

MALDI Deamidation Score (MDS): A fast and flexible method for assessing deamidation in ZooMS data and its application to the Denisova Cave bone assemblage

Fei Yang^{a,*}, Ismael Rodríguez Palomo^{b,*}, Bharath Anila Bhuvanendran Nair^b, Samantha Brown^c

^a Institute for Archaeological Sciences, University of Tübingen, Tübingen, Germany

^b McDonald Institute for Archaeological Research, University of Cambridge, Cambridge, UK

^c National Centre for Research on Human Evolution (CENIEH), Burgos, Spain

ARTICLE INFO

Keywords:

ZooMS
MALDI-TOF
MDS (MALDI Deamidation Score)
Denisova Cave
Deamidation
Taphonomy
Site formation process

ABSTRACT

Estimating deamidation from ZooMS spectra has frequently been achieved using the q2e method due to its high-throughput capacity and ease of use. Despite this, q2e only works with spectral data in txt format, operates slowly, and employs a genetic algorithm for fitting, which is stochastic and less interpretable. Furthermore, q2e only estimates deamidation at the peptide level and does not generate a sample-wide summary. The introduction of the Parchment Glutamine Index (PQI) presented an alternative method. Initially designed for a large ZooMS dataset of parchment, it utilises weighted least squares and a linear mixed-effects model (LME) to generate a peptide deamidation estimation and a sample level index, respectively. To address the limitations of q2e and expand the applicability of PQI to a wider range of archaeological tissues and MALDI-TOF-derived data (such as ZooMS data on bone collagen), we developed the MALDI Deamidation Score (MDS), an iteration of the PQI method optimised for handling large-scale datasets. MDS is more streamlined for analysing multi-species data with customisable peptide lists, offering dramatically reduced processing time. Using the published Denisova Cave ZooMS assemblage, we demonstrate that different peptides exhibit varying deamidation patterns over time, making the use of a single peptide to represent overall deamidation potentially biased. Such information is invaluable for investigating key questions such as protein preservation and site formation processes.

Significance: We introduce the MALDI Deamidation Score (MDS), a new model that estimates deamidation at both peptide and sample levels from MALDI-derived datasets. MDS streamlines multi-species analyses with customisable peptide lists and markedly reduces processing time, enabling robust and large-scale deamidation estimation. Before MDS, deamidation in ZooMS has often been inferred from a single peptide common to Eurasian terrestrial mammals. Applied to the published Denisova Cave ZooMS assemblage, MDS reveals peptide-specific temporal deamidation patterns, showing that single-peptide proxies can be biased as peptides differ in deamidation rates and may be subject to different taphonomic processes. These results have broad relevance for studies of protein preservation and site-formation processes related to molecular taphonomy across archaeological and paleontological contexts. Looking ahead, as MDS allows a customisable peptide list, it has the potential to extend deamidation analysis of ZooMS data beyond terrestrial mammals to marine mammals and non-mammalian vertebrates (e.g., birds), and to different proteinaceous material such as keratin. It also has the potential to be compatible with different instruments used in peptide mass fingerprinting, such as MALDI-TOF, MALDI-FTICR, and TIMS-TOF. We believe that MDS will contribute greatly to our understanding of deamidation on different peptide sequences under various environmental and burial conditions.

* Corresponding authors.

E-mail addresses: fei.yang@uni-tuebingen.de (F. Yang), ismael@palaeome.org (I. Rodríguez Palomo), bharath@palaeome.org (B. Anila Bhuvanendran Nair), samantha.brown@cenieh.es (S. Brown).

<https://doi.org/10.1016/j.jprot.2025.105577>

Received 15 September 2025; Received in revised form 3 November 2025; Accepted 26 November 2025

Available online 30 November 2025

1874-3919/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Review on application of deamidation in ZooMS

Zooarchaeology by Mass Spectrometry (ZooMS) is a widely used palaeoproteomic technique for identifying the taxonomic origin of collagen-rich archaeological, historical, and palaeontological materials, like bone, dentine, ivory, leather, and manuscripts, by analysing their peptide mass fingerprints and comparing the resulting spectra with established reference libraries [1–5]. Beyond taxonomic identification, researchers have increasingly sought to extract diagenetic information from ZooMS spectra, most notably through the analysis of glutamine deamidation [6–10], a spontaneous post-translational modification that accumulates over time, known to be influenced by multiple burial conditions, including pH, temperature, or humidity [11].

Deamidation is the process through which glutamine and asparagine convert to glutamic acid and aspartic acid, respectively. There are two different mechanisms that can induce spontaneous deamidation, the first is through direct hydrolysis of the side chain and the second is through the formation of a cyclical intermediate [12–14]. The process through which asparagine converts to aspartic acid has been well studied, as it is much more sensitive to deamidation than glutamine, and therefore can be observed under shorter timescales [15–18]. From an archaeological and palaeontological perspective, glutamine is therefore a more suitable measure of protein diagenesis and has now been routinely observed in bones and teeth dating back over millions of years [19,20]. Given that taphonomy has the potential to influence the rate of deamidation, there has long been a proposition that it may be used as a molecular clock [21] with several attempts to understand its mechanisms and kinetics (see for example [12,14,22–24]).

The development of ZooMS enabled the practical application of fast and low-cost deamidation analysis for fossil specimens and artifacts, which also required taxonomic identification. This advancement was further bolstered by two methodological advancements. The first was the development of the minimally-destructive ammonium bicarbonate (AmBic) extraction protocol for type-I collagen (COL1). This protocol was specifically designed to minimize lab-induced deamidation brought on through acid-based demineralisation [25]. The second was the introduction of an isotopic envelope deconvolution method for estimating deamidation levels on a series of peptides from ZooMS spectra [18], known as the ‘q2e method’. This method employs a genetic algorithm to estimate the abundance of the deamidated and non-deamidated peptide counterparts in the mixed isotopic envelope. Within the framework of the q2e method, two collagen peptides - COL1 α 1 508–519 (m/z 1105.58, GVQGPPGPAGPR) and COL1 α 1 435–453 (m/z 1706.77, DGEAGAQGPPGPAGPAGER) - are the most consistently detected in ZooMS spectra and are broadly conserved across terrestrial mammals common in the archaeological record. Therefore, these peptides are the best candidates for deamidation estimation on large faunal assemblage with diverse species to minimize amino acid sequence induced bias [26]. However, empirical data show that COL1 α 1 435–453 is only present in about half of the spectra [27], making COL1 α 1 508–519 the primary marker in most studies focused on ZooMS analyses of fossil bone assemblages.

Despite early research indicating a correlation between the deamidation levels of COL1 α 1 508–519 and the thermal age, exceptions due to various taphonomic factors have frequently been observed [18,28]. While thermal age tries to set an equivalent thermal history by encompassing multiple environmental factors on an equivalent or normalised time scale, it still misses most of the critical factors and their overall dominance and locality. This complicates the correlation between deamidation levels and chronological or thermal age, especially when samples from diverse regional, environmental and depositional settings are considered, or when there are large intra-site variations across time

that are not considered by the thermal age [29,30]. To understand the underlying mechanisms that may induce or accelerate deamidation in bone, numerous controlled experiments and empirical studies on archaeological material have been conducted to explore potential factors influencing deamidation beyond time, including sampling point, bone type, taphonomy and lab protocols for collagen extraction [7,27,31–33]. Together, they demonstrate the complexity of measuring protein diagenesis even within a single site as the deamidation of consecutive layers does not always follow a linear pattern. These findings, nevertheless, have not dissuaded recent ZooMS studies from continuing to publish deamidation data on COL1 α 1 508–519 using the q2e method, either as individual measurements or in an attempt to explain observed patterns in stratigraphy or biomolecular preservation [9,34–36]. Deamidation of glutamine remains the only post-mortem post-translational modification studied in ZooMS data and is frequently relied upon to demonstrate endogeneity and antiquity of specimens despite the complications with its use. So far, a limited number of ZooMS studies have successfully used deamidation rates to indicate stratigraphic integrity and sample’s authenticity (for example [34,37]).

As an alternative to the q2e method, a new method called Parchment Glutamine Index (PQI) was developed for assessing parchment quality via a sample-level deamidation value and to explore patterns of parchment production during the Medieval period [38]. To date there has been no publication on the application of PQI on other protein-based archaeological material.

1.2. From q2e to PQI and MDS

ZooMS peptide mass fingerprint spectrum is generated using Matrix-Assisted Laser Desorption/Ionisation (MALDI-TOF) mass spectrometry, a rapid yet low-resolution method of measuring proteins and peptides [1,39]. While MALDI-TOF MS may limit precise quantification of post-translational modifications, the rapidly growing volume of publicly available ZooMS data offers a valuable source for metadata analysis. These datasets can help reveal diachronic patterns from single archaeological sites or even large-scale geographic areas. The q2e method employs a genetic algorithm to estimate deamidation levels by stochastically generating a series of candidate values and selecting the one that best fits the isotopic envelope of a glutamine-containing peptide [18]. Despite its high-throughput nature and widespread use in the field, the q2e method has several limitations. Although the q2e method generates peptide-level deamidation values, it does not estimate an overall sample-level deamidation value based on multiple peptides. Most applications focus on a single peptide, COL1 α 1 508–519, as it is widely observed in MALDI-TOF spectra and the peptide sequence remains constant across multiple species. Van Doorn and colleagues [28] proposed a composite deamidation value based upon the average of different peptides, weighted by its logarithmic correlation with the peptide COL1 α 1 508–519. This again places more weight on this ubiquitous peptide without any consideration on the variation in peptide-specific deamidation kinetics. Moreover, values derived from sample triplicates are often averaged in a simplistic manner, without considering differences in spectral quality among the replicates. Additionally, the peptide list for the q2e lacks taxonomic information and is not easily configurable for the users to select different sets of peptides for different taxa. Furthermore, it uses a stochastic and non-deterministic genetic algorithm optimisation to estimate the deamidation, while the deconvolution can in fact be formalised as a linear least squares optimisation problem solvable through a mathematical expression [38]. Finally, the preprocessing and fitting steps are not modularised in a way that allows easy combination with other approaches, and the results of the peaks signal-to-noise ratios and quality of the fitting cannot be easily integrated into downstream analysis.

By contrast, the Parchment Glutamine Index (PQI) [38] uses weighted least-squares (WLS) to estimate deamidation at the peptide level and a linear mixed-effects model (LME) to combine the peptides and replicates estimates on a sample level index, based on an explicitly user-defined list of peptides. After spectral preprocessing, each peptide's deamidation is fitted by weighted least squares (WLS); replicates with higher signal-to-noise ratio (SNR) receive higher weights, and peptides shown to be more responsive to deamidation receive greater influence in the overall sample score. The peptide estimates for each replicate are then integrated and in turn weighted by their WLS fitting residuals, to yield a single PQI value representing sample level deamidation.

To assess this new method's suitability to the analysis of bone, we adapted the PQI statistical framework to develop the **MALDI Deamidation Score (MDS)** - a streamlined pipeline with WLS-LME architecture tailored for large ZooMS MALDI-TOF datasets. MDS retains the flexibility of PQI while offering improved usability and applicability to diverse peptide sets and taxa. We evaluate its performance against the q2e method and apply it to the Denisovan dataset [40], which is currently the largest open-access multi-layer ZooMS dataset from a single archaeological site. This application assesses the potential of MDS to improve deamidation estimation accuracy and to refine our understanding of collagen preservation in archaeological or palaeontological contexts. For clarity, in the following text, peptide-level deamidation estimates generated by MDS are referred to as **Weighted Least Squares q2e (WLS-q2e)** and are subsequently integrated into an MDS score representing sample-level deamidation, whereas deamidation estimates based on COL1 α 1 508–519 (*m/z* 1105) using the original q2e method developed by Wilson et al. [18] are referred to as **Genetic Algorithm q2e (GA-q2e)**. In MDS, higher values indicate higher levels of deamidation.

1.3. Material

Denisova Cave, located in the Altai region of Russian Siberia, has yielded valuable ZooMS data collected from its three chambers [40–42]. The deamidation analysis [40–42] specifically focused on samples from the East Chamber, where the majority of samples from the site had been analysed using ZooMS. This dataset spans a long stratigraphic sequence, encompassing seven layers over approximately 240,000 years [43,44]. Such a comprehensive time range makes it particularly suitable for assessing collagen preservation changes over extended periods and across diverse depositional environments. In total, over 8000 bone fragments from Denisova Cave were analysed using ZooMS [40–42]. Two different extraction protocols were used across the entire dataset, an acid-based protocol and the AmBic protocol developed specifically for the analysis of deamidation [25]. Brown and colleagues [35] selected 2459 samples for inclusion in their deamidation study, which focused specifically on the peptide COL1 α 1 508–509. These were chosen following four criteria; i) specimens were excavated from the East Chamber, to give some stratigraphic control, ii) collagen was extracted following a previously published AmBic protocol [25,27,45], iii) each specimen produced three replicate MALDI-TOF MS spectra (including peptide COL1 α 1 508–509), iv) each specimen could be taxonomically identified using ZooMS. Spectra were acquired at the Department of Archaeology, Max Planck Institute for the Science of Human History (MPI-SHH) in Jena, Germany. Peptides are reported following standard conventions [46]. A total of 2459 samples, each with three replicates, would have resulted in an expected dataset of 7377 spectra; however, only 7358 spectra were retrievable from the online database.

For the present analysis, we focused on the following nine taxonomic groups identified in Brown et al. [35]: *Bison/Yak*, *Capra*, *Ovis*, *Rhinocerotidae*, *Equidae*, *Cervidae/Gazella/Saiga*, *Ursidae*, *Canidae*, and *Crocota/Panthera*. These nine groups comprise the majority of terrestrial mammalian specimens in the dataset, including both herbivores and carnivores. Moreover, they share several peptides containing at least one glutamine residue, making them well-suited for evaluating the

performance of our novel MDS method across multiple taxa.

2. Method

2.1. Selection of peptides

We selected seven peptides for the MDS analysis of the nine major taxa groups included in this study (Table 1). Those peptides contain one glutamine (with their potential hydroxylated versions) and share the same positions by all nine taxa groups. PQI [38] has demonstrated that, when provided with peptide sequences, a theoretical isotopic distribution can be generated for peptides containing more than one glutamine, which can then be used to quantify deamidation. However, given the breadth of the database at use and to keep comparability across multiple taxa, we focus mainly on those peptides which are visible in ZooMS spectra and are already included in routine taxonomic identifications as they tend to ionise and be detected well by MALDI-TOF MS (COL1 α 1 508–519, COL1 α 2 757–789, COL1 α 2 757–789', COL1 α 2 978–990, and COL1 α 2 978–990'). Two additional peptides, which are not routinely included in ZooMS analysis, were also chosen as they were visible across ZooMS spectra and each contained one glutamine (COL1 α 1 435–453' and COL1 α 1 781–789). The peptide list for *Bison/Yak* is shown here as an example on Table 1, and the full list including all nine taxa is available with the code. See Data and code availability.

2.2. Spectra preprocessing

The aim of the preprocessing is to extract the isotopic distribution of the selected peptides in order to calculate the extent of deamidation for each. This follows the same steps undertaken by the previous PQI method established by Nair and colleagues [38]. Our new MDS leverages the infrastructure provided by the Spectra R package (v4.4.0, [47]) to load and process the spectra. This allows for batch loading and processing of the spectra, overcoming the need for excessive memory, and allowing faster parallel processing, which presented a challenge in PQI. Once loaded, the spectra are smoothed using the Savitzky-Golay filter [48] with a half-window size of 6. Following this, the baseline is estimated using the Statistics-sensitive Non-linear Iterative Peak-clipping algorithm (SNIP) with 50 decreasing iterations [49] on the weighted-moving-average smoothed (with half-window size 4) spectra to avoid the peak boundary artifacts introduced by the Savitzky-Golay filter [50]. This baseline is then subtracted from the Savitzky-Golay filtered spectra, which crucially is better at preserving peak shape. Peak picking is applied with a half-window size of 10 to detect local maxima. Noise is estimated using Friedman's SuperSmoother (Friedman, 1984) and peaks above a signal-to-noise ratio (S/N) above 2 are retained. Pseudo isotopic distributions are extracted for each of the peptides by locating the monoisotopic peak plus 4 consecutive peaks at distance of 1 Da with a fixed tolerance of 0.4 Da and 50 ppm. Incomplete isotopic patterns with gaps or less than 4 peaks are discarded as in Nair et al. [38]. Finally, the intensities of each peptide's isotopic envelope are normalised to its highest peak.

2.3. Taxon specific peptide selection

Based on the number of isotopic peaks extracted for each selected peptide, users should further select peptides that either (i) have a high proportion of spectra with at least four isotopic peaks in the envelope, or (ii) show consistently low isotopic peak detection but are commonly observed across the selected species. In either case, users are expected to justify their peptide selection based on the detection profile and the biological or taxonomic relevance of the peptides.

2.4. Estimation of peptide deamidation

We fitted the theoretical isotopic distribution of the non-deamidated

Table 1

List of common peptides chosen to calculate MDS for the 9 taxonomic groups and their corresponding pept_number. Peptides are ordered by increasing *m/z*. References refer to the first publication of this peptide in ZooMS literature.

Pept number	Pept name	<i>m/z</i>	Sequence	N Hyp	Current ZooMS Marker
pep1	COL1α1 781–789	898.51	GVVGLPGQR	1	F (Welker et al., 2016; 2017)
pep2	COL1α1 508–519	1105.58	GVQGPAGPAGPR	1	T (Buckley et al., 2009)
pep3	COL1α2 978–990	1192.68	IGQPGAVGPAGIR	0	T (Buckley et al., 2009)
pep4	COL1α2 978–990	1208.68	IGQPGAVGPAGIR	1	T (Buckley et al., 2009)
pep5	COL1α1 435–453	1706.77	DGEAGAQGGPPGAPAGER	1	F (Welker et al., 2016; 2017)
pep6	COL1α2 757–789	3017.47	GPSGEPGTAGPPGTPGQGLLGAPGFLGLPGSR	4	T (Buckley et al., 2009)
pep7	COL1α2 757–789	3033.47	GPSGEPGTAGPPGTPGQGLLGAPGFLGLPGSR	5	T (Buckley et al., 2009)

and deamidated version of the peptides to the data using WLS regression and obtained the coefficients, γ_0 and γ_1 , which represent the amount of each deamidation state respectively, calculated as:

$$\hat{\gamma} = (\hat{\gamma}_0, \dots, \hat{\gamma}_k) = \underset{\gamma}{\operatorname{argmin}} \sum_{i=0}^m w(y_i - X_i \gamma)^2 \quad (1)$$

Where *k* is the number of glutamines, or deamidation states, in the peptide, and the number of isotopic peaks used. y_i is the normalised intensity of the *i*-th peak in the observed envelope, X_i is a row vector with the theoretical peaks values for the *i*-th position and γ a column with the coefficients. For the case of *k* = 1 from the estimated $\hat{\gamma}_0$ and $\hat{\gamma}_1$, the fraction of deamidation is then

$$q2e = 1 - \frac{\hat{\gamma}_0}{\hat{\gamma}_0 + \hat{\gamma}_1} \quad (2)$$

The weights, *w*, used in eq. 1 are the signal-to-noise ratio and the inverse of the mass error of each peak of the isotopic series. The difference with the approach by Nair and colleagues [38] is that here, the observed intensities (y_i) are normalised for each isotopic envelope and

therefore the scaling factor is not required.

2.5. Quality control

Throughout the process of MDS we applied several quality control filters based on different measurements. First we removed the peptide envelopes from any sample in which an isotopic peak had a mass error above 0.4 Da. For WLS-q2e estimation, we calculate the reliability as the squared sum of the residuals and remove any peptide estimate from any sample with a reliability above 0.2. Finally, any WLS-q2e estimate below 0 or above 1 is discarded.

2.6. MDS estimation

MDS is calculated as described in Nair et al. [38]. Briefly, a linear mixed effect model (LME) is fitted using the $\log(q2e)$ values as the response variable, the peptide as the fixed effect, and the sample and replicate as the random effects. The squared sum of errors of the previous WLS step (reliability) is used as a weight or scaling factor for the residual variances (see [38]). The MDS is the sample-level random effect of the model, which is extracted after the model is fitted. In sum, the whole workflow of MDS is illustrated by Fig. 1.

2.7. Cluster analysis as another line of evidence

To explore whether other modifications, particularly hydrolysis, might have influenced the spectral patterns, we performed a hierarchical cluster analysis of the *Bison/Yak* samples. The results however are dominated by batch effects. For more information, see Supplementary Information.

3. Results

3.1. Improvement of MDS

3.1.1. More efficient preprocessing

We tested the processing times on the *Bison/Yak* samples (2178 spectral files according to the metadata from Brown et al. [35]), and found that GA-q2e, without considering any manual filtering or replicate averaging, was about 3 and 22 times slower than the complete PQI and MDS pipelines, respectively (Table 2).

Table 2

Comparison of running times for MDS, GA-q2e, and PQI on the *Bison/Yak* dataset. The total processing time for GA-q2e was 3032 s, excluding the time required for converting mzML files to TXT or CSV formats. For MDS, the reported preprocessing time includes all spectrum and aggregate files from the triplicate measurements (*).

Algorithm	Running time (s)				File format
	Preprocessing	WLS	LME	Total	
GA-q2e				3032	txt/csv
PQI	1955	7	20	1982	mzML
MDS	102/140*	10	20	132/170*	mzML

Fig. 1. Workflow for calculating MDS. The process begins with peptide list preparation, followed by spectra preprocessing including smoothing, baseline correction, and peak picking (requiring isotopic envelopes with ≥ 4 peaks). Based on the peak picking results, taxon-specific peptides are selected. WLS q2e estimation is then performed, followed by quality control (QC), and finally, the MDS is calculated.

The GA-q2e method relies on sequentially parsing text-based CSV files, which is more computationally demanding than reading spectral data stored in indexed mzML files. Additionally, GA-q2e employs slower preprocessing algorithms for signal-to-noise (S/N) estimation and peak detection. Most significantly, the genetic algorithm underlying GA-q2e greatly increases computation time compared to the WLS method, wherein values can be calculated directly using an analytical expression. In turn, MDS achieves substantial improvements in computational efficiency relative to the earlier PQI method by leveraging the Spectra package infrastructure [47] for handling MALDI-TOF spectral data. Thus, the most notable efficiency gains among these methods occur in the preprocessing step of MDS.

3.2. Normalization of isotopic series

The new MDS tool formats the extracted peptide isotopic peaks data to use standard R model functions (like 'lm') for peptide WLS-q2e estimation step, instead of an ad-hoc WLS fitting function as implemented in the MALDIpqi package (for the PQI method). This allows more flexibility and potentially the use of different models or fitting functions within the R environment. Furthermore, the observed intensities within each isotopic envelope are internally normalised to the same scale as the theoretical isotopic envelopes, which removes the need for an extra scaling factor as was necessary for the WLS step in the PQI method.

3.3. Differential detection of peptide isotopic envelopes

The estimation of q2e values by WLS in the MDS pipeline relies on the accurate detection of the selected peptide isotopic envelopes. From the peak picking step of our preprocessing pipeline, we observed differential detection of peptides, including the same or equivalent peptide marker across taxa (Fig. 2 and Table S1).

Pep1 (COL1 α 1 781–789) and pep2 (COL1 α 1 508–519) are detected with isotopic envelopes comprising at least four peaks in over 70 % and 90 % of spectra, respectively, across all taxa except for *Ovis* (pep1 and pep2 respectively observed in 63.67 % and 85.71 % of spectra) and *Bison/Yak* (pep1 observed in 68.08 %) (Fig. 2 and Table S1). Pep3 and pep4 correspond to the COL1 α 2 978–990 marker and its fully

hydroxylated version. These peptides are consistently recovered across all herbivores except *Equidae*. Among carnivores, pep3 and pep4 are only reliably observed in *Canidae*, while the homologous sequence in *Crocota/Panthera* (TGHPGSGVPAGVR) lacks a glutamine residue and is therefore excluded from MDS analysis. Even so, only about 25 % of *Crocota/Panthera* spectra exhibit a detectable isotopic envelope for either the hydroxylated or non-hydroxylated variant, indicating its low detection rates in MALDI-TOF MS. A similar pattern is observed in *Ursidae*, where the pep3 and pep4 markers were excluded due to the absence of a glutamine. Furthermore, when detected, this peptide in *Ursidae* predominantly appears in its hydroxylated form.

Pep6 and pep7 correspond to COL1 α 2 757–789 and its hydroxylated variant. In *Bison/Yak*, *Ovis*, and *Cervidae/Gazella/Saiga*, pep6 overlaps with a common conserved collagen peptide for bovid (GPPGAS-GAPGPQGFQPPGEPGPGQTGPAGAR; COL1 α 2 89–130) [51] and is frequently detected. However, for all other taxa in which pep6 does not overlap with COL1 α 2 89–130 (except *Ursidae*), typically only the hydroxylated version is detected, and the non-hydroxylated version is rarely observed. This suggests that pep6 sequences identified for *Bison/Yak*, *Ovis*, and *Cervidae/Gazella/Saiga* likely originate predominantly from the non-marker peptide. *Ursidae* exhibits a reversed pattern for pep6 and pep7, likely because the corresponding peptide sequence for *Ursidae* (GPSGEPGTAGPPGTAGPQGIIGAPGIIGIPGSR) contains only four prolines in the Y-position that can be easily hydroxylated (which generally tend to be fully hydroxylated), whereas in other taxa, the equivalent peptide contains five hydroxyproline residues.

Pep5 (COL1 α 1 435–453) is common across all taxa but is poorly represented in the Denisovan dataset. For instance, only around 25 % to 30 % of the pep5 detected for *Ovis* and *Crocota/Panthera* contain a clear isotopic envelope with at least four peaks; notably, this detection rate is already the highest among all taxa. Nevertheless, despite its relatively low spectral representation, pep5 was included in the MDS analysis because of its taxonomic ubiquity and its prior application in q2e-based deamidation estimations [27].

Therefore, based on the peak picking results, peptides were further refined from the preselected peptide list (Table 1) for downstream MDS estimation. Specifically, the selected peptides included pep1, pep2, pep3, pep4, pep5, and pep7 for *Bison/Yak*, *Ovis*, *Capra*, *Cervidae*/

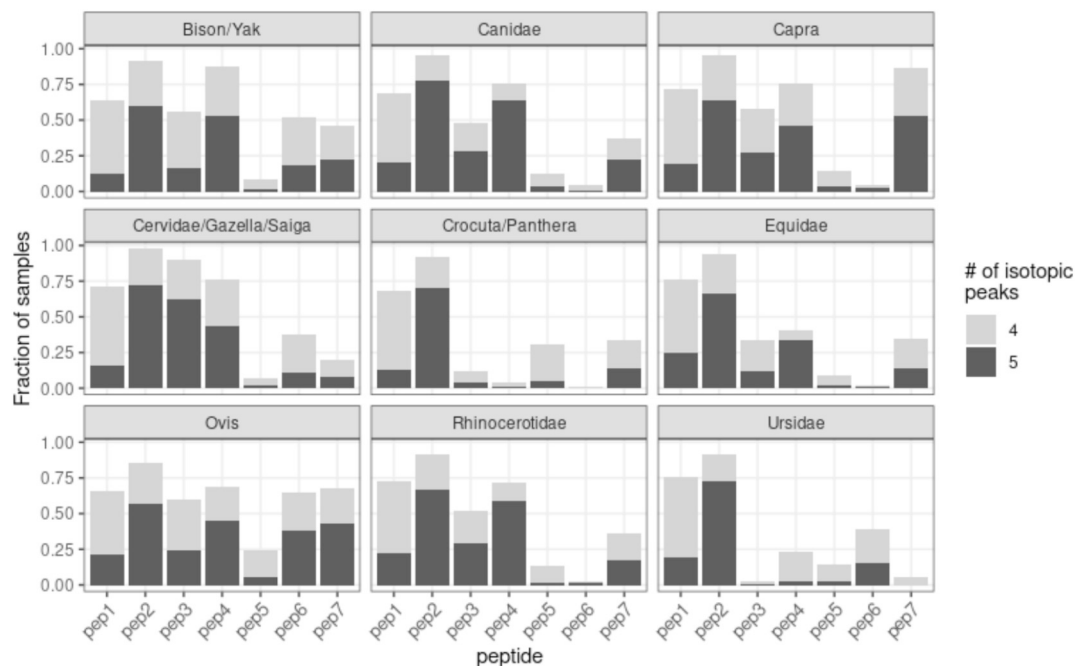


Fig. 2. Fraction of samples in each taxon where each peptide was detected with four or five isotopic peaks. Peptides with at least four isotopic peaks are eligible for inclusion in MDS-based deamidation estimation. Color shading indicates the number of peaks detected (light gray = 4, dark gray = 5).

Gazella/Saiga, *Rhinocerotidae*, *Equidae*, and *Canidae*; pep1, pep2, pep5, and pep7 for *Crocota/Panthera*; and pep1, pep2, pep5, and pep6 for *Ursidae*.

3.4. WLS-q2e peptide filtering

Following our quality control procedure, which filters out unreliable peak data based on WLS sum of squared residuals, WLS-q2e values were generated for each of the samples using MDS based on the selected peptides. At the sample level, triplicates were aggregated and all retained WLS-q2e with residual <0.2 fell within the (0,1) interval. For non-aggregated values, 4.38 % ($n = 813$ of 18,549) that passed the residual filter fell outside (0, 1). Samples with at least two valid WLS-q2e values were retained for further analysis. This filtering resulted in 2169 samples being included in the final MDS dataset, comprising 718 *Bison/Yak*, 750 *Cervidae/Gazella/Saiga*, 306 *Equidae*, 164 *Rhinocerotidae*, 63 *Canidae*, 63 *Capra*, 52 *Ursidae*, 27 *Ovis*, and 26 *Crocota/Panthera*.

Pep1 and pep2 are the two most reliable peptides. For pep2 specifically, WLS-q2e values were generated in 2137 samples, of which 2109 passed quality control and were retained for downstream analysis; slightly fewer pep1 values passed the quality control. Peptides with higher m/z tend to lose more values during the filtering. Interestingly, among the marker pair pep3 and pep4, pep3 showed the most significant data loss (>50 %) for *Cervidae/Gazella/Saiga* and *Canidae*, whereas pep4 exhibited greater data loss for *Equidae* and *Rhinocerotidae*. Fig. 3 summarizes the number of WLS-q2e values generated and retained for each peptide across taxa.

3.5. Peptide level deamidation

To evaluate the performance of the WLS regression on the selected peptides, we compared WLS-q2e values from all analysed peptides against GA-q2e values derived from pep2. Overall, a strong positive correlation was observed across taxa, demonstrating broad consistency between the two approaches (Fig. S1). However, we still observe substantial variability among taxa and peptides (Fig. 4). Pep1 and Pep2 are the most frequently recovered peptides in the MDS analysis. Across all taxa, Pep1 consistently shows higher deamidation levels than Pep2, with median values generally exceeding 0.5, whereas the median for Pep2 for cases is below 0.5 (Fig. 4). This suggests that within the analysed

dataset, Pep1 is deamidated more rapidly than Pep2.

The ZooMS marker pair pep3 and pep4 also show elevated deamidation values compared to pep2, with median always higher than 0.5 but with subtle variation between the two. In most taxa, pep4 is more deamidated than pep3 in terms of the median, except in *Cervidae/Gazella/Saiga*, where the reverse pattern is observed. In *Equidae*, pep4 displays a bimodal distribution, with one group of data clustered below 0.5 and one group above, potentially indicating the presence of two distinct sample preservation conditions. For discussion on this, please see *Supplementary Information*.

Pep5, although detected in fewer samples, shows relatively low deamidation levels (mostly <0.5) across all taxa, suggesting a slower deamidation rate for this peptide similar to pep2. In *Ursidae*, pep6 stands out as the most deamidated peptide, while pep7, present in multiple taxa, spans nearly the entire range (0–1), indicating substantial variation in degradation levels and possibly reflecting peptide-specific detectability on the MALDI spectrum.

Pep1 and pep2 are the two best candidates among the selected peptides for the deamidation estimation due to their consistent presence and their well-defined isotopic envelopes in MALDI-TOF spectra. Among these two, pep2 is better when dealing with assemblage covering deep time as it deamidates the slowest. However, we do not know if this is specific to the assemblage we are testing or a universal phenomenon. In addition, other factors such as the physicochemical properties of the peptides may affect their efficiency in extraction and detection during the lab process. For example, pep1 and pep2 may more easily bond with the C18 resin used in peptide purification in standard ZooMS protocols. They may also co-crystallise particularly efficiently with the matrix solution used to facilitate ionisation of the peptides to enable detectability in the TOF mass analyser [52]. The large variation for heavier peptides such as pep6 and pep7 may be the product of overlapping irrelevant peptide peaks on the spectra.

3.6. Diachronic patterns in peptide deamidation

Across stratigraphic layers (Fig. 5, Fig. S3), most peptides exhibit a general U-shaped pattern in WLS-q2e values, with decreasing deamidation from the oldest layers (e.g., layer 15 and, where available, 17.1) toward mid-sequence layers, followed by an increase in the youngest layers (9.3 to 9.2). However, both inter-peptide and inter-taxon



Fig. 3. Peptide counts per taxon before and after quality filtering. Each cell indicates the number of WLS-q2e estimates retained for each peptide. Values in parentheses represent the total counts before filtering.

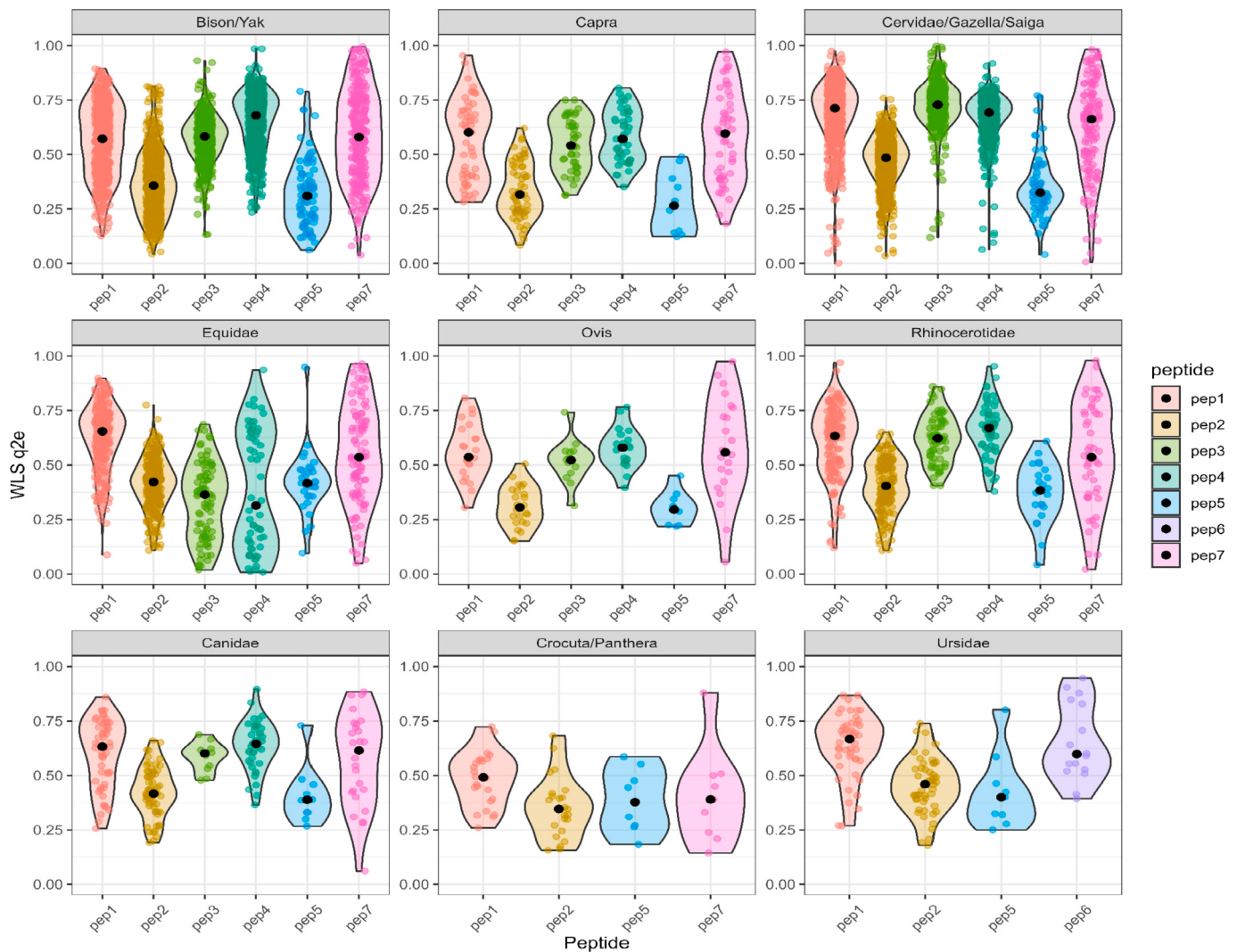


Fig. 4. Violin plots showing the distribution of WLS-q2e values across individual peptides (pep1–pep8) for all the nine taxa. Each panel represents a taxonomic group, with individual peptide distributions coloured consistently across panels. Median values are shown by the black dots.

variability are evident. Pep1 and pep2, which have the highest number of observations, clearly follow this U-shaped trend across most taxa. Pep3 and pep4 display similar distributions across layers when given sufficient sample size.

Pep5, pep6, and pep7 lack sufficient data to establish robust trends across all taxa. Nevertheless, for more abundant taxa such as *Bison/Yak*, pep5 and pep7 show similar trends compared to the well detected pep1 and pep2. However, the high variation of pep7 values, including many near-complete deamidation values close to 1, makes subtle interlayer differences difficult to discern.

Interestingly, peptide-specific deviations are observed within certain layers. For example, in *Bison/Yak*, the behaviour of pep1 and pep2 diverges in layer 11.3 and layer 13 (Fig. 6). For layer 11.3, although the sample size is small ($n = 5$), all five samples show WLS-q2e values for pep2 that are higher than the median values of both adjacent layers. This pattern is not observed in pep1, for which the WLS-q2e values instead show a gradual decreasing sequence from layer 11.4 to layer 11.2. Similarly, pep3 and pep4 also exhibit locally elevated values in layer 11.3, although based on even fewer observations (Fig. 5). On the contrary, pep2 shows a gradual decreasing trend from layer 14 to layer 11.2 with no layers showing statistically significantly higher values in between, while pep1 shows elevated WLS-q2e values in layer 13 which are

higher than the values in layer 12 and layer 14, although not significantly for the latter. The elevated WLS-q2e values for pep1 in layer 13 are also associated with a lower sample size compared to pep2 ($n = 18$ vs. $n = 22$), though the same association is not seen in layer 11.3. A similar, though subtler, pattern is evident in *Cervidae/Gazella/Saiga*. While no samples were recovered from layer 11.3, pep1 exhibits significantly elevated WLS-q2e values in layer 13 compared to layer 12, and slightly elevated values compared to layer 14, although the latter difference is not statistically significant. In contrast, pep2 in layer 13 shows only a slight but statistically significant elevation compared to layer 12 and is similar to layer 14 by visual assessment (Fig. 6).

Equidae appears to differ from this trend associated with layer 13. For pep2, the values in layer 13 are significantly higher than those in layer 12, a pattern not observed for pep1. However, this discrepancy might be attributed to the relatively small sample size available for *Equidae* in layer 13. In layer 11.3, only two samples were recovered for both peptides; notably, the pep2 values of these two samples exceeded the 75th percentile of values from the two neighbouring layers, whereas both pep1 values fall well below the median of layer 11.4, mirroring the pattern observed in *Bison/Yak*. (Fig. 6) In other taxa, limited sample sizes constrain a clear interpretation on the deamidation trends of each peptide, but general U-shape trends across all layers remain broadly

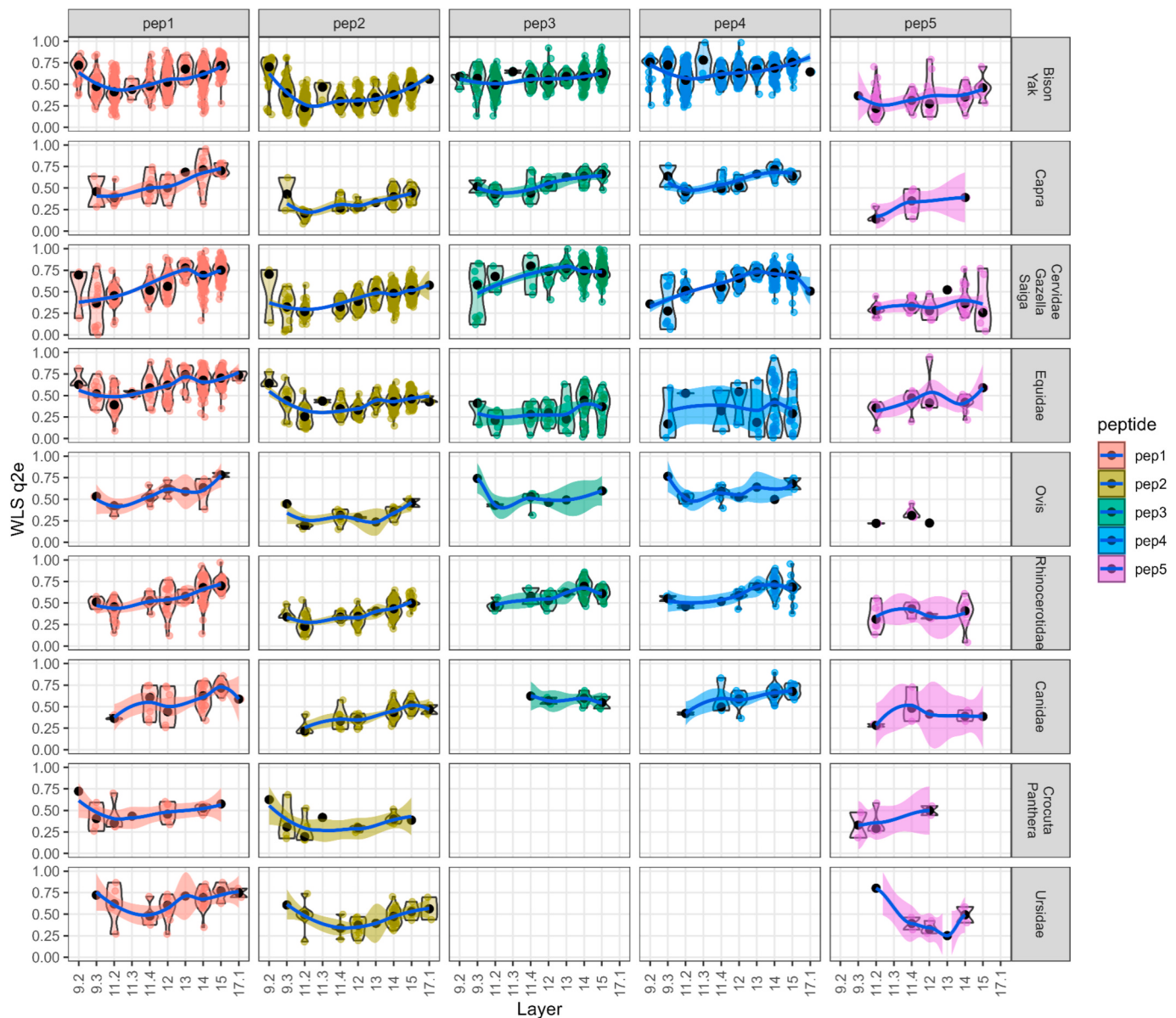


Fig. 5. Peptide deamidation patterns across layer and taxa on pep1 to pep5. For the same plot on pep6 and pep7, see supplementary Fig. S3.

consistent(Fig. 5, Fig. S3).

3.7. MDS

The overall MDS of each sample was calculated by integrating peptide-specific WLS-q2e values using a linear mixed-effects model (LME). MDS captures the overall deamidation trend while exhibiting varying levels of correlation with each peptide (Fig. S2). MDS has the highest correlation with pep2, followed by pep1, pep4 and pep3, while pep5, 6 and 7 show low correlation. Diachronic trends of MDS (Fig. 7) closely mirror those of GA-q2e from Brown et al., [35] (Fig. S4). Both methods show a decline in deamidation from Layer 15 to Layer 11.4, with a minor rise in Layer 13, followed by another decline at Layer 12. A small local increase is observed in Layer 11.3, though based on limited data. Both MDS and GA-q2e reach their minimum in Layer 11.2, after which deamidation levels rise sharply from layer 9.3 to 9.2. This similarity is further supported by a strong linear correlation between MDS and GA-q2e values across samples (Fig. S5). A global linear regression yields a slope of 1.05 ($R^2 = 0.62$, $p < 0.0001$), indicating a strong and statistically significant agreement between the two methods.

4. Discussion

4.1. Fast, streamlined, and flexible

Although the resolution of MALDI-TOF derived ZooMS spectra limits the precise quantification of post-translational modifications [8,53], deamidation estimations produced by the GA-q2e method have been already applied for assessing relative chronology, contamination, and taphonomy. A niche for ZooMS studies is the analysis of large and indeterminate osseous assemblages either to complement zooarchaeological records [36,54]; or as a pre-screening method of identifying hominin remains [6,41,55,56] or the remains of rare fauna [57,58]. These projects have generated a vast repository of ZooMS data at such a rapid speed that meta-analyses at site, regional, and time-series scales are now possible. This accumulation of ZooMS spectra will be further accelerated as laboratory robotics [59] and automated spectral post-processing advances [60–62]. Our benchmarking of processing times for MDS, GA-q2e, and PQI demonstrates that MDS significantly outperforms GA-q2e in terms of speed. It also surpasses PQI, which is its base model.

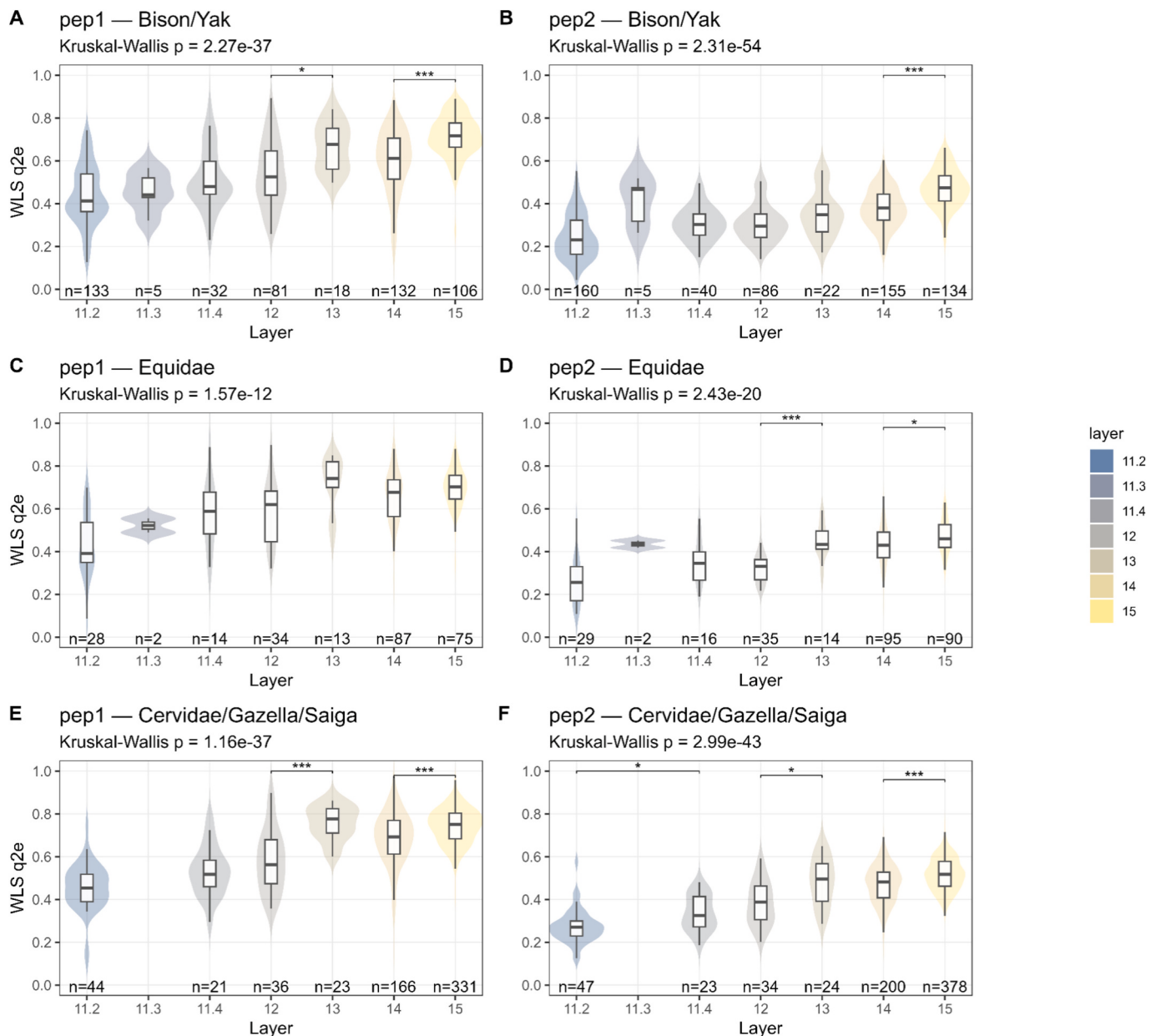


Fig. 6. Distributions of WLS q2e for pep1 and pep2 across layer 15–11.2. Panels are arranged by species (rows) and peptides (columns): A–B, *Bison/Yak*; C–D, *Equidae*; E–F, *Cervidae/Gazella/Saiga*; left column = pep1, right column = pep2. For each layer (11.2–15), violin plots show the kernel density, with overlaid boxplots indicating the median and interquartile range (whiskers = $1.5 \times \text{IQR}$); individual points are not displayed. Per-layer sample sizes (n) are printed near the x-axis. Layers with no observations for a given species (e.g., 11.3 in some groups) remain on the x-axis but have no violin/box. Panel subtitles report the global Kruskal–Wallis p -value across layers for that species–peptide combination. Pairwise tests are limited to adjacent layers (Wilcoxon rank-sum with Bonferroni correction): $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***); non-significant contrasts are omitted.

A significant portion of the processing speed improvement comes from the elimination of numerous time-consuming data preparation steps required before running the deamidation model, as well as from streamlining the overall workflow. These time savings are often masked and difficult to quantify directly. For example, sampling strategy coupled with GA-q2e analysis often involves manual filtering of samples with failed triplicates [35] or additional processing such as averaging spectral triplicates [34], making GA-q2e difficult to scale with the potential and growing need for meta-analysis and lose valuable information from the differences within triplicates. Although PQI shares the same underlying statistical structure, it is less streamlined and requires additional time to execute each step of the workflow. MDS overcomes all these problems by integrating data preprocessing, quality control, and deamidation estimation into a single, automated pipeline. This

streamlined design minimizes manual intervention and significantly improves overall processing efficiency, making it particularly well-suited for high-throughput analysis of large datasets.

MDS offer great flexibility for the users wishing to work on multiple taxa with different targeted peptides. The configuration of peptide lists is significantly simplified in MDS, allowing users to tailor taxa-specific peptide sets within a single Excel sheet that incorporates both m/z and sequence information. Although GA-q2e also permits user-defined peptide lists, its format includes only mass and sequence data without taxonomic specificity. This makes it difficult to process datasets involving multiple taxa with distinct peptide sets, unless further modifications are made to the underlying C++ code [18]. MDS also introduces rigorous preprocessing steps that enable more refined screening of usable peptides, in contrast to GA-q2e, which directly

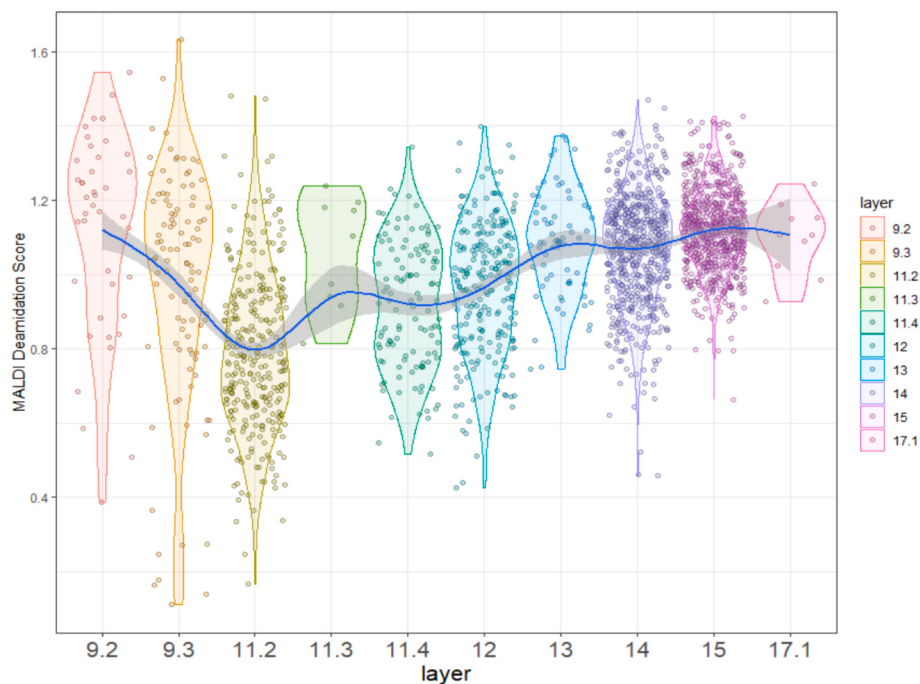


Fig. 7. Diachronic patterns of MDS across stratigraphic layers.

outputs deamidation values. We observed that peptides with higher molecular masses are more frequently filtered out, possibly due to lower spectral visibility, higher noise or overlapping isotopic envelopes with non-target peptides. The observed difference in filtering rates between non-hydroxylated and hydroxylated peptide versions in certain taxa may be attributed to peptide overlapping effects in taxa-specific regions. With MDS's built-in quality control features based on mass errors, SNR or reliability, it is now feasible to map peptide visibility more systematically and to provide informed guidance on the selection of suitable peptides for deamidation analysis using MALDI data and open doors for further analysis to inspect performance of individual peptides in MALDI-TOF MS.

4.2. WLS-q2e reveal nuance pattern obscured by single peptide

A key feature of MDS is that it, for the first time, formalizes the output of WLS-q2e on each individual peptide selected per taxon. This enabled us to detect subtle fluctuations in deamidation rates between groups such as *Bison/Yak*, *Cervidae/Gazella/Saiga*, and *Equidae*, where sufficient sample sizes were available. While the overall deamidation trends across the full stratigraphic sequence remain broadly similar, distinct shifts were observed in layer 13 and layer 11.3, with elevated values for *pep1* and *pep2*, respectively. These layer-specific patterns could not have been identified using GA-q2e only targeting *pep2*, nor by relying solely on the final MDS output that summarizes the overall deamidation rate for each sample. We focused on *pep1* and *pep2* as these peptides were the most well-defined in the spectra based on our quality control criteria, and their sequences remain conserved across most terrestrial mammal taxa. The deamidation trend across all layers of *pep1* are very similar to both WLS-q2e on *pep2* and GA-q2e, suggesting that *pep1* may serve as an additional reliable marker for single-peptide deamidation analysis alongside the widely used *pep2*. *Peptide 2* (COL1 α 1 508–519), although ubiquitously observed with relatively good SNR and peak shape, is a slow deamidating peptide. Just relying on it might decrease the temporal resolution power. The inconsistency of other peptides can be due to instrumental and extraction-specific constraints (e.g., extraction solution and matrix interferences, ionisation efficiency, etc.). As these constraints are relaxed in the future due to improved

instruments and wet lab procedure, more stable detection of these peptides is expected. Furthermore, tools like PAMPA [63] or the sequence-driven approaches such as the work by Lawrence and Buckley [64] provide a way of crafting peptide marker sets from sequence data, potentially expanding the peptides suitable for estimation of deamidation by our MDS framework. In MDS, the user does not select a priori a stable peptide to which the others are referenced as in Wilson et al. [28], and the strong correlation between MDS and *pep2* (Fig. S2) is an outcome of it being less noisy, having less variance across replicates and detected across more samples. Overall MDS simplifies the study of deamidation by aggregating the deamidation estimates into a single value that captures the trend and makes use of the peptide and replicate level information.

4.3. Archaeological implication

Overall, our MDS model not only reproduced the broad deamidation pattern observed with GA-q2e but also yielded more nuanced insight. In both approaches, a gradual decrease in deamidation is evident from layer 15 (and, where available, layer 17) through to layer 11.2, followed by a rapid increase beginning at layer 9.3. The shift between layers 9.3 and 9.2 is attributed to the influx of diagenetic agents, notably phosphates and humic acids derived from bat guano in the overlying Holocene layers [65,66]. These acidic leachates infiltrated layer 9, substantially degrading bone preservation and promoting the formation of secondary phosphatic minerals such as taranakite and leucophosphite but did not penetrate beyond layer 11 [40,65–67], leaving the interval from layer 15 to layer 11 characterised by a steady temporal decline in deamidation with only localised diagenetic alterations ([68]; [40,65–67]; Goldberg, personal communication).

Within this broader trend, peptide-specific patterns for *pep1* and *pep2* reveal additional complexities. Unexpectedly high deamidation was observed for *pep2* in layer 11.3 and for *pep1* in layer 13. While these results may simply reflect the small sample sizes for these layers, they could also suggest diagenetic processes that differentially affected the two peptides. The divergence between *pep1* and *pep2* trends in these layers may reflect highly localised processes acting on individual bones and different protein sequences. The phosphatic stratigraphy of the East

Chamber provides some possible context. Layer 11.3 lies just beneath the strongly phosphatised zone of layer 9, where bat-guano-derived acidic leachates and phosphates are abundant [67]. Although the bulk geochemistry suggests that the phosphate front did not extend beyond layer 11, micro-scale seepage along fractures or porous pockets could have locally altered the chemistry in parts of layer 11.3, selectively accelerating the deamidation of more chemically susceptible peptides such as pep2. In contrast, layer 13 is well within the phosphate-free zone (layers 11.3–17; [67]), yet sedaDNA [69] showed that even in these deeper units, preservation varies markedly between samples, likely due to patchy microenvironments influenced by organic inputs or localised pH fluctuations. These observations indicate that local geochemical processes at Denisova Cave might have affected individual collagen peptides in distinct ways. Applied more broadly, MDS enables such peptide-level patterns to be detected and related to site-specific diagenetic processes, providing a tool to link molecular preservation with depositional and post-depositional environments.

While the lack of taphonomic data for the Denisova assemblage analysed by ZooMS limits our ability to interpret these peptide-specific anomalies on individual bone samples, there is a growing trend to document taphonomic observations, where possible, for indeterminate bones in large-scale ZooMS studies [54,70,71]. Incorporating such information enables more detailed analyses of deamidation in relation to taphonomy and bone type (e.g., [27,31]). Further application of MDS to datasets that include detailed morphological and taphonomic records would allow direct comparison of estimated deamidation rates with specific diagenetic variables, thereby greatly improving our understanding of potential links between deamidation and bone preservation.

5. Conclusion

In this study, we introduce the MALDI Deamidation Score (MDS), a new statistical framework for estimating bone collagen deamidation from MALDI-TOF spectra and demonstrate its application on the large faunal assemblage from Denisova Cave. Compared to the widely used q2e method, MDS integrates peptide-level estimates through weighted least squares and linear mixed-effects models, enabling the resolution of peptide-specific variation that could otherwise bias results. This formalisation provides a flexible framework that enables MDS to employ additional regression methods capable of better mitigating the stochastic noise characteristic of MALDI-TOF mass spectra. By accommodating user-defined peptide sets, triplicate runs, and partially failed spectra without manual filtering, MDS provides a transparent, scalable, and efficient solution for analysing large and taxonomically diverse ZooMS datasets. The application to Denisova Cave demonstrates how MDS can recover both broad deamidation trends and finer peptide-specific features, such as the elevated pep2 deamidation in layer 11.3 and pep1 in layer 13. These patterns, potentially linked to localised diagenetic microenvironments, illustrate the added interpretative value of peptide-level resolution. Such insights would have been obscured under aggregate deamidation estimates, underscoring the importance of integrating peptide-specific analysis into routine workflows. In practice, users may choose to use MDS as a summary measurement where diagenetic pathways are well understood (e.g., parchment) or to leverage WLS-q2e patterns to explore more complex taphonomy.

A particular strength of MDS lies in its adaptability to ongoing advances in mass spectrometry technology. The framework's flexibility in peptide selection, modular design and its robust statistical approach make it well-positioned for future applications as instrumentation continues to improve. As higher-resolution, lower-noise platforms such as MALDI-FTICR (*Fourier-transform ion cyclotron resonance*) MS and TIMS-TOF MS become more widely available, spectral quality will improve, enabling reliable inclusion of peptides currently excluded due to poor isotopic patterns or low signal-to-noise ratios. The current implementation of MDS is optimised for MALDI-TOF and future work will be needed to accommodate ultra-high-resolution instruments, which may

allow individual isotopic peaks to be distinguished from the 0.984 Da deamidation shift.

Realising the full potential of MDS will require pairing molecular-level properties (e.g., physicochemical) with detailed zooarchaeological and taphonomic records (e.g., bioclimatic and geochemical variables). Systematic contextual documentation could allow MDS outputs to be interpreted in relation to bone type, surface modifications, and diagenetic pathways.

MDS offers a flexible, scalable, and easy to interpret framework, thereby enabling us to measure glutamine deamidation in large-scale zooarchaeological datasets in a rapid and efficient manner. Even with limited metadata, our results show that MDS can yield valuable insights into diagenetic and taphonomic patterns from ZooMS data. As measuring deamidation becomes routine in palaeoproteomic studies, MDS provides a powerful approach to explore biomolecular preservation, with the potential to be further strengthened through future integration with high-resolution mass spectrometry and detailed archaeological records.

CRediT authorship contribution statement

Fei Yang: Writing – review & editing, Writing – original draft, Visualization, Validation, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Ismael Rodríguez Palomo:** Writing – review & editing, Writing – original draft, Visualization, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation. **Bharath Anila Bhuvanendran Nair:** Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Investigation, Data curation. **Samantha Brown:** Writing – review & editing, Writing – original draft, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Data and code availability

The R package maldiMDS used to calculate the MALDI Deamidation Scores is available on GitHub (<https://github.com/ismaRP/maldiMDS>) and archived on Zenodo (<https://doi.org/10.5281/zenodo.16938781>). The scripts used to analyse the data and produce these results can be found on GitHub (https://github.com/ismaRP/MDS_Denisova) and archived on Zenodo (<https://doi.org/10.5281/zenodo.16946802>). It also contains the peptide table and metadata used. The data used in this study was obtained from the repositories <https://doi.org/10.17632/bjvt47wb9c.1>, <https://doi.org/10.17632/wfsv6bt64w.1>, <https://doi.org/10.17632/h79h6t77tj.1> as they were archived by Brown et al. (2021).

Acknowledgement

We thank Paul Goldberg for his advice on the geoarchaeology of Denisova Cave. We thank Matthew Collins for engaging in discussions about deamidation and collagen preservation with us. We thank all the archaeologists (especially Maxim Kozlikin, Michael Shunkov, Anatoly Derevianko) for recovering the material from Denisova Cave, and, together with many others who published the initial data study in 2021, for their commitment to open science, which is vital to the advancement of the scientific community. Fei Yang was supported by The University of Tübingen's Excellence Initiative. Ismael Rodríguez Palomo and Bharath Anila Bhuvanendran Nair worked on this study in their free time.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jprot.2025.105577>.

References

- [1] M. Buckley, M. Collins, J. Thomas-Oates, J.C. Wilson, Species identification by analysis of bone collagen using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry, *Rapid Commun. Mass Spectrom.* 23 (2009) 3843–3854.
- [2] S. Fiddymment, B. Holsinger, C. Ruzzier, A. Devine, A. Binois, U. Albarella, R. Fischer, E. Nichols, A. Curtis, E. Cheese, M.D. Teasdale, C. Checkley-Scott, S. J. Milner, K.M. Rudy, E.J. Johnson, J. Vnouček, M. Garrison, S. McGrory, D. G. Bradley, M.J. Collins, Animal origin of 13th-century uterine vellum revealed using noninvasive peptide fingerprinting, *Proc. Natl. Acad. Sci. USA* 112 (2015) 15066–15071.
- [3] A.N. Coutu, G. Whitelaw, P. le Roux, J. Sealy, Earliest Evidence for the Ivory Trade in Southern Africa: Isotopic and ZooMS Analysis of Seventh–Tenth Century ad Ivory from KwaZulu-Natal, *Afr. Archaeol. Rev.* 33 (2016) 411–435.
- [4] L.O. Brandt, J.A. Ebsen, K. Haase, Leather shoes in early Danish cities: choices of animal resources and specialization of crafts in Viking and medieval Denmark, *Eur. J. Archaeol.* 23 (2020) 428–450.
- [5] K.K. Richter, M.C. Codlin, M. Seabrook, C. Warinner, A primer for ZooMS applications in archaeology, *Proc. Natl. Acad. Sci. USA* 119 (2022) e2109323119.
- [6] F. Welker, M. Hajdinjak, S. Talamo, K. Jaouen, M. Dannemann, F. David, M. Julien, M. Meyer, J. Kelso, I. Barnes, S. Brace, P. Kamminga, R. Fischer, B.M. Kessler, J. R. Stewart, S. Pääbo, M.J. Collins, J.-J. Hublin, Palaeoproteomic evidence identifies archaic hominins associated with the Châtelperronian at the Grotte du Renne, *Proc. Natl. Acad. Sci. USA* 113 (2016) 11162–11167.
- [7] M. Pal Chowdhury, R. Wogelius, P.L. Manning, L. Metz, L. Slimak, M. Buckley, Collagen deamidation in archaeological bone as an assessment for relative decay rates, *Archaeometry* 61 (2019) 1382–1398.
- [8] F. Bray, I. Fabrizio, S. Flament, J.-L. Locht, P. Antoine, P. Auguste, C. Rolando, Robust high-throughput proteomics identification and deamidation quantitation of extinct species up to Pleistocene with ultrahigh-resolution MALDI-FTICR mass spectrometry, *Anal. Chem.* 95 (2023) 7422–7432.
- [9] M. Antonosyan, E. Hill, M. Jodry, N. Amano, S. Brown, T. Rick, N. Boivin, A new legacy: potential of zooarchaeology by mass spectrometry in the analysis of north American megafaunal remains, *Front. Mamm. Sci.* 3 (2024), <https://doi.org/10.3389/fmamm.2024.1399358>.
- [10] C. Peters, Y. Wang, V. Vakil, J. Cramb, J. Dortch, S. Hocknull, R.A. Lawrence, T. Manne, C. Monks, G.E. Rössner, H. Ryan, M. Siversson, T. Ziegler, J. Louys, G. Price, N. Boivin, M. Collins, Bone collagen from subtropical Australia is preserved for more than 50,000 years, *Commun. Earth Environ.* 4 (2023) 438.
- [11] J.W. Scotcher, A.B. Robinson, Deamidation of glutamyl residues: dependence on pH, temperature, and ionic strength, *Anal. Biochem.* 59 (1974) 319–322.
- [12] S. Catak, G. Monard, V. Aviyente, M.F. Ruiz-López, Deamidation of asparagine residues: direct hydrolysis versus succinimide-mediated deamidation mechanisms, *J. Phys. Chem. A* 113 (2009) 1111–1120.
- [13] A.B. Joshi, M. Sawai, W.R. Kearney, L.E. Kirsch, Studies on the mechanism of aspartic acid cleavage and glutamine deamidation in the acidic degradation of glucagon, *J. Pharm. Sci.* 94 (2005) 1912–1927.
- [14] D.L. Riggs, J.W. Silzel, Y.A. Lyon, A.S. Kang, R.R. Julian, Analysis of glutamine Deamidation: products, pathways, and kinetics, *Anal. Chem.* 91 (2019) 13032–13038.
- [15] R.M. Daniel, M. Dines, H.H. Petach, The denaturation and degradation of stable enzymes at high temperatures, *Biochem. J.* 317 (Pt 1) (1996) 1–11.
- [16] N.E. Robinson, Z.W. Robinson, B.R. Robinson, A.L. Robinson, J.A. Robinson, M. L. Robinson, A.B. Robinson, Structure-dependent nonenzymatic deamidation of glutamyl and asparagyl pentapeptides, *J. Pept. Res.* 63 (2004) 426–436, <https://doi.org/10.1111/j.1399-3011.2004.00151.x>.
- [17] T.C. Terwilliger, S. Clarke, Methylation of membrane proteins in human erythrocytes. Identification and characterization of polypeptides methylated in lysed cells, *J. Biol. Chem.* 256 (1981) 3067–3076.
- [18] J. Wilson, N.L. van Doorn, M.J. Collins, Assessing the extent of bone degradation using glutamine deamidation in collagen, *Anal. Chem.* 84 (2012) 9041–9048.
- [19] P.P. Madupe, C. Koenig, I. Patramanis, P.L. Rüther, N. Hlazo, M. Mackie, M. Tawane, J. Krueger, A.J. Taurozzi, G. Troché, J. Kibii, R. Pickering, M. R. Dickinson, Y. Sahle, D. Kgotleng, C. Musiba, F. Manthi, L. Bell, M. DuPlessis, C. Gilbert, B. Zipfel, L.F.K. Kuderna, E. Lizano, F. Welker, P. Kyriakidou, J. Cox, C. Mollereau, C. Tokarski, J. Blackburn, J. Ramos-Madriral, T. Marques-Bonet, K. Penkman, C. Zanolli, L. Schroeder, F. Racimo, J.V. Olsen, R.R. Ackermann, E. Cappellini, Enamel proteins reveal biological sex and genetic variability in southern African Paranthropus, *Science* 388 (2025) 969–973.
- [20] E. Cappellini, F. Welker, L. Pandolfi, J. Ramos-Madriral, D. Samodova, P.L. Rüther, A.K. Fotakis, D. Lyon, J. Victor Moreno-Mayar, M. Bukhsianidze, R.R. Jersie-Christensen, M. Mackie, A. Ginolhac, R. Ferring, M. Tappen, E. Palkopoulou, M. R. Dickinson, T.W. Stafford, Y.L. Chan, A. Götherström, S.K.S. Nathan, P. D. Heintzman, J.D. Kapp, I. Kirillova, Y. Moodley, J. Agusti, R.-D. Kahlke, G. Kiladze, B. Martínez-Navarro, S. Liu, M.S. Velasco, M.-H.S. Sinding, C. D. Kelstrup, M.E. Allentoft, L. Orlando, K. Penkman, B. Shapiro, L. Rook, L. Dalén, M.T.P. Gilbert, J.V. Olsen, D. Lordkipanidze, E. Willerslev, Early Pleistocene enamel proteome from Dmanisi resolves Stephanorhinus phylogeny, *Nature* 574 (2019) 103–107.
- [21] A.B. Robinson, J.H. McKerron, P. Cary, Controlled deamidation of peptides and proteins: an experimental hazard and a possible biological timer, *Proc. Natl. Acad. Sci. USA* 66 (1970) 753–757.
- [22] A.B. Robinson, C.J. Rudd, Deamidation of glutamyl and asparagyl residues in peptides and proteins, *Curr. Top. Cell. Regul.* 8 (1974) 247–295.
- [23] P. Hao, S.S. Adav, X. Gallart-Palau, S.K. Sze, Recent advances in mass spectrometric analysis of protein deamidation, *Mass Spectrom. Rev.* 36 (2017) 677–692.
- [24] A. Ramsøe, M. Crispin, M. Mackie, K. McGrath, R. Fischer, B. Demarchi, M. J. Collins, J. Hendy, C. Speller, Assessing the degradation of ancient milk proteins through site-specific deamidation patterns, *Sci. Rep.* 11 (2021) 7795.
- [25] N.L. van Doorn, H. Hollund, M.J. Collins, A novel and non-destructive approach for ZooMS analysis: ammonium bicarbonate buffer extraction, *Archaeol. Anthropol. Sci.* 3 (2011) 281–289, <https://doi.org/10.1007/s12520-011-0067-y>.
- [26] N.E. Robinson, A.B. Robinson, Prediction of protein deamidation rates from primary and three-dimensional structure, *Proc. Natl. Acad. Sci. USA* 98 (2001) 4367–4372.
- [27] F. Welker, M.A. Soressi, M. Roussel, I. van Riemsdijk, J.-J. Hublin, M.J. Collins, Variations in glutamine deamidation for a Châtelperronian bone assemblage as measured by peptide mass fingerprinting of collagen, *Sci. Technol. Archaeol. Res.* 3 (2017) 15–27.
- [28] N.L. van Doorn, J. Wilson, H. Hollund, M. Soressi, M.J. Collins, Site-specific deamidation of glutamine: a new marker of bone collagen deterioration, *Rapid Commun. Mass Spectrom.* 26 (2012) 2319–2327.
- [29] E.R. Schroeter, T.P. Cleland, Glutamine deamidation: an indicator of antiquity, or preservation quality? *Rapid Commun. Mass Spectrom.* 30 (2016) 251–255.
- [30] M. Pal Chowdhury, M. Buckley, Trends in deamidation across archaeological bones, ceramics and dental calculus, *Methods* 200 (2022) 67–79.
- [31] R.D. Åsmundsdóttir, J. Hansen, Z. Fagernäs, G. Troché, J.V. Olsen, M. Saña Seguí, F. Welker, Early Holocene preservation differences between cortical and trabecular bone proteomes, *J. Archaeol. Sci. Rep.* 57 (2024) 104643.
- [32] J.P. Simpson, K.E.H. Penkman, B. Demarchi, H. Koon, M.J. Collins, J. Thomas-Oates, B. Shapiro, M. Stark, J. Wilson, The effects of demineralisation and sampling point variability on the measurement of glutamine deamidation in type I collagen extracted from bone, *J. Archaeol. Sci.* 69 (2016) 29–38.
- [33] D. Mylopotamitaki, F.S. Harking, A.J. Taurozzi, Z. Fagernäs, R.M. Godinho, G. M. Smith, M. Weiss, T. Schüller, S.P. McPherron, H. Meller, J. Cascalheira, N. Bicho, J.V. Olsen, J.-J. Hublin, F. Welker, Comparing extraction method efficiency for high-throughput palaeoproteomic bone species identification, *Sci. Rep.* 13 (2023) 18345.
- [34] H. Xia, D. Zhang, J. Wang, Z. Fagernäs, T. Li, Y. Li, J. Yao, D. Lin, G. Troché, G. M. Smith, X. Chen, T. Cheng, X. Shen, Y. Han, J.V. Olsen, Z. Shen, Z. Pei, J.-J. Hublin, F. Chen, F. Welker, Middle and late Pleistocene Denisovan subsistence at Baishiya karst cave, *Nature* 632 (2024) 108–113.
- [35] S. Brown, M. Kozlikin, M. Shunkov, A. Derevianko, T. Higham, K. Douka, K. K. Richter, Examining collagen preservation through glutamine deamidation at Denisova cave, *J. Archaeol. Sci.* 133 (2021) 105454.
- [36] V. Sinet-Mathiot, G.M. Smith, M. Romandini, A. Wilcke, M. Peresani, J.-J. Hublin, F. Welker, Combining ZooMS and zooarchaeology to study late Pleistocene hominin behaviour at Fumane (Italy), *Sci. Rep.* 9 (2019) 12350.
- [37] N.L. Martisius, F. Welker, T. Dogandžić, M.N. Grote, W. Rendu, V. Sinet-Mathiot, A. Wilcke, S.J.P. McPherron, M. Soressi, T.E. Steele, Non-destructive ZooMS identification reveals strategic bone tool raw material selection by Neandertals, *Sci. Rep.* 10 (2020) 7746.
- [38] B. Nair, I.R. Palomo, B. Markussen, C. Wiuf, S. Fiddymment, M.J. Collins, Parchment glutamine index (PQI): a novel method to estimate glutamine deamidation levels in parchment collagen obtained from low-quality MALDI-TOF data, *Peer Comm. J.* 3 (2023), <https://doi.org/10.24072/pcjournal.230>.
- [39] M. Buckley, S. Whitcher Kansa, S. Howard, S. Campbell, J. Thomas-Oates, M. Collins, Distinguishing between archaeological sheep and goat bones using a single collagen peptide, *J. Archaeol. Sci.* 37 (2010) 13–20.
- [40] S. Brown, N. Wang, A. Oertle, M.B. Kozlikin, M.V. Shunkov, A.P. Derevianko, D. Comeskey, B. Jope-Street, V.L. Harvey, M.P. Chowdhury, M. Buckley, T. Higham, K. Douka, Zooarchaeology through the lens of collagen fingerprinting at Denisova cave, *Sci. Rep.* 11 (2021) 15457.
- [41] S. Brown, T. Higham, V. Slon, S. Pääbo, M. Meyer, K. Douka, F. Brock, D. Comeskey, N. Procopio, M. Shunkov, A. Derevianko, M. Buckley, Identification of a new hominin bone from Denisova cave, Siberia using collagen fingerprinting and mitochondrial DNA analysis, *Sci. Rep.* 6 (2016) 23559.
- [42] S. Brown, D. Massilani, M.B. Kozlikin, M.V. Shunkov, A.P. Derevianko, A. Stoessel, B. Jope-Street, M. Meyer, J. Kelso, S. Pääbo, T. Higham, K. Douka, The earliest Denisovans and their cultural adaptation, *Nat. Ecol. Evol.* 6 (2022) 28–35.
- [43] K. Douka, V. Slon, Z. Jacobs, C.B. Ramsey, M.V. Shunkov, A.P. Derevianko, F. Mafessoni, M.B. Kozlikin, B. Li, R. Grün, D. Comeskey, T. Deviese, S. Brown, B. Viola, L. Kinsley, M. Buckley, M. Meyer, R.G. Roberts, S. Pääbo, J. Kelso, T. Higham, Age estimates for hominin fossils and the onset of the upper Palaeolithic at Denisova cave, *Nature* 565 (2019) 640–644.
- [44] Z. Jacobs, B. Li, M.V. Shunkov, M.B. Kozlikin, N.S. Bolikhovskaya, A. K. Agadjanian, V.A. Uliyanov, S.K. Vasiliev, K. O'Gorman, A.P. Derevianko, R. G. Roberts, Timing of archaic hominin occupation of Denisova cave in southern Siberia, *Nature* 565 (2019) 594–599.
- [45] S. Brown, S. Hebestreit, N. Wang, N. Boivin, K. Douka, K. Korzow Richter, Zooarchaeology by mass spectrometry (ZooMS) for bone material, *Acid Insoluble Protocol v1* (2020), <https://doi.org/10.17504/protocols.io.bf43jyqn>.
- [46] S. Brown, K. Douka, M.J. Collins, K.K. Richter, On the standardization of ZooMS nomenclature, *J. Proteome* 235 (2021) 104041.

- [47] J. Rainer, A. Vicini, L. Salzer, J. Stanstrup, J.M. Badia, S. Neumann, M.A. Stravs, V. Verri Hernandez, L. Gatto, S. Gibb, M. Witting, A modular and expandable ecosystem for metabolomics data annotation in R, *Metabolites* 12 (2022) 173.
- [48] A. Savitzky, M.J.E. Golay, Smoothing and differentiation of data by simplified least squares procedures, *Anal. Chem.* 36 (1964) 1627–1639.
- [49] C.G. Ryan, E. Clayton, W.L. Griffin, S.H. Sie, D.R. Cousens, SNIP, a statistics-sensitive background treatment for the quantitative analysis of PIXE spectra in geoscience applications, *Nucl. Inst. Methods Phys. Res. B* 34 (1988) 396–402.
- [50] M. Schmid, D. Rath, U. Diebold, Why and how Savitzky-Golay filters should be replaced, *ACS Meas. Sci. Au* 2 (2022) 185–196.
- [51] A. Janzen, K.K. Richter, O. Mwebi, S. Brown, V. Onduso, F. Gatwiri, E. Ndiema, M. Katongo, S.T. Goldstein, K. Douka, N. Boivin, Distinguishing African bovines using Zooarchaeology by mass spectrometry (ZooMS): new peptide markers and insights into Iron age economies in Zambia, *PLoS One* 16 (2021) e0251061.
- [52] M. Körsen, A. Pelster, K. Dreisewerd, H.F. Arlinghaus, 3D ToF-SIMS analysis of peptide incorporation into MALDI matrix crystals with sub-micrometer resolution, *J. Am. Soc. Mass Spectrom.* 27 (2016) 277–284.
- [53] J.P. Simpson, M. Fascione, E. Bergström, J. Wilson, M.J. Collins, K.E.H. Penkman, J. Thomas-Oates, Ionisation bias undermines the use of matrix-assisted laser desorption/ionisation for estimating peptide deamidation: synthetic peptide studies demonstrate electrospray ionisation gives more reliable response ratios, *Rapid Commun. Mass Spectrom.* 33 (2019) 1049–1057.
- [54] L.E. Lee-Michaelis, S.C. Münzel, K. Kitagawa, M. Saunders, F. Yang, B.M. Starkovich, N.J. Conard, S. Brown, The influence of fragmentation on a combined Zooms and Zooarchaeological study at Geißenklösterle cave in the Swabian Jura, Southwestern Germany, *SSRN Electron. J.* (n.d.). <https://ssrn.com/abstract=5357240> or <https://doi.org/https://doi.org/10.2139/ssrn.5357240>.
- [55] T. Deviese, I. Karavanić, D. Comeskey, C. Kubiak, P. Korlević, M. Hajdinjak, S. Radović, N. Procopio, M. Buckley, S. Pääbo, T. Higham, Direct dating of Neanderthal remains from the site of Vindija cave and implications for the middle to upper Paleolithic transition, *Proc. Natl. Acad. Sci. USA* 114 (2017) 10606–10611.
- [56] D. Myopotamitaki, M. Weiss, H. Fewlass, E.I. Zavala, H. Rougier, A.P. Sümer, M. Hajdinjak, G.M. Smith, K. Ruebens, V. Sinet-Mathiot, S. Pederzani, E. Essel, F. S. Harking, H. Xia, J. Hansen, A. Kirchner, T. Lauer, M. Stahlschmidt, M. Hein, S. Talamo, L. Wacker, H. Meller, H. Dietl, J. Orschiedt, J.V. Olsen, H. Zeberg, K. Prüfer, J. Krause, M. Meyer, F. Welker, S.P. McPherron, T. Schüller, J.-J. Hublin, Homo sapiens reached the higher latitudes of Europe by 45,000 years ago, *Nature* 626 (2024) 341–346.
- [57] C.A. Hofman, T.C. Rick, J.M. Erlandson, L. Reeder-Myers, A.J. Welch, M. Buckley, Collagen fingerprinting and the earliest marine mammal hunting in North America, *Sci. Rep.* 8 (2018) 10014.
- [58] M. Buckley, J. Herman, Species identification of late Pleistocene bat bones using collagen fingerprinting, *Int. J. Osteoarchaeol.* 29 (2019) 1051–1059.
- [59] E.-M. Oldfield, M.S. Dunstan, M.P. Chowdhury, L. Slimak, M. Buckley, AutoZooMS: integrating robotics into high-throughput ZooMS for the species identification of palaeontological remains at Grotte Mandrin, France, *Archaeol. Anthropol. Sci.* 17 (2025), <https://doi.org/10.1007/s12520-024-02073-7>.
- [60] M. Gu, M. Buckley, Semi-supervised machine learning for automated species identification by collagen peptide mass fingerprinting, *BMC Bioinformatics* 19 (2018), <https://doi.org/10.1186/s12859-018-2221-3>.
- [61] S. Hickinbotham, S. Fiddymant, T.L. Stinson, M.J. Collins, How to get your goat: automated identification of species from MALDI-ToF spectra, *Bioinformatics* 36 (2020) 3719–3725.
- [62] E.I. Végh, K. Douka, SpecieScan: semi-automated taxonomic identification of bone collagen peptides from MALDI-ToF-MS, *Bioinformatics* 40 (2024), <https://doi.org/10.1093/bioinformatics/btae054>.
- [63] F. Bray, H. Touzet, PAMPA: A Software for Peptide Markers and Taxonomic Identification for ZooMS Samples in Archaeology and Paleontology, *bioRxiv* (2025), <https://doi.org/10.1101/2025.03.26.645441>.
- [64] T. Lawrence, M. Buckley, Sequence-driven species identification of ZooMS collagen peptide mass fingerprints, *J. Proteome* 321 (2025) 105525.
- [65] M.V. Shunkov, A.Y. Fedorchenko, M.B. Kozlikin, A.P. Derevianko, Initial upper Palaeolithic ornaments and formal bone tools from the east chamber of Denisova cave in the Russian Altai, *Quat. Int.* (2020), <https://doi.org/10.1016/j.quaint.2020.07.027>.
- [66] M.V. Shunkov, N.A. Kulik, M.B. Kozlikin, E.V. Sokol, The phosphates of pleistocene–holocene sediments of the eastern gallery of denisova cave, *Dokl. Earth Sci.* (2018), https://doi.org/10.1134/S1028334X18010270&casa_token=yNG0n6JzF08AAAAA:o4mb3G9jPF0IU7IEczx9TVFUSy6tgWhl0l8Y8E5ipmpexMQAYlDdQ2TpD6cT3K04BJAuiCnT6X3_bY..
- [67] E.V. Sokol, M.B. Kozlikin, S.N. Kokh, A.V. Nekipelova, N.A. Kulik, V.A. Danilovsky, P.V. Khvorov, M.V. Shunkov, Phosphate record in Pleistocene–Holocene sediments from Denisova cave: formation mechanisms and archaeological implications, *Minerals (Basel)* 12 (2022) 553.
- [68] M.W. Morley, P. Goldberg, V.A. Uliyanov, M.B. Kozlikin, M.V. Shunkov, A. P. Derevianko, Z. Jacobs, R.G. Roberts, Hominin and animal activities in the microstratigraphic record from Denisova cave (Altai Mountains, Russia), *Sci. Rep.* 9 (2019) 13785.
- [69] E.I. Zavala, Z. Jacobs, B. Vernot, M.V. Shunkov, M.B. Kozlikin, A.P. Derevianko, E. Essel, C. de Filippo, S. Nagel, J. Richter, F. Romagné, A. Schmidt, B. Li, K. O’Gorman, V. Slon, J. Kelso, S. Pääbo, R.G. Roberts, M. Meyer, Pleistocene sediment DNA reveals hominin and faunal turnovers at Denisova cave, *Nature* 595 (2021) 399–403.
- [70] G.M. Smith, K. Ruebens, V. Sinet-Mathiot, F. Welker, Towards a deeper integration of ZooMS and zooarchaeology at Palaeolithic sites: current challenges and future directions, *PaleoAnthropology* 2024 (2024) 186–211.
- [71] K. Ruebens, E. Discamps, G.M. Smith, J. Hublin, Integrating ZooMS and Zooarchaeology to Assess the châtelperronian and Carnivore Occupations at Cassenade (Dordogne, France), *PaleoAnthropology* 2024 (2) (2024) 263–281. <http://hal.science/hal-04768810/>.