to illustrate our point, we would like to draw attention to Figure 3. The mirror plots depicted there demonstrate the similarity of the patterns of fragment ion intensity and coverage across *both* the experimentally-derived spectrum from CMNF-59632 and the corresponding *in silico* predicted one, including the lack of c-term b-ions in this peptide sequence. Despite their reliability however, admittedly these software tools cannot deal yet with some of the amino acid modifications commonly identified in our ancient samples. Regardless, we believe that globally the amount of data we provide in our manuscript, and their interpretation to validate its robustness and authenticity, are sufficient to confidently support, beyond reasonable doubt, the reliability of the evidence at the base of our conclusions.

Figure 2d should have two y-axes: one for amino acid mutability, whose approach should be explained in the method section, and another for peptide count, whose y-axis is missing.

Response: We thank the Referee for pointing this oversight on our part, and we have amended the figure labels accordingly. We have removed the plotting of amino acid mutability.

Extended Figure 1A: when the authors refer to “Before filtering” and “After filtering”, which filter strategy do they mean, minimally-filtered or strictly-filtered? I suggest including a comparative analysis between these two filter strategies.

Response: This is a good point concerning the ambiguity here, and so we have amended both the Figure caption, to explicitly note we refer to our ‘strict filtering step’, and the main text, to clarify which filtering we are referencing (“Strictly-filtered” in this case) with additional detail on the criteria employed. The main text now reads as follows:

> “At this stage, peptide-spectrum matches are discarded if they present a match to any reasonable candidate contaminants, if they are also identified in the blank, or if they present poorly-covered ion series. The resulting peptide-spectrum matches are used to reconstruct sequences for the ‘minimally-filtered dataset’.”

Regarding a comparative analysis, Extended Figure 1 provides the amino acid coverage levels comparing our “Minimally-filtered” and “Strictly-filtered” datasets. We would also like to draw attention to Supplementary Information section 5, which includes all phylogenetic trees produced for both “Minimally-filtered” and “Strictly-filtered” datasets.

Extended Figure 8: Please include the overlapping peptides for the other 9 SAPs.

Response: We thank the Referee for pointing out this omission. It is very important for us too to share these figures, and this recommendation is not only useful here, but could very well represent a reasonable standardized criterion more broadly for palaeoproteomic studies, so the validity of reconstructed sequences can be assessed without having to rerun analyses. To fulfill this request and rectify our omission, we have added additional figures for the SAPs of interest, displaying PSMs with ion series coverage, and highlighting the SAPs (Figures S1-S5).

Referee #2 (Remarks to the Author):

The paper now reads very well. The suggestions and comments were addressed and included in the paper, where needed. Congradulations for the manuscript!

Response: We thank the Referee for their kind words, and again for recommending recontextualizing our manuscript to more directly comment on the significance of our findings.

Referee #3 (Remarks to the Author):

I thank the authors for their diligence in replying to the reviewers concerns.
I have one final comment on the draft:
While I think the manuscript is technically well prepared, I feel the authors need to be more cautious in their interpretation of these results and their implications for biomolecular preservation outside of this single fossil. The inclusion of thermal age data is interesting but statements like those found in Line 238 (the description of the formation as a palaeomolecular lagerstaette) seem like an overstatement. Especially if there has only been one fossil analysed, as stated by the authors in their reply document. One tooth with exceptional biomolecular preservation does not demonstrate a trend for an entire geological formation when it has been demonstrated in previous research that preservation can be variable between fossils from the same sites and the same geological conditions. At least it does not seem appropriate without the study of additional fossils. Certainly, this supports a larger trend that high latitude and cold regions are likely to preserve proteins in deep time, and I agree that the authors' findings should encourage further studies (as stated in their conclusion) but with such a small sample size I think more caution should be advised in the wider interpretation.

Response: We thank the Referee for their critical recommendations for our manuscript, particularly with the finer details in the first round, and also here for raising this important point. We absolutely agree, and have modified the language accordingly (Lines 236-240), encouraging further experimental work to confirm if the exceptional preservation extends across the formation, and making our statements less bold. We are also happy to not use the term "palaeomolecular lagerstaette" here, as it is perhaps premature, though we are keen on proposing such a concept, generally.

Referee #5 (Remarks to the Author):

The authors have sufficiently addressed my two main concerns from the previous draft - 1.The lack of thermal ages calculations which, in my view, was necessary in order to substantiate the

suggestions of investigating fossils from present cold climates, and 2. An appropriate discussion of the diagenetic context of the preserved proteins. The authors have also included some references that were previously excluded in the initial manuscript, but which I felt were necessary to add historical context to significance of their findings.

I recommend that the manuscript be published.

Response: We thank the Referee for the kind comments, and again for recommending a reframing of the manuscript, which we believe strengthened it considerably compared to the initial submission.