Low-Invasive Sampling Method with Tape-Disc Sampling for the Taxonomic Identification of Archeological and Paleontological Bones by Proteomics

Isabelle Fabrizi, Stéphanie Flament, Claire Delhon, Lionel Gourichon, Manon Vuillien, Tarek Oueslati, Patrick Auguste, Christian Rolando, Fabrice Bray

Low-invasive sampling method with tape disc sampling for the taxonomic identification of archaeological and paleontological bones by proteomics

*Isabelle Fabrizi[1], Stéphanie Flament[1], Claire Delhon[2], Lionel Gourichon[2], Manon Vuillien[2], Tarek Oueslati[3], Patrick Auguste[4], Christian Rolando[1,5], Fabrice Bray[1,]**

[1] Univ. Lille, CNRS, UAR 3290 - MSAP - Miniaturisation pour la Synthèse, l'Analyse et la Protéomique, F-59000 Lille, France

[2] Univ. Côte d'Azur, CNRS, UMR 7264 – CEPAM - Cultures et Environnements Préhistoire, Antiquité, Moyen Âge, F-06300 Nice, France

[3] Univ. Lille, CNRS, UMR 8164 – HALMA - Histoire, Archéologie et Littérature des Mondes Anciens, F-59000 Lille, France

[4] Univ. Lille, CNRS, UMR 8198 – EEP - Evolution, Ecology and Paleontology, F-59000 Lille, France

[5] Shrieking Sixties, Villeneuve d'Ascq F-59650, France

KEYWORDS. Palaeoproteomics, Taxonomy, Dermatological skin tape-discs sampling, MALDI FTICR MS, LC-MS/MS.

ABSTRACT

Collagen from paleontological bones is an important organic material for isotopic measurement, radiocarbon and paleoproteomics analyzes, to provide information on diet, dating, taxonomy and phylogeny. Current paleoproteomics methods are destructive and require from a few milligrams to several tens of milligrams of bone for analysis. In many cultures, bones are raw materials for artefact which are conserved in museum which hampers to damage these precious objects during sampling. Here, we describe a low-invasive sampling method that identifies collagen, taxonomy and post-translational modifications from Holocene and Upper Pleistocene bones dated to 130,000 and 150 BC using dermatological skin tape-discs for sampling. The sampled bone micro-powders were digested following our highly optimized eFASP protocol, then analyzed by MALDI FTICR MS and LC-MS/MS for identifying the genus taxa of the bones. We show that this low-invasive sampling does not deteriorate the bones and achieves results similar to those obtained by more destructive sampling. Moreover, this sampling method can be carried out at archaeological sites or in museums.

INTRODUCTION

Bone is composed of 60% inorganic component, 30% organic component and 10% water.[1] Collagen represent 90% of organic component and is an extremely important material for revealing the past of bones.[2] But sample preparation remains one of the most difficult aspects of paleoproteomics experiments, as the mineral part of the bone made from hydroxyapatite must be removed to access proteins. Removing hydroxyapatite while preserving protein integrity requires special additional steps that are not necessary for other tissue types. Collagen isolation methods are destructive and consumes from a few milligrams to several tens of milligrams of bone. The bone hydroxyapatite is eliminated by acid or basic demineralization and the residual collagen can be gelatinized, before enzymatic digestion and peptide analysis. In paleoproteomics, the most commonly used method is Zooarchaeology by mass spectrometry (ZooMS).[3] This method uses biomarkers peptides detection for taxonomic identification. Others, several methods of preparation have been used such as in solution digestion after proteins extraction, FASP method (Filter Aided Samples Preparation), eFASP (enhanced Filter Aided Samples Preparation) method, an approach based on paramagnetic beads with surfaces modified by resin bearing carboxylate group called SP3 (Single-pot, Solid-phase-enhanced Sample-Preparation), and on beads with weak cation-exchanger called SPIN (Species by Proteome Investigation), for shotgun proteomics workflow for analyzing archaeological bone.[4-10] We previously optimized the sample preparation by eFASP protocol using 96 well plates.[11] All these methods provide answers to various questions, such as taxonomic identification, the proteome of bone, teeth, etc., the metaproteome, phylogeny and post-translational modifications.[2, 12-19] However, the destructive nature of these methods is undesirable when analyzing archaeological material, such as bone remains or rare artifacts made from bone. Several important considerations must be taken into account before destructive sampling of an

artefact is carried out, such as the probability of successful analysis, the choice of sampling technique to minimize traces on the object, the amount of material to be sampled, and the effects of current sampling on future research and artifacts conservation.[20]

The development of a method of species identification that does not damage the object or leaves very little visible trace on the object and which can be used for the analysis of museum collections is still a challenge despite the facts that several methods has been described recently with this objective in mind. One of the first developed method is ammonium bicarbonate buffer extraction without demineralization which was performed on bones to identify taxa, even if the cold acid demineralization allows a better yield during the extraction of proteins.[21] A second method which allows low invasive analysis of samples is based on the triboelectric effect of a PVC (polyvinyl chloride) eraser.[22] This method was originally set up for the analysis of parchment and on other archaeological materials such as bone and ivory.[23-26] The triboelectric effect has been applied to bones contained in plastic storage bags allowing identifying them.[27] In 2019, Kirby *et al*. used polishing films to sample photographs.[28-30] Then Evans *et al.* applied this technique for the taxonomic identification of bone artefacts, with successful results on 5,000-year-old remains.[31] More recently, proteins on skin and bones on the surfaces of cranial bone of a mummified Egyptian from the 26th Dynasty (664–525 BC) has been identified using dermatology grade skin sampling strips.[32] Recently, Hansen et al. compared different minimally invasive sampling methods and showed that bone preservation influences the proteomic result and that the bone surfaces are modified according to the specific sampling method.[33] We describe here the low-invasive proteomics identification of Iron Age, Neolithic and Upper Pleistocene bones from 120,000-150 BC based on sampling with dermatological skin tape discs. Firstly, we adapted our new sensitive digestion method based on both demineralization and digestion in 96 well plates followed by

MALDI FTICR MS analysis which has a sensitivity below the milligram of bones to dermatological skin tape discs sampling. We then compared dermatological skin tape discs sampling with previously describe minimally invasive methods using our optimized protocol. Finally, we showed that dermatological skin tape discs sampling method allows to obtain sufficient collagen to be obtained for correct taxonomic identification without affecting the appearance of the bone demonstrating it can be used in archaeological, paleontological or museum sites.

EXPERIMENTAL SECTION

Chemicals and biochemicals

All aqueous solutions were prepared from ultrapure grade water obtained by water filtration with a two stages Millipore system (Milli-Q® Academic with a cartouches Q-Gard 1 and Progard 2, Merck Millipore, Burlington, Massachusetts, United States). All chemicals, biochemicals and solvents were purchased from Merck (Merck KGaA, Darmstadt, Germany) and used without purification. All solvents were MS analytical grade.

Samples

Neolithic samples came from Tremblay-en-France (Île-de-France, France) (260 - 150 BC), Bouchain (Nord, France) (3200-2900 cal. BC) and the cave of Pertus II (Alpes-de-Haute-Provence, France) (3800-3300 cal. BC).[34-37] Upper Pleistocene samples came from the Sarrasins cave (Loverval, Belgium) (63 ka BC) and Waziers (Nord, France) (132 – 123 ± 8 ka BC).[38] Modern samples came from the laboratories (EEP and HALMA) of the authors specialized in archaeology or paleontology. Samples Mod 1, 3 and 4 were cleaned after maceration with water at 35 °C for 2 days. The sample Mod 2 was cleaned after maceration with water and detergent at 35 °C for 2 days

then were bleached with diluted hydrogen peroxide (5 % v/v, 16 h at 4 °C). All information on bones are given in Supplementary Information 1, Table S1 and pictures of bones are in Supplementary Information 1, Figure S1-4.

Triboelectric rubbing in bag sampling

The sampling method is based on the protocol described by McGrath *et al*.[23] First, entire bone artifacts were placed in a clean, labeled Minigrip 60 µm, 100 × 150 mm (PlanetArcheo, Marcilly-le-Châtel, France) sample bags made of neutral low density polyethylene. The bone artifacts were rubbed gently for 60 s in the sealed bag to create triboelectric friction between the bag and the bone artifact. Bone artifacts were removed from the bag and stored in new, sterile sample bags. 300 µL of warmed AmBic (50 mM ammonium bicarbonate pH 8.8, 65 °C) was pipetted into the empty sample bag and massaged gently for two minutes. The AmBic solution was extracted from the bag and placed in a 1.5 mL Eppendorf™ tubes (Eppendorf, Hamburg, Germany). Gelatinization and in solution trypsin digestion was then applied following the protocol described below.

Eraser sampling by rubbing

An area of 1 cm² from bone artifacts were rubbed with a small polyvinyl chloride (PVC) eraser (Staedtler, UGAP, France) on a flat section of the artifact's surface as described in McGrath *et al.*[23] If the outer surface of the artifact was dirty, the initial eraser shavings were discarded and the area was erased again with a clean section of the eraser. Approximately, 15 mg of eraser shavings were collected and placed into 200 µL of Ambic solution. Samples were vortexed at 2,000 rpm for 5 min (Vortex genie® 2, Scientific Industries, Inc., USA) and the solution was placed in a new 1.5

mL Eppendorf™ tubes. Gelatinization and in solution trypsin digestion were then applied following the protocol described below.

Swab mopping sampling

An area of 1 cm² from bone artifacts was rubbed with a small swab (Foamtec UltraSOLV™ 1700 Series Swabs, Merck Life Science S.A.S, France) on a flat section of the artifact's surface. A different zone was used for each method involving rubbing. The swab was placed in an Eppendorf™ 5 mL tube and 300 µL of AmBic solution were added. Swab was vortexed at a low speed for 5 min and the solution was placed in a new 1.5 mL Eppendorf™ tubes. Gelatinization and in solution trypsin digestion were then applied following the protocol described below.

Ammonium bicarbonate buffer etching sampling

This method was based on Van Doorm *et al.*[21] A piece of bone weighing approximately 50 mg was placed in a 1.5 mL Eppendorf™ tube and covered with AmBic solution for 1 hours. With the piece of bone still in the Eppendorf™ tube, gelatinization and in solution trypsin digestion were then applied following the protocol described below.

Destructive ZooMS sampling

This method was based on Bray *et al.*[11] Approximately 1 mg of bone was removed by scraping with a scalpel blade and demineralized in 100 µL of 5% TFA solution during 4 h at room temperature with vortexing (at 2,000 rpm). The demineralization solution was recovered, 8 µL NaOH 6 M was added to neutralize TFA and then 100 µL of 100 mM AmBic. In solution trypsin digestion was then applied following the protocol described below. The bone powder was rinsed with 200 µL of 50 mM AmBic and the solution was discarded. This step was repeated twice. Then

200 µL AmBic was added on bone powder. Gelatinization and in solution trypsin digestion were then applied following the protocol described below.

Dermatological skin tape-discs sampling on Iron Age, Neolithic and Upper Pleistocene bones

Sample preparation was conducted according to established guidelines for working on ancient protein, to minimize exogenous laboratory contamination.[2] The surface of the bones where the samples were going to be taken was first cleaned with a brush to remove post-depositional deposits. Then the area was cleaned with a wipe soaked with mQ water (Kimberly Clark Kimtech Science, Nanterre, France). The samples were taken after the bone surface had dried. Each tape-disk (D-Squame®, 22 mm diameter, 3.8 cm^2, Monaderm, Principality of Monaco) was pressed manually against the bone for 3 s, before being stripped from the bone. For experiments involving several samplings, each tape-disk was placed at the same place on the bone. Tape-disks were cut into quarters and the four pieces were placed in individual 1.5 mL Eppendorf™ (Eppendorf, Hamburg, Germany). Extraction of bone particles was realized by adding 500 µL of extraction buffer (8 M urea, 100 mM ammonium bicarbonate pH 8.8), followed by 15 min sonification in iced water (0–4 °C) in an ultrasonic bath (Advantage Lab, Switzerland) followed by 1 h on vortex at 4 °C and 2,000 rpm. The tape-disk quarters were scraped in the 1.5 mL Eppendorf™ with a spatula to release all bone particles in the solution and the tape-disk quarters were discarded. The solutions containing the bone particles were combined in a 5 mL Eppendorf™, then concentrated with a 0.5 mL Amicon® ultra centrifugal filters with a cut-off of 10 kDa (EMD Millipore, Darmstadt, Germany) by centrifugation for 20 min at 10,000 g using an Eppendorf™ centrifuge 5430R (Eppendorf, Hambourg, Germany). Before use, 0.5 mL Amicon® ultra centrifugal filters were freshly incubated overnight with the passivation solution containing 5% (v/v) Tween®-20 and were washed in a water bath during 20 min four times before use. After concentration, 100 µL of

denaturation buffer (8 M urea, 50 mM DTT, 100 mM ammonium bicarbonate pH 8.8) were added and the Amicon® was incubated at 4 °C overnight.

eFASP digestion optimized for skin tape-disc sampling

The proteomic method is based on the protocol described by Bray *et al.* which has been optimized for skin tape-disc sampling.[9] The bone powder suspension in the Amicon® obtained in the previous step was centrifuged during 20 min at 10,000 g, then washed with 100 µL of exchange buffer (8 M urea, 100 mM ammonium bicarbonate pH 8.8) and then the exchange buffer was eliminated by centrifugation during 20 min at 10,000 g. The filtrates were discarded, then 200 µL of exchange buffer were added again into the Amicon® filter which was centrifuged. This step was repeated twice. Proteins were alkylated during 1 h at room temperature in the dark using 100 µL of alkylation buffer (8 M urea, 50 mM iodoacetamide, and 100 mM ammonium bicarbonate, pH 8.8). The Amicon® filter was centrifuged for 20 min at 10,000 g and the filtrate was discarded. After the alkylation step, 200 µl of exchange buffer was added to the Amicon® filter which was centrifuged for 20 min at 10,000 g and the filtrate was discarded. 200 µl of AmBic solution (50 mM ammonium bicarbonate pH 8.8) were added to the Amicon® filter and then centrifuged. This step was repeated twice, discarding the filtrate at each step. The Amicon® filter was transferred to a new 1.5 mL microcentrifuge collection tube. 100 µL of AmBic solution and 40 µl of trypsin (0.05 µg/µl, Promega, Madison, USA) were added and incubated with shaking at 400 rpm in a heating block tube (MHR23, Hettich, Netherlands) overnight at 37 °C. After this step, the peptides present in the Amicon® filter were recovered in the lower collection tube by centrifugation during 15 min at 10,000 g. In order to collect a maximum of peptides, the Amicon® filters was washed twice with 50 µl of AmBic solution. The filtrates containing all the peptides were transferred to 1.5 mL Eppendorf™ tubes and were evaporated to dryness at room temperature with a

SpeedVac™ Concentrator (Eppendorf, Hamburg, Germany). Tryptic peptides were desalted on a 96-well plate C18 (Affinisep, Petit-Couronne, France) following the protocol described below.

Gelatinization and in solution trypsin digestion

Following the sampling procedures described above, all samples were gelatinized, except the demineralization solution from the destructive sampling, in AmBic at 65 °C for one hour with shaking at 400 rpm in a heating block tube. Samples were incubated overnight (12–18 h) at 37 °C with 10 µL of trypsin solution (0.05 µg / µl in 50 mM AmBic solution) with shaking at 400 rpm on a heating stirrer MHR23 (Hettich, Tuttlingen, Germany). After digestion, 300 µL of 0.5% acetic acid were added and peptides were purified and desalted using a 96-well plate C18.

Purification of peptides

Briefly, each well of the 96-well plate C18 (Affinisep, Normandy, France) was washed two times with 100 µL of acetonitrile (ACN) followed by a second washing step repeated 3 times with 100 µL of 80% ACN, H_2O 0.5% acetic acid and a third washing repeated also 3 times with with 100 µL of H_2O alone 0.5% acetic acid. Tryptic peptides from bone powder were resuspended in 200 µL of a H_2O, 0.5% acetic acid solution. Tryptic peptides were transferred to the C18 96-well plate and eluted with a vacuum manifold. Each well was washed 6 times with 100 µL of H_2O, 0.5% acetic acid. Peptides were recovered in a V-bottom well collecting plate (Sarstedt, Nümbrecht, Germany) using 100 µL of a 80% ACN, 0.1% acetic acid solution followed by 100 µL of ACN. The plate was evaporated on TurboVap 96 Evaporator (Caliper LifeScience, Hopkinton, USA). For mass spectrometry analysis, the samples were dissolved again in 10 µl of solvent A of LC (see below). The concentration was then estimated by measuring the OD at 215 nm using 1 µl of the

solution using a droplet UV spectrometer (DS-11+, Denovix, Wilmington, USA). Samples were diluted at a concentration of 1 µg/µL before LC-MS/MS analysis.

Liquid chromatography-tandem mass spectrometry

LC-MS/MS analyses were performed on an Orbitrap Q Exactive plus mass spectrometer hyphenated to a U3000 RSLC Microfluidic HPLC System (ThermoFisher Scientific, Waltham, Massachusetts, USA). 1 µl of the peptide mixture at a concentration of 1 µg/µL was injected with solvent A (5% acetonitrile, 0.1% formic acid v/v) for 3 min at a flow rate of 10 µl.min-1 on an Acclaim PepMap100 C18 pre-column (5 µm, 300 µm i.d. × 5 mm) from ThermoFisher Scientific. The peptides were then separated on a C18 Acclaim PepMap100 C18 reversed phase column (3 µm, 75 mm i.d. × 500 mm), using a linear gradient (5-40%) of solution B (75% acetonitrile and 0.1% formic acid) at a rate of 250 nL.min-1. The column was washed with 100% of solution B during 5 minutes and then re-equilibrated with buffer A. The column and the pre-column were placed in an oven at a temperature of 45°C. The total duration of the analysis was 140 min. The LC runs were acquired in positive ion mode with MS scans from *m/z* 350 to 1,500 in the Orbitrap mass analyzer at 70,000 resolution at *m/z* 200. The automatic gain control was set at $1e10^6$. Sequentially MS/MS scans were acquired in the high-energy collision dissociation cell for the 15 most-intense ions detected in the full MS survey scan. Automatic gain control was set at $5e10^5$, and the normalized collision energy was set to 28 eV. Dynamic exclusion was set at 90 s and ions with only 1 charge and more than 8 charges were excluded.

MALDI FTICR analysis

Desalted peptides (1 µL) were deposited on 384 ground steel MALDI plates (Bruker Daltonics, Bremen, Germany), then 1 µL of HCCA matrix at 10 mg/mL in ACN/H_2O 70:30 v/v 0.1% TFA

was added for each sample spot and dried at room temperature. MALDI FTICR experiments were carried out on a Bruker 9.4 Tesla SolariX XR FTICR mass spectrometer controlled by FTMS Control software and equipped with a CombiSource and a ParaCell (Bruker Daltonics, Bremen, Germany). A Bruker Smartbeam-II Laser System was used for irradiation at a frequency of 1,000 Hz using the "Minimum" predefined shot pattern. MALDI FTICR spectra were generated from 500 laser shots in the *m/z* range from 693.01 to 5,000. 2M data points were used per spectrum which corresponds to a transient duration of 5.0332 s. Twenty spectra were averaged. The transfer time to the ICR cell was set to 1.2 ms and the quadrupole mass filter set at *m/z* 600 was operated in RF-only mode.

Bioinformatics

MS raw data from MALDI FTICR were processed using DataAnalysis 5.0. For deisotoping and extracting the monoistopic peaks the SNAP algorithm was employed with the following parameters of S/N > 3 and quality 0.6. The procedure for the deamidation value calculation from MALDI FTICR was based on Bray *et al.* 2023.[11] The identification is supported by all peptide markers presented in previous reports.[3, 39-44]

LC-MS/MS proteomics data were firstly processed with Mascot v 2.5.1 against NCBI database mammalian (NCBI_2022_01, 9,016,701 sequences) and Aves (NCBI_2022_01, 2,494,584,292 sequences). Three missed cleavages, 10 ppm mass error for MS and MS/MS were applied. Cysteine carbamidomethylation (+57.02 Da) was set as fixed modification. Methionine oxidation, hydroxylation of proline (+15.99 Da) and asparagine, glutamine deamidation (+0.98 Da) were selected as variable modifications.

A second bioinformatics analysis was performed with PEAKS X plus (Bioinformatics software, Waterloo, Canada) against a home-made database containing 1,765 collagen sequences (collagens sequences with signal peptide and propeptides) extracted from NCBI database (All_Collagen, from Bray *et al.* 2023) restricted to Mammalian and Aves.[11] Precursor's mass tolerance was fixed to 10 ppm and fragment ion mass tolerance to 0.05 Da. Three missed cleavages were allowed. The same post-translational modifications (PTMs) above were allowed plus hydroxylation of amino acids (RYFPNKD) (+15.99 Da) as variable modifications. Five variable PTMs were allowed per peptides. PEAKS PTM and SPIDER ran with the same parameters. Results were filtered using the following criteria: protein score -10logP ≥ 20`, 1% peptide False Discovery Rate (FDR)`, PTM with Ascore = 20`, mutation ion intensity = 5% and Denovo ALC ≥ 50%. Peptides with amino acids substitutions was filtered with a minimal intensity threshold set as 1e107. Peptides identified on collagen I alpha 1 and I alpha 2 were aligned against the NCBI non redundant protein sequences (all non-redundant GenBank CDS translations + PDB + SwissProt + PIR + PRF`, excluding environmental samples from WGS projects`, which contains 308`,570`,119 sequences) to find similarity with BLASTp (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins). The scoring parameter alignment used BLOSUM62 matrix.[45] The specific peptides were validated with a score of 100% of identity and full query coverage.

The mass spectrometry proteomics data have been deposited on the ProteomeXchange Consortium (http://proteomecentral. proteomexchange.org) via the PRIDE partner repository with the data set identifier PXD044039.[46]

Calculation of deamidation and hydroxyproline level

MS raw data were also processed by MaxQuant (MQ) software 1.5.8.3 (https://www.maxquant.org/) to further investigate about deamidation level of the proteins

identified.[47] In this search, for each identified species specific COL1A1 and COL1A2 sequences with signal peptide and propeptides from Uniprot and NCBi were used.[11] Database search was carried out using the following parameters: (i) full tryptic peptides with a maximum of 3 missed cleavage sites; (ii) cysteine carbamidomethylation as a fixed modification and (iii) oxidation of methionine, the transformation of *N*-terminal glutamine and *N*-terminal glutamic acid residue to pyroglutamic acid, hydroxyproline oxidation, and the deamidation of asparagine and glutamine as variable modifications. Match type was "match from and to" and the decoy mode was set to "revert". PSM (Peptide-Spectrum Matches) and Protein and Site decoy fraction False Discovery Rate (FDR) were set at 0.01 as threshold for peptide and protein identifications. Match between run was used with 2 min for match time windows and 20 min for alignment time windows. Dependent peptides FDR was set at 0.01 All the other parameters were set to their default values.

An estimation of the percentage of deamidation for N and Q residues for each sample was calculated using the freely available command-line script "deamidation" (https://github.com/dblyon/deamidation), which use the MaxQuant "evidence.txt" file. The calculations were done separately for potentially original peptides and potential contaminants peptides as previously reported in Mackie *et al.* 2018.[48]

For the determination of deamidation of mammals bones following the traditional ZooMS approach the COL1A1 peptide named $P1^{m}$ (position COL1A1-508-519) with the sequence $GVQ^{dem}GP^{ox}PGPAGPR$ was used.[3, 49] For bird bones the peptide sequence $GVQ^{dem}GP^{ox}PGPQGPR$ named $P1^{b}$ with the same position on COL1A1 compared to mammals and which appears conserved across all birds, Australian marsupials and some reptiles was used.[44] In MALDI FTICR the percentage of deamidation of the peptide from COL1A1 is calculated simply by dividing the intensity of the monoisotopic peak of the deamidated peptide by the sum

of the intensity of the monoisotopic peak of the non-deamidated and that of the deamidated peptide without any assumption. In LC-MS/MS analyses without database search the same percentage is calculated from the MS spectrum of the relevant peptides identified by their MS/MS spectra and retention times. The MS/MS spectra are given in Supplementary Information 3, Figure S1-S4. The *m/z* of the peptides with and without deamidation used for the calculation of deamidation percentage are presented in Supplementary Information 1, Table S3 which contains also the *m/z* expected in the MALDI FTICR and LC-MS/MS analyses for native and deamidated peptides

An estimation of the percentage of hydroxyproline residues for each sample was calculated with the PEAKS X+ PTM results. The calculation method is the same as for the calculation of deamidation on the COL1A1 peptide from MALDI FTICR MS data. (1) The percentage of hydroxyproline for a given peptide is equal to the ratio in percent of the intensities of the hydroxyproline modified peptide divided by the sum of the native peptide plus the hydroxyproline modified peptide. The percentage of hydroxyproline is calculated for each peptide from COL1A1 and COL1A2 of the species identified then the average is calculated according to equation (1).

$$\text{Percentage of hydroxyproline} = \text{average}\left(\frac{Intensity\ of\ modified\ peptide}{Intensity\ of\ modified\ peptide + intensity\ of\ native\ peptide}\right) \times 100 \quad (1)$$

An alternative method for calculating the percentage of hydroxyproline residue may be used according to equation 2. This percentage is calculated by dividing the number of identified hydroxyprolines by the total sum of the number of identified hydroxyprolines plus number of identified prolines by LC-MS/MS.

$$\text{Percentage of hydroxyproline} = \frac{Number\ of\ modified\ peptides}{Number\ of\ modified\ peptides + number\ of\ native\ peptides} \times 100 \quad (2)$$

The percentage of deamidation on COL type 1 for N and Q residues from Maxquant, the percentage of deamidation of the peptide from COL1A1 and COL1A2 for each sample, the

percentage of hydoxyproline residues from PEAKS X+ by using method (1) and (2) are presented in Supplementary Information 4, Table S1.

RESULTS AND DISCUSSION

Bone surface modifications induced by the different samplings

The classical methods of sampling bones in archaeology and paleontology is performed conventionally by destructive sampling with a miniaturized grinder or a scalpel under conditions that avoid contaminations.[2, 50] However several low invasive samplings have been introduced recently (i) the triboelectric rubbing in bag sampling, (ii) the PVC eraser sampling by rubbing, (iii) the swab mopping sampling and (iv) the ammonium bicarbonate buffer etching sampling. Foremost, we compared the effect on the surface of modern *Bos taurus* (samples Mod 1) of these low invasive samplings except ammonium bicarbonate buffer etching which works on small fragments and of the destructive ZooMS method, with dermatological skin tape-disc stripping. There is no visible effect on the surface of the bone by triboelectric rubbing and PVC eraser sampling methods (Supplementary Information 1, Figure S5 A, B, C, D). A whitish deposit is observed after mopping with a swab (Supplementary Information 1, Figure S5 E, F). Sampling with scalpel blade on the surface of the bone removes material and leaves a mark on the bone surface (Supplementary Information 1, Figure S5, G, H). Dermatological skin tape-disc sampling was tested on the same mandible from modern *Bos taurus* (Figure 1) to compare the effect of the sampling on bone surface. The skin tape-disc measures 22 mm in diameter. It is placed on the surface of the cleaned bone and pressure is applied manually to the patch for 3 seconds. Then the skin tape-disc is gently removed using clamps. When several patches are used the following patches are placed at the same place to avoid affecting a large area of bone. Figure 1, A shows a

photograph of *Bos taurus* mandible with the dermatological skin tape-discs attached on the surface and Figure 1, B after removing it.

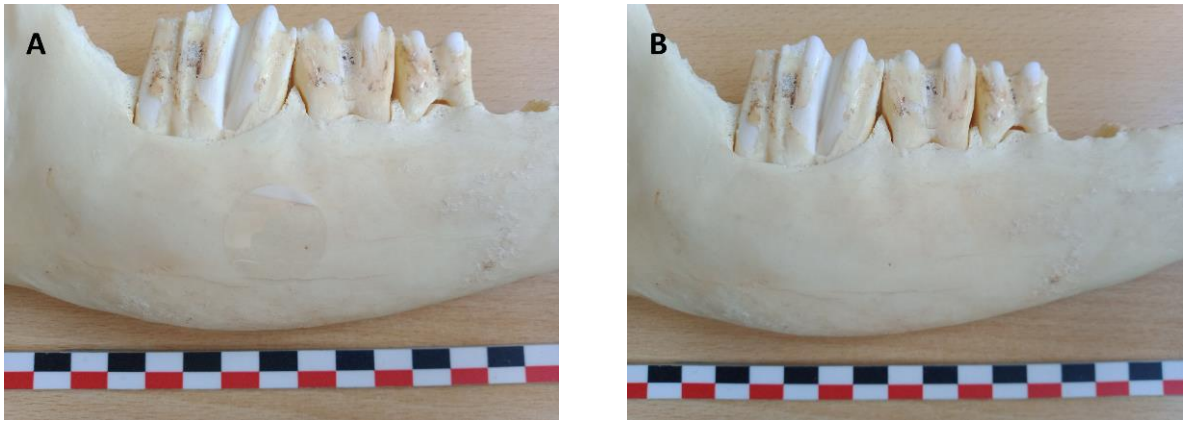

Figure 1. Image of modern *Bos taurus* mandible (Mod 1) with the dermatological skin tape-discs attached to the surface after pressing it (A) and after stripping (B). The black, white and red marks on the scale bar have a length of 1 cm.

The structure of the bone is not affected by stripping with 1 or 5 dermatological skin tape-disc, but after stripping with 10 and 20 dermatological skin tape-discs, the appearance of the bone surface is modified (Supplementary Information 1, Figure S5, I, J, K, L). The amount of material removed by our low sampling method is very small but it can leave a mark on the bone such as Sinet-Mathiot *at al.* showed in her study on the PVC eraser method, even if in our study, the PVC eraser method left no trace, but this depends on the state of conservation of the bone.[25] It is important to take samples from areas where there is no trace of fracturing, cutting to avoid erasing traces of human activity. Supplementary Information 1, Figure S6, show the thin layer on the dermatological skin tape-disc after the first stripping and the powder obtained after scraping of the tape-disk which corresponding to 1 mg of bone powder. Olesen *et al*. showed that the use of an adhesive strip tape removes 1.3 µm of epidermis by stripping.[51]

We tested next our methodology on a Neolithic bone. Supplementary 1, Figure S7, shows the effect of the sampling method on wild boar tibia from the Bouchain site (sample Bou 1). The image shows no effect on the surface of the archeological bone after triboelectric rubbing in forced bag sampling and eraser sampling by rubbing (Supplementary 1, Figure S7, A, B, C, D). With the swab methods, a slight abrasion was observed (Supplementary 1, Figure S7, E, F). With the destructive ZooMS sampling, the sampling location is visible (Supplementary 1, Figure S7, G, H). On the contrary, the dermatological skin tape-discs sampling leaved no visible trace on the surface of the bone (Supplementary 1, Figure S7, I, J).

Optimization of the digestion for dermatological skin tape-disc sampling

During sampling, each dermatological skin tape-disc was cut in four part which were deposited in a 1.5 mL Eppendorf® so that the extraction buffer can cover the disc quarters to recover the bone powder. The eFASP digestion was used to avoid the loss of bone powder present on the skin tape-discs and increase the numbers of identified peptides and so the protein coverage and their identification. We wish to emphasize that this method does not require the acidic demineralization and gelatinization steps of the sample which preserves all existing post-translational modification (PTM) and avoid the creation of artifactual ones. When the stripping was performed with 5 dermatological skin tape-disc each disc was extracted separately and the analysis was performed both on individual digest or on the 5 combined digests for increasing the sensitivity. Figure 2, shows the MALDI FTICR spectrum of the digestion of the combined 5 dermatological skin tape-discs for modern *Bos taurus* (sample Mod 1) with zoom on specific ZooMS peptides markers. The resolution of molecular peak of the mammal peptide $P1^{m}$ (for nomenclature see [3, 49]) which is used to measure the deamidation at *m/z* 1105.574 was 315,000. Figure 2 shows that the glue present on the patch was not found in the MALDI FTICR spectrum after digestion. The same observation

was made for or the LC-MS/MS mass spectrometry analyses. Indeed, the digestion in Amicon® filters enable the glue polymer if any to be eliminated during the washing stages.

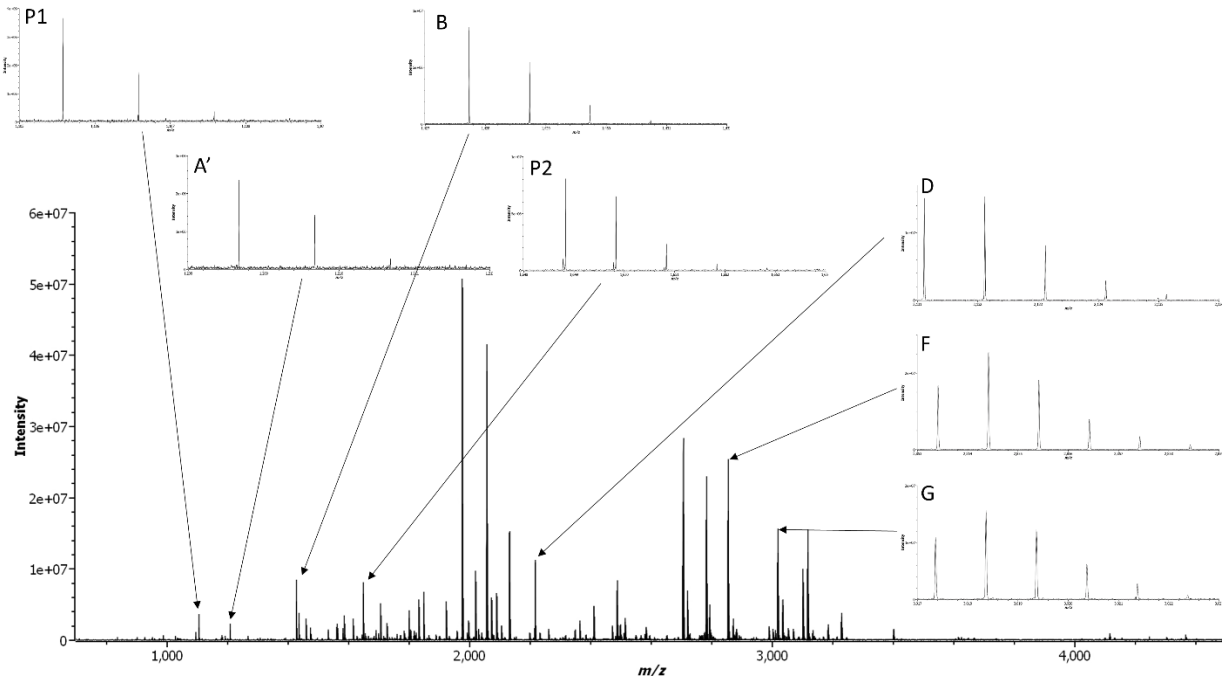

Figure 2. MALDI FTICR spectrum of *Bos taurus* bone (Mod 1) tryptic digest using 5 dermatological skin tape-discs. Zoom on 6 ZooMS peptide markers (P1, A', B, P2, D, F, G). The *m/z* of each peptides were indicated in the Supplementary Information 4, Table S1.

Each spectrum obtained from the individual and combined 5 dermatological skin tape-discs allowed to identify the genus *Bos* by the identification of the 12 ZooMS markers peptides (Supplementary Information 4, Table S2). Thus, to obtain a robust identification and low damage to the bones, the dermatological skin tape-disc sampling were limited to 5 tape-discs in the following studies.

Comparison with other destructive and low invasive methods

We compared dermatological sampling with 1 and 5 skin tape-discs with the 5 other methods (i) triboelectric rubbing in forced bag, (ii) rubbing with a PVC eraser, (iii) mopping with swab, (iv) etching with ammonium bicarbonate buffer, and (v) conventional sampling by the destructive method taking 1 mg of bone on the 4 modern samples and 3 archaeological samples (Bou 1, Bou 2, Trem 2). For the destructive method, 2 fractions were analyzed, the acid demineralization fraction and the bone powder one. A major difference between our method and others is the absence of demineralization and gelatinization. The amount of peptide obtained after digestion evaluated by the absorbance at 215 nm is indicated in Supplementary Information 1, Table S2. The entries for 1 and 5 dermatological skin tape-disc sampling show that increasing the number of stripping increased the amount of peptide after digestion. In the literature Multari *et al*. showed that no direct relationship exists between the number of strips used for sampling the skin on the forearm of a volunteer and the number of identified peptides.[32] The authors indicate that the problem stems from overcrowding in the tubes that occurred during the extraction process. For avoiding this saturation during extraction, we put each strip in a separate Eppendorf™ and extracted the bone powder and proteins from each skin tape-disc before combining them together for digestion. The amount of peptides after digestion for the eraser method, acid fraction were similar to using 1 dermatological skin tape-discs. The amount of peptides was the lowest for the forced bag and ammonium bicarbonate etching methods. The mopping with swabs provided an amount equivalent to 5 dermatological skin tape-discs. Bone powder fraction of 1 milligram of bone yielded a greater amount of peptides. As stripping by 5 dermatological skin tape-discs method gave a better result than 1 dermatological skin tape-discs without modifying the appearance of the bone it was kept for further analyses.

The peptides obtained by digesting with trypsin the bone powder obtained by stripping with 5 skin tape-discs, the 4 low destructive samplings and the destructive ZooMS methods of 4 modern and 3 Neolithic bones were analyzed by MALDI FTICR. For each method, ZooMS markers peptide were identified and allowed to identify the taxonomic genus (Supplementary Information 4, Table S2). The number of peptide markers for each sample and method is given in Supplementary Information 1, Figure S8. As expected, marker peptides were more easily detected in modern than Neolithic samples and their number was higher. The triboelectric charge methods (Eraser and forced bag) produced low quality spectra on archaeological samples but allowed taxa to be identified. Fewer marker peptides were identified compared to other methods. Several studies have shown that these methods can be used on ancient objects such as the St. Lawrence Iroquoian bone points (middle of the 14th to the late 16th centuries AD) analyzed by McGrath *et al*. or the Neanderthal bone artifacts analyzed by Martisius *et al.*[23, 27] Coutu *et al*. noted that the surface of the bones can influence the results. If the bones are smooth and well preserved, the triboelectric charge methods have difficulty recovering bone particles, resulting in a low quantity of collagen and spectra with few marker peaks.[52] However if the sample is crumbled, the forced bag method is most suitable.[31] The type of plastic used have also an impact on the results as the triboelectric charge density varied by more than three order of magnitudes from polytetrafluoroethylene (PTFE) or polyethylene(PE) to polyethyleneterephthalate(PET).[53] For the swab mopping method, the number of marker peptides identified per sample was between 8 and 12. This method was never used to collect material but rather to clean objects. Surface analysis by microscopy showed a deposit of material on the surface of the bone. The ammonium bicarbonate buffer etching method produced good results even though only 6 marker peptides were identified in the Trem2 sample. For the other samples, the number of peptides was 12 or 11. This method is less efficient when the

bones are highly mineralized, which reduces extraction efficiency. Naihui *et al.* showed that the addition of a demineralization step increased the number of identifiable samples.[54] The results of the destructive ZooMS method for the acid fraction and bone powder fractions allowed better identification and detection of peptides markers than ammonium bicarbonate buffer etching method. In all the samples the 5 dermatological skin tape-disc stripping and the destructive sampling allowed identifying the most peptide markers. The 5 dermatological skin tape-disc stripping and destructive ZooMS method allowed the identification of the genus but the 5 dermatological skin tape-disc stripping causes little degradation of the samples contrary to destructive ZooMS method. Therefore, we then applied our dermatological skin tape-disc method to 19 bones from Holocene and Upper Pleistocene bones dated to 130,000 and 150 BC.

Tape stripping method applied to palaeontological and archaeological bones: taxa identification

We applied our minimal invasive sampling approach based on dermatological skin tape-disc stripping on 19 bones from Iron Age, Neolithic and Upper Pleistocene bones dated from 130,000 to 100 BC. Firstly, after digestion, a MALDI FTICR analysis was achieved. Table S1 in Supporting Information 4 contains the identified peptide biomarkers. The taxonomy of each bone has been identified by looking at the 12 ZooMS peptide biomarkers. MALDI FTICR MS analyses identified the taxonomy of all of the 19 samples. Despite the small amount removed by the tape-disc and the age of the bones, it is very easy to identify the taxonomy.

Only one sample Wa24 was identified as *Ursus* sp. while the identification made from osteomorphology by the paleontologists is *Cervus* sp. Although we used a MALDI FTICR MS, the more commonly used MALDI TOF MS can also be employed. It should be noted that the main difference between the two instruments is the 100 times higher resolution of the MALDI FTICR

MS, which allows for the unambiguous detection of monoisotopic peak of native and deamidated peptides allowing the precise quantification of the deamidation rate.

The same samples were also analyzed by LC-MS/MS to validate the MALDI FTICR MS identifications and check whether the bone proteome is not affected by the sampling method. Sample identification was performed using the NCBI mammalian and aves database which contains protein sequences from several sources, including annotated coding region translations in GenBank, RefSeq and TPA, as well as SwissProt records. The mammalian and aves databases were chosen because MALDI identifications only identified mammal and bird species in the same way as morphological identifications by archaeologists and palaeontologists. [3, 39-44] In this database, there are few protein sequences of extinct species such as cave bears and aurochs which complicates the identification of extinct species. To allow their identification, it is necessary to use the closest species at the phylogenetic level. For example, for the aurochs *Bos primigenius*, it is necessary to use the sequences of its closest ancestor which is *Bos taurus*.

LC-MS/MS proteomic analysis of Iron age, Neolithic or Upper Pleistocene bones indicates that the major proteins identified are type I cytoskeleton keratin, type II cytoskeleton keratin, and type I collagen as in the control bone of modern *Bos taurus*. This information shows that the use of the patch allows the extraction of bone proteins despite the fossilization of the bone. The average number of identified peptides for the best identified protein (collagen 1 alpha 1) on modern bones is 645, 458 on the Iron age, Neolithic samples and 190 on the Upper Pleistocene samples (Table 1). The preservation of the Iron age, Neolithic samples in the Tremblay-en-France and Bouchain, Pertus sites is exceptional, which explains the number of significant peptides identified in the samples. Sequence coverage for collagen 1 alpha 1 varies between 59, 50 and 40% for modern, Iron age, Neolithic and Upper Pleistocene samples respectively. The number of peptides and

sequence coverage are lower in older bones due to collagen degradation over time. Supplementary Information 2 shows the identification of all proteins in the analyzed samples. The number of peptides for modern *Struthio camelus* samples is lower than others modern samples due to the processing of the bone. Indeed, this bone has been boiled to remove the flesh elements then cleaned with hydrogen peroxide. Despite this treatment, the analysis with the dermatological skin tape-discs and the preparation allows the identification of many peptides. This shows that the sampling method and proteomic analysis could be used to study bones contained in museums which have been cleaned sometimes with aggressive protocols.

Table 1. LC-MS/MS results for modern, Holocene and Upper Pleistocene bones sampled with dermatological skin tape-discs stripping. The table contains name of sample, taxonomic identification by archaeozoologists and paleontologists, the species of best identified protein collagen alpha-1(I) chain (COL1A1) from NCBI database, its accession number, Mascot score, number of peptides and coverage.

Name	ID archaeologist	Protein	Species form NCBI database	Accession number	Mascot score	Peptides	Coverage (%)
Mod 1	*Bos taurus*	Collagen alpha-1(I) chain	*Bos taurus*	AAI05185.1	9655	624	66
Mod 2	*Struthio camelus*	Collagen alpha-1(I) chain	*Struthio camelus*	XP_009685373.1	5192	254	56
Mod 3	*Capra ibex*	Collagen alpha-1(I) chain	*Capra hircus*	XP_017920382.1	11187	646	62
Mod 4	*Cervus elaphus*	Collagen alpha-1(I) chain	*Cervus canadensis*	XP_043327093.1	22864	1044	59
PE-F01	*Bos taurus*	Collagen alpha-1(I) chain	*Bos taurus*	AAI05185.1	12118	831	60
PE-F21	*Capra hircus*	Collagen alpha-1(I) chain	*Capra hircus*	XP_005678993.1	6701	396	58
PE-F46	*Sus* sp.	Collagen alpha-1(I) chain	*Sus scrofa*	XP_020922812.1	6613	488	52
PE-F55	*Ovis aries*	Hypothetical protein JEQ12_002924	*Ovis aries*	KAG5203341.1	7575	512	49
PE-F72	*Cervus elaphus*	Collagen alpha-1(I) chain	*Cervus canadensis*	XP_043327093.1	5175	365	45

Bou1	*Sus scrofa*	Collagen alpha-1(I) chain	*Sus scrofa*	XP_020922812.1	6272	422	49
Bou2	*Bos taurus*	Collagen alpha-1(I) chain	*Bos taurus*	AAI05185.1	6637	399	57
Bou3	*Cervus elaphus*	Collagen alpha-1(I) chain	*Cervus canadensis*	XP_043327093.1	6456	359	63
Bou4	*Castor fiber*	Collagen alpha-1(I) chain	*Castor canadensis*	JAV39636.1	4707	350	51
Bou5	*Capreolus capreolus*	Hypothetical protein FD754_009325	*Muntiacus muntjac*	KAB0365169.1	8645	567	53
Bou6	*Sus scrofa*	Collagen alpha-1(I) chain	*Sus scrofa domesticus*	BAX02568.1	5357	350	62
Bou7	*Bos taurus*	Collagen alpha-1(I) chain	*Bos taurus*	AAI49096.1	3639	250	44
Trem 1	*Equus* sp.	Collagen alpha-1(I) chain	*Equus asinus*	ACM24774.1	6219	400	49
Trem 2	*Bos taurus*	Collagen alpha-1(I) chain	*Bos taurus*	AAI05185.1	7388	410	62
Lov11	*Ursus spelaeus*	Collagen alpha-1(I) chain	*Ursus maritimus*	XP_040496745.1	4157	293	47
Wa21	*Equus* sp.	Collagen alpha-1(I) chain	*Equus asinus*	ACM24774.1	2349	178	45

Wa22	*Bos primigenius*	Collagen alpha-1(I) chain	*Bos taurus*	AAI49096.1	1654	130	33
Wa23	*Cervus elaphus*	Collagen alpha-1(I) chain	*Cervus canadensis*	XP_043327093.1	1348	81	35
Wa24	*Cervus elaphus*	Collagen alpha-1(I) chain	*Ursus maritimus*	XP_040496745.1	4163	274	38

1

The *N*- and *C*-terminal parts are not identified because they are removed during collagen maturation.[55] The highest number of the identified peptides correspond to the type I collagen sequence. Proteomic identification correlates with taxonomic identification by archaeologists and paleontologists using comparative anatomy. Analyses are 100% consistent with the family level and 95% consistent with the genus level of taxonomic identifications (Table 1). The sample Wa24 was not identified, as the size and preservation of the bone did not permit this. The LC-MS/MS analysis showed that the Wa24 sample corresponded to an Ursidae, as did the MALDI FTICR. Unique peptides identified by LC-MS/MS on collagen 1 alpha 1, 1 alpha 2, and 3 alpha 1 (major bone proteins), with Mascot software were selected and further analyzed with BlastP to validate family, genus and species. Table 2 lists the specific peptides for the species studied and the Supplementary Information 3 gathers the MS/MS of the unique peptides (Supplementary Information 3 Figure S5 to S16). This analysis was facilitated by the preliminary MALDI analysis that allowed targeting of the genus. The LC-MS/MS results show that dermatological skin tape-discs stripping allows identifying taxonomy of the bones using the collagen, but identification of the taxonomy is influenced by the collagen (I) reference database, and some taxa may be poorly resolved due to a lack of available sequences. Currently, the NCBI database contains 271,985,047 million sequences for 138,491 organisms but for example the COL1A1 and COL1A2 collagens of species of archaeological interest *Castor fiber* and *Capreolus capreolus* are not present. A second problem is the similarity of type I collagen sequences which does not allow discrimination between species as for *Ursus arctos horibilis, Ursus maritimus, Ursus americanus*. This implies that discrimination between species should be done using other proteins present in the bones or teeth.[56]

Table 2. List of specific peptides of genus taxa for archaeological and paleontological samples identified by LC-MS/MS from bones sampled with dermatological skin tape-discs stripping and validated by BlastP with 100% identity and 100% cover query. The table includes the taxonomic name, rank, the collagen accession number, the peptide sequence and the position of the peptide on the sequence.

Taxonomic name	Taxonomic rank	Accession number	Specific peptide sequence for collagen type I	Position in COL sequence
Castoridae	Familly	XP_020019043.1	GEVGLPGLSGPVGPPGNPGANGLAGSK, COL1A2 GLVGEPGPAGSKGETGSK, COL1A2	283 - 309 343 - 360
Suinae	Familly	BAX02569.1	GNDGSVGPVGPAGPIGSAGPPGFPGAPGPK, COL1A2 GEVGLPGVSGPVGPPGNPGANGLPGAK, COL1A2	235 – 264 283 - 309
Cervidae	Familly	XP_043750223.1	GAPGAQGPPGAPGPLGIAGVTGAR, COL1A3	933 - 956
Equus	Genus	XP_023508478.1	GDAGPPGPAGPAGPPGPIGSVGAPGPK, COL1A1	835 – 861
Caprinae (Ovis)	Sub familly	XP_004007775.1	TGEPGAAGPPGFVGEK, COL1A2 GEPGPVGAVGPAGAVGPR, COL1A2	831 – 846 977 - 994
Bos	Genus	NP_776945.1	SGETGASGPPGFVGEK, COL1A2	829-844
Struthio camelus	Species	XP_009672566.1	GLHGEFGAPGPAGPR, COL1A2	574 - 588
Capra	Genus	XP_005678993.1	TGEPGAAGPPGFVGEKGPSGEPGTAGPPGTPGPQGFLGP PGFLGLPGSR, COL1A2	829 - 877
Ursus	Genus	XP_008684476.1	GESGNKGEPGSVGPQGPPGPSGEEGK, COL1A2	356 - 381

The database mining of the LC-MS/MS proteomics data led also to the identification of non-collagenous proteins (NCP) such as biglycan, chondroadherin, osteomodulin, calmodulin.[57] These protein has been identified in bones from different periods and for different species such as humans, mammoths, moas, cattle, horses, turkeys, rabbits, squirrels, extinct rhinoceros.[6, 39, 56-64] Because many of the NCPs and some collagens have higher mutation rates than collagen type 1, they are better targets for phylogenetic reconstructions, especially between closely related species.[10, 39] The number of NCP proteins identified is higher in the Holocene samples than in the Upper Pleistocene. Although the stripping with dermatological skin tape-discs is a low-invasive method, it does identify minor proteins such as NCPs. The decrease in the number of these proteins is correlated with the smaller number of peptides identified in the Upper Pleistocene samples due to the faster degradation of these proteins compared to collagen.[62-63] Both biglycan and lumican which exhibits collagen-binding properties were identified in modern, Iron Age and Neolithic samples but not in Pleistocene sample. This indicates that these proteins have been degraded in Pleistocene samples. Studies show that it is possible to detect NCP proteins in samples as old as 650 Ka, but using only 50 mg of sample.[62-63] The increase in yield obtained with our method would make it possible to use only 5 mg of sample to obtain the same results.

We detected also in the LC-MS/MS proteomics data a variety of *in vivo* and diagenetic post-translational modifications (PTMs) in the samples. Analysis of post-translational modifications such as glutamine deamidation provides information on the conservation status of bone. The degree of glutamine deamidation, has been previously reported to hold promising as an indicator of degradation in ancient materials.[19, 65-67] Older or degradated samples have a higher level of deamidation.[19, 65-71] Deamidation is used in many publications to validate that proteins found in samples are ancient and not the result of a contamination. Unfortunately, comparing deamidation

data between publications is complex due to technical issues and condition of preservation of the bones. The first technical issue is that the rate of deamidation obtained for bones from the same period depends on the precise demineralization and trypsin digestion protocol as deamidation is easily induced during protein digestion.[19, 65, 69] The second technical issue is that the calculation methods and software are different and do not focus on the same deamidation rate values.[48, 66-67, 69, 71-72]

Several values of deamidation are obtained according to the analytical protocol. In ZooMS based on MALDI MS analysis only the percentage of glutamine deamidation from the peptide P1 is used. LC-MS/MS gives access to the percentage of glutamine and asparagine deamidation of all the peptides identified on type I collagen and other proteins. In 2012, Van Doorm *et al.* showed that the deamidation rate of glutamine according to their position on the collagen 1 alpha 1 are not identical.[19] Figure 3 compares the percentages of deamidation between samples using MALDI FTICR or LC-MS/MS analyses, for deamidation of glutamine peptides $P1^{m}$ or $P1^{b}$ for mammals and birds, global deamidation of asparagine and glutamine and their specific values for COL1A1 and COL1A2 and proline hydroxylation. The percentage of deamidation on the glutamine based on peptide $P1^{m}$ for mammals and $P1^{b}$ for birds increases significantly (Student's *t test*, $p<0.001$) between modern vs Neolithic or Neolithic vs Pleistocene both for the MALDI FTICR and LC-MS/MS analyses but no significant difference between Iron age and Neolithic samples is observed. Although the two analytical methods MALDI FTICR and nanoLC-MS/MS Orbitrap have completely different ionization processes, the quantification data obtained on deamidation are similar. The percentage of deamidation on the glutamine based on the peptide P1 by both methods are strongly correlated ($R^2 = 0.91$, Supplementary Information 1, Figure S9). This deamidation on glutamine percentage for the peptide P1 determined by MALDI FTICR and LC-MS/MS correlates

also well with the global deamidation percentage for all glutamines calculated from Maxquant quantification data by the deamidation software package for COL1A1 and COL1A2 (R^2 = 0.88 and R^2 = 0.94 respectively, Supplementary Information S1, Figure S10 and S11 respectively). The better correlation between peptide P1 and COL1A1 is not surprising as peptide P1 belongs to COL1A1. Furthermore Figure 3 shows that the percentage of deamidation asparagine and glutamine follows the same trend for both COL1A1 and COL1A2. A significant increase is observed between modern and Pleistocene with Maxquant method (Student's *t test*, p<0.001, Figure 3). This indicates that the result of the calculation of the deamidation on glutamine using the peptide P1 or all potentially deamidated positions on COL1A1 or COL1A2 are similar, even if some positions have a faster or slower deamidation rate, this has little impact on the result compared with the peptide P1 analysis alone.

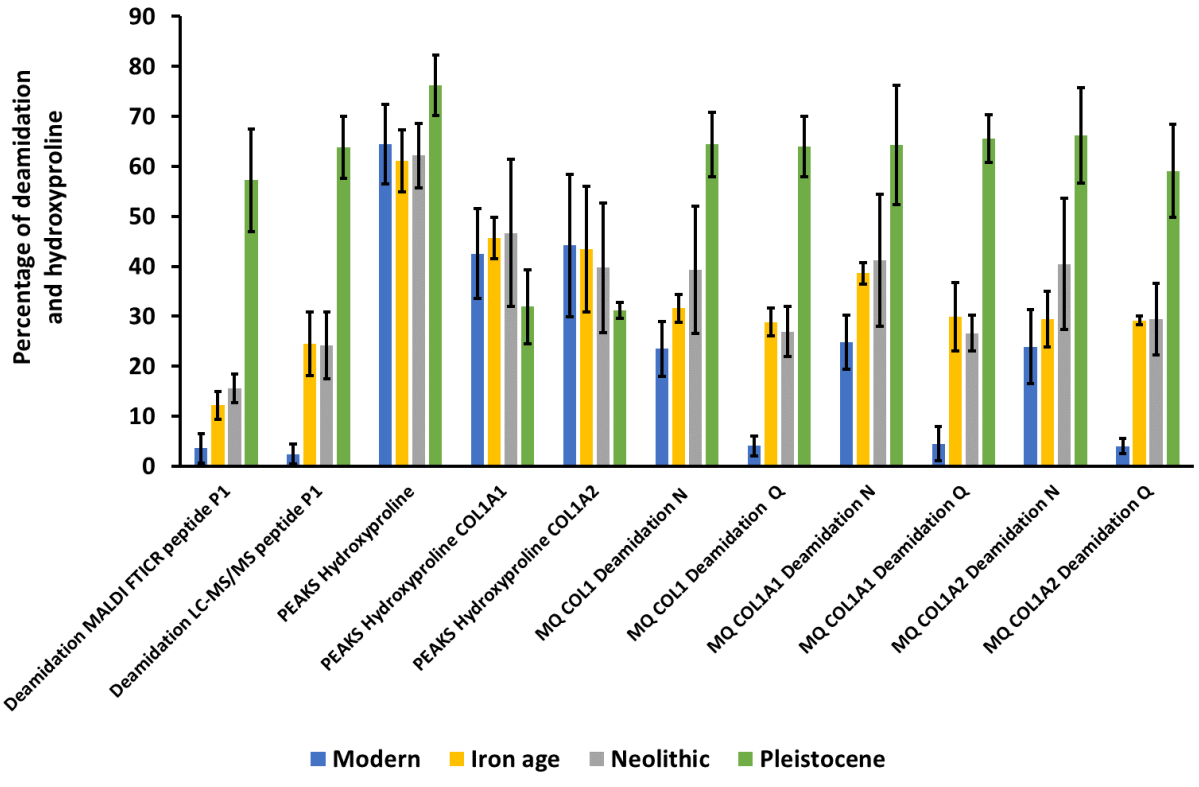

Figure 3. Percentage of deamidation and hydroxyproline for archeological samples by MALDI FTICR and LC-MS/MS. Bleu: modern sample, orange: Iron Age, grey: Neolithic and green Pleistocene. The deamidation percentage rates were calculated from MALDI FTICR and LC-MS/MS data based on the peptide $P1^m$ and $P1^b$, deamidation percentage rates on N and Q were calculated from the LC-MS/MS by processing Maxquant (MQ) with the deamidation package, the percentage of hydroxylation on peptides values are based on LC-MS/MS with PEAKS X+ PTM (method 1) and the percentage of hydroxyproline residues with PEAKS X+ (method 2).

The percentage of the deamidation on the asparagine increases between modern and plesitocene and iron age, Neolithic and pleistocene (Student's *t test*, $p<0.001$, Figure 3). There is no significant difference between Iron Age and Neolithic sample for the deamidation on the asparagine, however an increasing trend is observed. The percentage of deamidation on asparagine is slightly higher than glutamine but without significant statistical value in agreement with Pal Chowdhury *at al*, who showed that the glutamine is much more stable towards deamidation than asparagine.[67] The correlation of the global deamidation percentage of all asparagines with the deamidation on glutamine based on the peptide P1 obtained for MALDI and LC-MS/MS methods is slightly lower (R^2 = 0.60 and R^2 = 0.70, Supplementary Information 1, Figure S12, S13 respectively) when compared to the correlation of glutamine deamidation (Supplementary Information 1, Figure S10; Figure S11). Finally, the correlation between the global percentages of glutamine and asparagine deamidation determined by LC-MS/MS and calculated from Maxquant quantitative data by the deamidation software package is a bit better (R^2 = 0.71, Supplementary Information 1, Figure S14).

It is worth to notice that is no significant difference between the glutamine deamidation percentages between the bones from Temblay-en-France and Bouchain (both open air sites, Iron Age 2^{nd}-1^{st} century BC, Neolithic 3200-2900 BC) and the bones from Pertus II (cave site, Neolithic 3800 - 3300 BC). The deamidation values for Bouchain, Pertus and Tremblay sites are 29.8% +/- 6.1%, 52.5% +/- 4.5%, 31.5% +/- 2.7% for asparagine, 24.2% +/- 2.8%, 30.6% +/- 5.3%, 28.8%

+/- 2.7% for glutamine deamidation values respectively (Supplementary Information 4, Table S1). Comparing the deamidation rates between the three sites, it seems counter-intuitive that the deamidation rate for the Iron-Age samples is not the lowest. However, it has been described that storage conditions such as humidity and oxygen levels have an influence on the rate of deamidation.[66]

Coutu *et al.* studied the domestication of sheep in southern Africa around 2000 BP through paleoproteomics using LC-MS/MS.[52] For archeological sample, overall percentage of deamidation for asparagine (N) and glutamine (Q) residues for tryptic peptides was calculated. The results show 69% of deamidation for asparagine and 37% for glutamine deamidation for archeological samples. The Bouchain and Pertus samples, which are closest in terms of dates, show different values for glutamine and asparagine deamidation. Using the deamidation measurement by processing Maxquant quantification data with the deamidation package, they obtained 39.3% of deamidation for asparagine and 26.9% for glutamine deamidation for Neolithic samples. However, for modern samples Coutu *et al.* obtained deamidation percentage rates of 50% and 10% for asparagine and glutamine respectively. Deamidation values for the modern *Bos taurus* sample using our method are much lower. (23.5% +/- 5.5% for asparagine and 4.1% +/- 2.0% for glutamine, Supplementary Information 4, Table S1). These differences between the data indicate that our method affords a gentler extraction of proteins from bone and a milder digestion without acidic demineralization which preserves fragile post-translational modifications. It should also be pointed out that many factors such as temperature, humidity and pH can influence results.[66, 70, 72]

The de novo analysis of the PTMs present in the samples using the PEAKS X+ software shows the presence of hydroxyproline. Hydroxyproline is naturally present in collagen to stabilize the structure.[73-74] It has been shown that the occupancy of proline hydroxylation on the sequences of

COL1 varies according to amino acid positions.[75] Approximately 50% of proline residues in collagen are hydroxylated.[76] The average percentage of hydroxyproline present on tryptic collagen peptides identified by LC-MS/MS is 65% using calculation method (1) while the percentage of hydroxyproline residues is 42% and 39% on COL1A1 and COL1A2 respectively (method 2). There is an increase in hydroxyproline frequency and a decrease of hydroxyproline residues in Pleistocene bones but this change appears to be not correlated with sample age (Figure 3). The reduced sequence coverage for Pleistocene bones may explain the decrease of hydroxyproline residues while hydroxyproline frequency percentage (method 1) is unaffected by sequence coverage.

CONCLUSION

Our work shows that dermatological skin tape-disc stripping, leaves very little trace on bones. Analysis on the stripped samples either by MALDI FTICR MS peptide fingerprinting or LC-MS/MS bottom-up proteomics allows identifying the proteins contained in bones giving access to the taxonomy identification of bones from today up to Pleistocene. The comparison of the different sampling methods shows that using dermatological skin tape-disc gives similar results to destructive sampling followed by acid demineralization, suggesting that dermatological skin tape-disc is a minimally invasive alternative to destructive sampling. Our method does not use demineralization and gelatinization steps of bones which preserves existing all existing post-translational modification (PTM). LC-MS/MS analyses show that the proteome obtained using our method is in line with literature data, and that this sampling method has no impact on the information obtained, even when only a small quantity of material is used. The deamidation rate that we obtained shows that it is possible to differentiate modern, Neolithic and Plesitocene period from bones. Elucidating the origin and significance of a bone fragment or artefact, while preserving its integrity, is a crucial issue for archaeologists, and the use of dermatological skin tape-disc is an alternative for these studies. Our study shows a potential application for museum analysis of reference and archaeological bones or artefact analysis.

ASSOCIATED CONTENT

SUPPORTING INFORMATION 1 - PDF - contains the list of bone samples, the pictures of the bones before and after sampling, the comparison of sampling methods with peptide quantity and number of ZooMS markers, the correlation of the percentages of deamidation between MALDI FTICR MS and LC-MS/MS

Supplementary Figure S1. Picture of modern bones. A- *Capra ibex* rib; B- *Bos taurus* mandible; C- *Struthio camelus* femur; D- *Cervus elaphus* mandible. One square of the scale bar represents 1 cm. 10

Supplementary Figure S2. Picture of archaeological bones from Bouchain, Tremblay-en-France. A: *Sus scrofa* tibia (Bouchain); B: *Bos primigenius* femur (Bouchain); C: *Equus sp*. scapula (Tremblay-en-France); D: *Bos taurus* tibia (Tremblay-en-France); E: *Sus scrofa* mandible (Bouchain); F: *Bos taurus* first phalanx (Bouchain); G: *Cervus elaphus* radius (Bouchain); H: *Castor fiber* mandible (Bouchain); I: *Capreolus capreolus* metatarsal (Bouchain). The scale bar represents 1 cm. 11

Supplementary Figure S3. Picture of archaeological bones from Pertus. A- *Bos taurus* radius; B – *Capra hircus* ulna; C – *Sus sp*. mandibule; D – *Cervus elaphus* occipital bone; E – *Ovis sp*. mandible. The black bar represents 1 cm. 12

Supplementary Figure S4. Picture of archaeological bones from Loverval and Waziers. A- *Ursus spelaeus* mandibule; B – *Equus achenheimensis* humerus; C –*Bos primigenius* humerus; D – *Cervus elaphus* metapodial; E – *Cervus elaphus* long bone. The black bar represents 1 cm. 13

Supplementary Figure S5. Optical microscopy (magnification ×4, Optika B-383PLi, Ponteranica (BG), Italy) of *a Bos taurus* bone (Mod 1). A, C, E, G, I: before sampling; B, D, F, H, J, K, L: after sampling. B: Triboelectric rubbing in forced bag sampling; D: Eraser sampling by rubbing, F: Swab mopping sampling; H: Destructive ZooMS sampling; J, K, L: 1 D-Squame® skin tape-disc method. J: 5 D-Squame® skin tape-disc; K: 10 D-Squame® skin tape-disc; L: 20 D-Squame® skin tape-disc. The black circle indicates a modification of the surface after sampling. The black bar represents 1 cm. 15

Supplementary Figure S6. Optical microscopy (magnification ×4, Optika B-383PLi, Ponteranica (BG), Italy) of A: D-Squame® skin tape-disc after sampling of a *Bos taurus* bone (Mod 1) and B: the powder obtained after scraping with a spatula and collected in an Eppendorf™ tube. A thin layer of bone powder is visible. 16

Supplementary Figure S7. Opticala microscopy (magnification ×4, Optika B-383PLi, Ponteranica (BG), Italy) of a wild boar, *Sus scrofa* tibia (Bou 1). A, C, E, G, I: before sampling;

B, D, F, H, J: after sampling. A, B: Triboelectric rubbing in forced bag sampling; C, D: Eraser sampling by rubbing, E, F: Swab mopping sampling, G, H: Destructive ZooMS sampling, I, J: 5 D-Squame® skin tape-disc. The black circle indicates a modification of the surface after sampling. The black bar represents 1 cm. 17

Supplementary Figure S8. Number of ZooMS peptide markers for modern (4 samples, Mod 1 to 4) and Neolithic (3 samples, Bou 1, Bou 2, Trem 2) for each tested method. For each sample the number of ZooMS peptides markers are indicated. The cumulative numbers of peptides markers are found in abscissa and the six sampling methods are ordonate. 18

Supplementary Figure S9. Glutamine deamidation percentages by LC-MS/MS versus MALDI FTICR MS based on the peptide P1 sequence from COL1A1 GVQdemGPoxPGPAGPR for mammals and GVQdemGPoxPGPQGPR for birds. Each of the 23 points corresponds to the analysis of the 4 moderns and 19 archaeologicals samples by 5 D-Squame® skin tape-disc. 19

Supplementary Figure S10. Glutamine deamidation percentages by LC-MS/MS based on Maxquant quantitative data processed with the deamidation package versus the glutamine deamidation percentages by MALDI FTICR MS based on the peptide P1 sequence from COL1A1 GVQdemGPoxPGPAGPR for mammals and GVQdemGPoxPGPQGPR for birds. Each of the 23 points corresponds to the analysis of the 4 moderns and 19 archaeologicals samples by 5 D-Squame® skin tape-disc. 20

Supplementary Figure S11. Glutamine deamidation percentages by LC-MS/MS based on Maxquant quantitative data processed with the deamidation package versus glutamine deamidation percentages by LC-MS/MS based on the peptide P1 sequence from COL1A1 GVQdemGPoxPGPAGPR for mammals and GVQdemGPoxPGPQGPR for birds. Each of the 23 points corresponds to the analysis of the 4 moderns and 19 archaeologicals samples by 5 D-Squame® skin tape-disc. 21

Supplementary Figure S12. Asparagine deamidation percentages by LC-MS/MS based on Maxquant quantitative data processed with the deamidation package versus glutamine deamidation percentages by MALDI FTICR based on the peptide P1 sequence from COL1A1 GVQdemGPoxPGPAGPR for mammals and GVQdemGPoxPGPQGPR for birds. Each of the 23

<u>SUPPORTING INFORMATION 2</u> - Excel file - contains the identification of proteins by LC-MS/MS.

<u>SUPPORTING INFORMATION 3</u> – PDF - contains MS/MS spectra of relevant peptides obtained by LC-MS/MS.

SUPPORTING INFORMATION 4 **-Excel file-** contains the identification using ZooMS markers.

AUTHOR INFORMATION

Corresponding Author

Fabrice Bray - Univ. Lille, CNRS, UAR 3290 - MSAP - Miniaturisation pour la Synthèse, l'Analyse et la Protéomique, F-59650 Lille, France ; ORCID : https://orcid.org/0000-0002-4723-8206; Phone: +33 (0)3 20 33 71 12; Email: fabrice.bray@univ-lille.fr

Authors

Isabelle Fabrizi - Univ. Lille, CNRS, UAR 3290 - MSAP - Miniaturisation pour la Synthèse, l'Analyse et la Protéomique, F-59650 Lille, France ; ORCID : https://orcid.org/0000-0001-8934-4683; Phone: +33 (0)3 20 33 71 12; Email: isabelle.fabrizi@univ-lille.fr

Stéphanie Flament - Univ. Lille, CNRS, UAR 3290 - MSAP - Miniaturisation pour la Synthèse, l'Analyse et la Protéomique, F-59650 Lille, France ; ORCID : https://orcid.org/0000-0002-0135-675X; Phone: +33 (0)3 20 33 71 12; Email: stephanie.flament@univ-lille.fr

Christian Rolando - Univ. Lille, CNRS, UAR 3290 - MSAP - Miniaturisation pour la Synthèse, l'Analyse et la Protéomique, F-59000 Lille, France ; Shrieking Sixties, F-59650 Villeneuve-d'Ascq, France. ORCID : https://orcid.org/0000-0002-3266-8860; Phone: + +33 (0)3 20 43 49 77; Email: christian.rolando@univ-lille.fr

Claire Delhon – Univ Côte d'Azur, CNRS, CEPAM – Culture Environnement Préhistoire Antiquité Moyen Age, F-06300 Nice, France ; ORCID : https://orcid.org/0000-0002-7216-0148; Phone: + +33 (0)4 89 15 24 02 ; Email : claire.delhon@cepam.cnrs.fr

Lionel Gourichon – Univ Côte d'Azur, CNRS, CEPAM – Culture Environnement Préhistoire Antiquité Moyen Age, F-06300 Nice, France ; ORCID : https://orcid.org/0000-0002-5160-5902 ; Email : lionel.gourichon@cepam.cnrs.fr

Manon Vuillien – Univ Côte d'Azur, CNRS, CEPAM – Culture Environnement Préhistoire Antiquité Moyen Age, F-06300 Nice, France ; ORCID : https://orcid.org/0000-0001-7657-7613 ; Email : manon.vuillien@mnhn.fr

Tarek Oueslati - Univ. Lille, CNRS, UMR8164 – HALMA - Histoire Archéologie Littérature des Mondes Anciens, F-59650 Lille, France ; ORCID : https://orcid.org/0000-0002-2886-085X; Phone : +33 (0)3 20 41 66 13 ; Email : tarek.oueslati@univ-lille.fr

Patrick Auguste - Univ. Lille, CNRS, UMR 8198 - EEP – Evolution, Ecology and Paleontology, F-59650, Lille, France ORCID: https://orcid.org/0000-0002-8302-6307; Email: patrick.auguste@univ-lille.fr

Author Contributions

I.F, SF. Methodology , writing-original draft. F.B. Conceptualization, methodology, writing-original draft, project administration, funding acquisition. C.R: writing-review-editing, funding acquisition. C.D, L. G, M. V: review and access of the bones. T.O, P.A: review, access of the bones and funding acquisition. All authors have given approval to the final version of the manuscript

Funding Sources

The authors acknowledge the IBiSA network for financial support of the UAR 3290 (MSAP) proteomics facility. The mass spectrometers were funded by University of Lille, CNRS, Région Hauts-de-France and the European Regional Development Fund. The authors deeply thank CNRS - Mission pour l'Interdisciplinarité for the funding of the Prot_HR_DAT project (2021-2022) and the AgroPastCN project (2020-2021). The authors are grateful to the Région Hauts-de-France for the funding of the ProtéOsHdF project (2021-2022) and the financial support from the IR INFRANALYTICS FR2054 CNRS for conducting the research is gratefully. The EU_FTICR _MS project Grant Agreement 731077, and the IPERION HS project Grant Agreement 871034 funded by EU Horizon 2020 Research and Innovation Program are profoundly thanked.

Notes

Any additional relevant notes should be placed here.

ACKNOWLEDGMENT

The authors thank warmly Cédric Lepère, (ÉVEHA -Études et Valorisations Archéologiques, Limoges, France and Université Côte d'Azur, CNRS, UMR 7264 – CEPAM - Cultures et Environnements Préhistoire, Antiquité, Moyen Âge, Nice, France) for allowing the authors to sample material from the cave Pertus II (Méailles, Alpes-de-Haute-Provence, France).

REFERENCES

1. Feng, X., Chemical and biochemical basis of cell-bone matrix interaction in health and disease. *Current chemical biology* **2009,** *3* (2), 189-196.
2. Hendy, J.; Welker, F.; Demarchi, B.; Speller, C.; Warinner, C.; Collins, M., A guide to ancient protein studies. *Nature Ecology & Evolution* **2018,** *2* (5), 791-799.
3. Buckley, M.; Collins, M.; Thomas-Oates, J.; Wilson, J. C., Species identification by analysis of bone collagen using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry* **2009,** *23* (23), 3843-54.
4. Schroeter, E. R.; DeHart, C. J.; Schweitzer, M. H.; Thomas, P. M.; Kelleher, N. L., Bone protein "extractomics": comparing the efficiency of bone protein extractions of Gallus gallus in tandem mass spectrometry, with an eye towards paleoproteomics. *PeerJ* **2016,** *4*, e2603.
5. Schroeter, E. R.; Blackburn, K.; Goshe, M. B.; Schweitzer, M. H., Proteomic method to extract, concentrate, digest and enrich peptides from fossils with coloured (humic) substances for mass spectrometry analyses. *Royal Society open science* **2019,** *6* (8), 181433.
6. Kostyukevich, Y.; Bugrova, A.; Chagovets, V.; Brzhozovskiy, A.; Indeykina, M.; Vanyushkina, A.; Zherebker, A.; Mitina, A.; Kononikhin, A.; Popov, I.; Khaitovich, P.; Nikolaev, E., Proteomic and lipidomic analysis of mammoth bone by high-resolution tandem mass spectrometry coupled with liquid chromatography. *European Journal of Mass Spectrometry* **2018,** *24* (6), 411-419.
7. Cleland, T. P., Human Bone Paleoproteomics Utilizing the Single-Pot, Solid-Phase-Enhanced Sample Preparation Method to Maximize Detected Proteins and Reduce Humics. *Journal of proteome research* **2018,** *17* (11), 3976-3983.
8. Palmer, K. S.; Makarewicz, C. A.; Tishkin, A. A.; Tur, S. S.; Chunag, A.; Diimajav, E.; Jamsranjav, B.; Buckley, M., Comparing the Use of Magnetic Beads with Ultrafiltration for Ancient Dental Calculus Proteomics. *Journal of proteome research* **2021,** *20* (3), 1689-1704.
9. Bray, F.; Flament, S.; Abrams, G.; Bonjean, D.; Rolando, C.; Tokarski, C.; Auguste, P., Extinct species identification from late Middle Pleistocene and earlier Upper Pleistocene bone fragments and tools not recognizable from their osteomorphological study by an enhanced proteomics protocol. *Archaeometry* **2022**.
10. Rüther, P. L.; Husic, I. M.; Bangsgaard, P.; Gregersen, K. M.; Pantmann, P.; Carvalho, M.; Godinho, R. M.; Friedl, L.; Cascalheira, J.; Taurozzi, A. J.; Jørkov, M. L. S.; Benedetti, M. M.; Haws, J.; Bicho, N.; Welker, F.; Cappellini, E.; Olsen, J. V., SPIN enables high throughput species identification of archaeological bone by proteomics. *Nature Communications* **2022,** *13* (1), 2458.
11. Bray, F.; Fabrizi, I.; Flament, S.; Locht, J. L.; Antoine, P.; Auguste, P.; Rolando, C., Robust High-Throughput Proteomics Identification and Deamidation Quantitation of Extinct Species up to Pleistocene with Ultrahigh-Resolution MALDI-FTICR Mass Spectrometry. *Analytical Chemistry* **2023,** *95* (19), 7422-7432.
12. Taurozzi, A. J.; Rüther, P. L.; Patramanis, I.; Koenig, C.; Sinclair Paterson, R.; Madupe, P. P.; Harking, F. S.; Welker, F.; Mackie, M.; Ramos-Madrigal, J., Deep-time phylogenetic inference by paleoproteomic analysis of dental enamel. *Nature Protocols* **2024**, 1-32.
13. Warinner, C.; Korzow Richter, K.; Collins, M. J., Paleoproteomics. *Chemical reviews* **2022,** *122* (16), 13401-13446.
14. Richter, K. K.; Codlin, M. C.; Seabrook, M.; Warinner, C., A primer for ZooMS applications in archaeology. *Proc Natl Acad Sci U S A* **2022,** *119* (20), e2109323119.
15. Schroeter, E. R.; Cleland, T. P.; Schweitzer, M. H., Deep Time Paleoproteomics: Looking Forward. *Journal of proteome research* **2021**.

16. Cucina, A.; Cunsolo, V.; Di Francesco, A.; Saletti, R.; Zilberstein, G.; Zilberstein, S.; Tikhonov, A.; Bublichenko, A. G.; Righetti, P. G.; Foti, S., Meta-proteomic analysis of the Shandrin mammoth by EVA technology and high-resolution mass spectrometry: what is its gut microbiota telling us? *Amino Acids* **2021,** *53* (10), 1507-1521.
17. Cleland, T. P.; Schroeter, E. R.; Colleary, C., Diagenetiforms: a new term to explain protein changes as a result of diagenesis in paleoproteomics. *Journal of proteomics* **2021,** *230*, 103992.
18. Welker, F.; Smith, G. M.; Hutson, J. M.; Kindler, L.; Garcia-Moreno, A.; Villaluenga, A.; Turner, E.; Gaudzinski-Windheuser, S., Middle Pleistocene protein sequences from the rhinoceros genus Stephanorhinus and the phylogeny of extant and extinct Middle/Late Pleistocene Rhinocerotidae. *PeerJ* **2017,** *5*, e3033.
19. van Doorn, N. L.; Wilson, J.; Hollund, H.; Soressi, M.; Collins, M. J., Site-specific deamidation of glutamine: a new marker of bone collagen deterioration. *Rapid Communications in Mass Spectrometry* **2012,** *26* (19), 2319-27.
20. Pálsdóttir, A. H.; Bläuer, A.; Rannamäe, E.; Boessenkool, S.; Hallsson, J. H., Not a limitless resource: ethics and guidelines for destructive sampling of archaeofaunal remains. *Royal Society open science* **2019,** *6* (10), 191059.
21. van Doorn, N. L.; Hollund, H.; Collins, M. J., A novel and non-destructive approach for ZooMS analysis: ammonium bicarbonate buffer extraction. *Archaeological and Anthropological Sciences* **2011,** *3* (3), 281-289.
22. Fiddyment, S.; Holsinger, B.; Ruzzier, C.; Devine, A.; Binois, A.; Albarella, U.; Fischer, R.; Nichols, E.; Curtis, A.; Cheese, E.; Teasdale, M. D.; Checkley-Scott, C.; Milner, S. J.; Rudy, K. M.; Johnson, E. J.; Vnoucek, J.; Garrison, M.; McGrory, S.; Bradley, D. G.; Collins, M. J., Animal origin of 13th-century uterine vellum revealed using noninvasive peptide fingerprinting. *Proceedings of the National Academy of Sciences (Proceedings of the National Academy of Sciences of the United States of America)* **2015,** *112* (49), 15066-71.
23. McGrath, K.; Rowsell, K.; Gates St-Pierre, C.; Tedder, A.; Foody, G.; Roberts, C.; Speller, C.; Collins, M., Identifying Archaeological Bone via Non-Destructive ZooMS and the Materiality of Symbolic Expression: Examples from Iroquoian Bone Points. *Scientific Reports* **2019,** *9* (1), 11027.
24. Righetti, P. G.; Zilberstein, G.; Zilberstein, S., EVA Technology and Proteomics: A Two-Pronged Attack on Cultural Heritage. *Journal of proteome research* **2020,** *19* (8), 2914-2925.
25. Sinet-Mathiot, V.; Martisius, N. L.; Schulz-Kornas, E.; van Casteren, A.; Tsanova, T. R.; Sirakov, N.; Spasov, R.; Welker, F.; Smith, G. M.; Hublin, J. J., The effect of eraser sampling for proteomic analysis on Palaeolithic bone surface microtopography. *Scientific Reports* **2021,** *11* (1), 23611.
26. Barberis, E.; Manfredi, M.; Ferraris, E.; Bianucci, R.; Marengo, E., Non-Invasive Paleo-Metabolomics and Paleo-Proteomics Analyses Reveal the Complex Funerary Treatment of the Early 18th Dynasty Dignitary NEBIRI (QV30). *Molecules* **2022,** *27* (21), 7208.
27. Martisius, N. L.; Welker, F.; Dogandzic, T.; Grote, M. N.; Rendu, W.; Sinet-Mathiot, V.; Wilcke, A.; McPherron, S. J. P.; Soressi, M.; Steele, T. E., Non-destructive ZooMS identification reveals strategic bone tool raw material selection by Neandertals. *Scientific Reports* **2020,** *10* (1), 7746.
28. Kirby, D. P.; Manick, A.; Newman, R., Minimally invasive sampling of surface coatings for protein identification by peptide mass fingerprinting: a case study with photographs. *Journal of the American Institute for Conservation* **2020,** *59* (3-4), 235-245.

29. Chen, P., Comparison of Sandpaper and Polishing Film in Minimally-Invasive ZooMS. *The Ethnograph: Journal of Anthropological Studies* **2023,** *5* (1), 48-57.
30. Gilbert, C.; Krupicka, V.; Galluzzi, F.; Popowich, A.; Bathany, K.; Claverol, S.; Arslanoglu, J.; Tokarski, C., Species identification of ivory and bone museum objects using minimally invasive proteomics. *Science Advances* **2024,** *10* (4), eadi9028.
31. Evans, Z.; Paskulin, L.; Rahemtulla, F.; Speller, C. F., A comparison of minimally-invasive sampling techniques for ZooMS analysis of bone artifacts. *Journal of Archaeological Science: Reports* **2023,** *47*.
32. Multari, D. H.; Ravishankar, P.; Sullivan, G. J.; Power, R. K.; Lord, C.; Fraser, J. A.; Haynes, P. A., Development of a novel minimally invasive sampling and analysis technique using skin sampling tape strips for bioarchaeological proteomics. *Journal of Archaeological Science* **2022,** *139*, 105548.
33. Hansen, J.; Dekker, J.; Troché, G.; Fagernäs, Z.; Olsen, J. V.; Seguí, M. S.; Welker, F., A comparative study of commercially available, minimally invasive, sampling methods on Early Neolithic humeri analysed via palaeoproteomics. *Journal of Archaeological Science* **2024,** *167*, 106002.
34. Battentier, J.; Lepère, C.; Théry-Parisot, I.; Carré, A.; Delhon, C., La grotte de Pertus II (Méailles, Alpes-de-Haute-Provence): exploitation du couvert forestier au chasséen récent (3850-3650 cal. BC). 2016.
35. Broutin, P., Tremblay-en-France–Zac Sud Charles-de-Gaulle, rû du Sausset (tranche 1). Opération préventive de diagnostic (2018). *ADLFI. Archéologie de la France-Informations. une revue Gallia* **2021**.
36. Salvador, P. G.; Boulen, M.; Leroy, G.; Oueslati, T., Le site de Bouchain (France): apports d'une étude pluridisciplinaire à l'évolution paléoenvironnementale de la vallée de l'Escaut durant l'Holocène moyen. *Quaternaire. Revue de l'Association française pour l'étude du Quaternaire* **2021,** *32* (4), 253-280.
37. Oueslati, T.; Leroy, G.; Salvador, P.-G., Fowling on the banks of the Scheldt river in the recent Neolithic (France, 3300-2900 cal BC). *Quaternary International* **2022,** *626*, 52-61.
38. Hérisson, D.; Deschodt, L.; Antoine, P.; Locht, J. l.; Lacroix, S.; Angélique, S.; Yann, P.; Vallin, L.; Rorive, S.; Simon, F., Waziers, Le Bas-Terroir: historique de dix années de recherches archéologiques et géomorphologiques dans un marais pléistocène de la plaine de la Scarpe (2011-2021). *Quaternaire. Revue de l'Association française pour l'étude du Quaternaire* **2022,** *33* (4), 225-246.
39. Welker, F.; Hajdinjak, M.; Talamo, S.; Jaouen, K.; Dannemann, M.; David, F.; Julien, M.; Meyer, M.; Kelso, J.; Barnes, I., Palaeoproteomic evidence identifies archaic hominins associated with the Châtelperronian at the Grotte du Renne. *Proceedings of the National Academy of Sciences (Proceedings of the National Academy of Sciences of the United States of America)* **2016,** *113* (40), 11162-11167.
40. Buckley, M., Zooarchaeology by Mass Spectrometry (ZooMS) Collagen Fingerprinting for the Species Identification of Archaeological Bone Fragments. In *Zooarchaeology in Practice*, 2018; pp 227-247.
41. Buckley, M.; Herman, J., Species identification of Late Pleistocene bat bones using collagen fingerprinting. *International Journal of Osteoarchaeology* **2019,** *29* (6), 1051-1059.
42. Eda, M.; Morimoto, M.; Mizuta, T.; Inoué, T., ZooMS for birds: Discrimination of Japanese archaeological chickens and indigenous pheasants using collagen peptide fingerprinting. *Journal of Archaeological Science: Reports* **2020,** *34*.

43. Janzen, A.; Richter, K. K.; Mwebi, O.; Brown, S.; Onduso, V.; Gatwiri, F.; Ndiema, E.; Katongo, M.; Goldstein, S. T.; Douka, K., Distinguishing African bovids using Zooarchaeology by Mass Spectrometry (ZooMS): New peptide markers and insights into Iron Age economies in Zambia. *PloS one* **2021,** *16* (5), e0251061.
44. Codlin, M. C.; Douka, K.; Richter, K. K., An application of zooms to identify archaeological avian fauna from Teotihuacan, Mexico. *Journal of Archaeological Science* **2022,** *148*, 105692.
45. Henikoff, S.; Henikoff, J. G., Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences* **1992,** *89* (22), 10915-10919.
46. Vizcaíno, J. A.; Côté, R. G.; Csordas, A.; Dianes, J. A.; Fabregat, A.; Foster, J. M.; Griss, J.; Alpi, E.; Birim, M.; Contell, J., The PRoteomics IDEntifications (PRIDE) database and associated tools: status in 2013. *Nucleic acids research* **2012,** *41* (D1), D1063-D1069.
47. Cox, J.; Mann, M., MaxQuant enables high peptide identification rates, individualized ppb-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology* **2008,** *26* (12), 1367-1372.
48. Mackie, M.; Rüther, P.; Samodova, D.; Di Gianvincenzo, F.; Granzotto, C.; Lyon, D.; Peggie, D. A.; Howard, H.; Harrison, L.; Jensen, L. J., Palaeoproteomic profiling of conservation layers on a 14th century Italian wall painting. *Angewandte Chemie International Edition* **2018,** *57* (25), 7369-7374.
49. Brown, S.; Douka, K.; Collins, M. J.; Richter, K. K., On the standardization of ZooMS nomenclature. *Journal of proteomics* **2021,** *235*, 104041.
50. Desmond, A.; Barton, N.; Bouzouggar, A.; Douka, K.; Fernandez, P.; Humphrey, L.; Morales, J.; Turner, E.; Buckley, M., ZooMS identification of bone tools from the North African Later Stone Age. *Journal of Archaeological Science* **2018,** *98*, 149-157.
51. Olesen, C. M.; Fuchs, C. S. K.; Philipsen, P. A.; Haedersdal, M.; Agner, T.; Clausen, M. L., Advancement through epidermis using tape stripping technique and Reflectance Confocal Microscopy. *Scientific Reports* **2019,** *9* (1), 12217.
52. Coutu, A. N.; Taurozzi, A. J.; Mackie, M.; Jensen, T. Z. T.; Collins, M. J.; Sealy, J., Palaeoproteomics confirm earliest domesticated sheep in southern Africa ca. 2000 BP. *Scientific Reports* **2021,** *11* (1), 6631.
53. Liu, S.; Zheng, W.; Yang, B.; Tao, X., Triboelectric charge density of porous and deformable fabrics made from polymer fibers. *Nano energy* **2018,** *53*, 383-390.
54. Naihui, W.; Samantha, B.; Peter, D.; Sandra, H.; Maxim, K.; Sindy, L.; Oshan, W.; Stefano, G.; Michael, C.; Liora, H. K., Testing the efficacy and comparability of ZooMS protocols on archaeological bone. *Journal of proteomics* **2021,** *233*, 104078.
55. Onursal, C.; Dick, E.; Angelidis, I.; Schiller, H. B.; Staab-Weijnitz, C. A., Collagen biosynthesis, processing, and maturation in lung ageing. *Frontiers in Medicine* **2021,** *8*, 593874.
56. Cappellini, E.; Welker, F.; Pandolfi, L.; Ramos-Madrigal, J.; Samodova, D.; Ruther, P. L.; Fotakis, A. K.; Lyon, D.; Moreno-Mayar, J. V.; Bukhsianidze, M.; Rakownikow Jersie-Christensen, R.; Mackie, M.; Ginolhac, A.; Ferring, R.; Tappen, M.; Palkopoulou, E.; Dickinson, M. R.; Stafford, T. W., Jr.; Chan, Y. L.; Gotherstrom, A.; Nathan, S.; Heintzman, P. D.; Kapp, J. D.; Kirillova, I.; Moodley, Y.; Agusti, J.; Kahlke, R. D.; Kiladze, G.; Martinez-Navarro, B.; Liu, S.; Sandoval Velasco, M.; Sinding, M. S.; Kelstrup, C. D.; Allentoft, M. E.; Orlando, L.; Penkman, K.; Shapiro, B.; Rook, L.; Dalen, L.; Gilbert, M. T. P.; Olsen, J. V.; Lordkipanidze, D.; Willerslev, E., Early Pleistocene enamel proteome from Dmanisi resolves Stephanorhinus phylogeny. *Nature* **2019,** *574* (7776), 103-107.

57. Lanigan, L. T.; Mackie, M.; Feine, S.; Hublin, J.-J.; Schmitz, R. W.; Wilcke, A.; Collins, M. J.; Cappellini, E.; Olsen, J. V.; Taurozzi, A. J., Multi-protease analysis of Pleistocene bone proteomes. *Journal of proteomics* **2020,** *228*, 103889.
58. Chen, F.; Welker, F.; Shen, C. C.; Bailey, S. E.; Bergmann, I.; Davis, S.; Xia, H.; Wang, H.; Fischer, R.; Freidline, S. E.; Yu, T. L.; Skinner, M. M.; Stelzer, S.; Dong, G.; Fu, Q.; Dong, G.; Wang, J.; Zhang, D.; Hublin, J. J., A late Middle Pleistocene Denisovan mandible from the Tibetan Plateau. *Nature* **2019,** *569* (7756), 409-412.
59. Sawafuji, R.; Cappellini, E.; Nagaoka, T.; Fotakis, A. K.; Jersie-Christensen, R. R.; Olsen, J. V.; Hirata, K.; Ueda, S., Proteomic profiling of archaeological human bone. *Royal Society open science* **2017,** *4* (6), 161004.
60. Cappellini, E.; Jensen, L. J.; Szklarczyk, D.; Ginolhac, A.; da Fonseca, R. A.; Stafford, T. W.; Holen, S. R.; Collins, M. J.; Orlando, L.; Willerslev, E.; Gilbert, M. T.; Olsen, J. V., Proteomic analysis of a pleistocene mammoth femur reveals more than one hundred ancient bone proteins. *Journal of proteome research* **2012,** *11* (2), 917-26.
61. Cleland, T. P.; Schroeter, E. R.; Schweitzer, M. H., Biologically and diagenetically derived peptide modifications in moa collagens. *Proceedings of the National Academy of Sciences (Proceedings of the National Academy of Sciences of the United States of America)* **2015,** *282* (1808), 20150015.
62. Wadsworth, C.; Buckley, M., Proteome degradation in fossils: investigating the longevity of protein survival in ancient bone. *Rapid Communications in Mass Spectrometry* **2014,** *28* (6), 605-15.
63. Buckley, M.; Wadsworth, C., Proteome degradation in ancient bone: diagenesis and phylogenetic potential. *Palaeogeography, Palaeoclimatology, Palaeoecology* **2014,** *416*, 69-79.
64. Orlando, L.; Ginolhac, A.; Zhang, G.; Froese, D.; Albrechtsen, A.; Stiller, M.; Schubert, M.; Cappellini, E.; Petersen, B.; Moltke, I.; Johnson, P. L.; Fumagalli, M.; Vilstrup, J. T.; Raghavan, M.; Korneliussen, T.; Malaspinas, A. S.; Vogt, J.; Szklarczyk, D.; Kelstrup, C. D.; Vinther, J.; Dolocan, A.; Stenderup, J.; Velazquez, A. M.; Cahill, J.; Rasmussen, M.; Wang, X.; Min, J.; Zazula, G. D.; Seguin-Orlando, A.; Mortensen, C.; Magnussen, K.; Thompson, J. F.; Weinstock, J.; Gregersen, K.; Roed, K. H.; Eisenmann, V.; Rubin, C. J.; Miller, D. C.; Antczak, D. F.; Bertelsen, M. F.; Brunak, S.; Al-Rasheid, K. A.; Ryder, O.; Andersson, L.; Mundy, J.; Krogh, A.; Gilbert, M. T.; Kjaer, K.; Sicheritz-Ponten, T.; Jensen, L. J.; Olsen, J. V.; Hofreiter, M.; Nielsen, R.; Shapiro, B.; Wang, J.; Willerslev, E., Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* **2013,** *499* (7456), 74-8.
65. Wilson, J.; van Doorn, N. L.; Collins, M. J., Assessing the extent of bone degradation using glutamine deamidation in collagen. *Analytical Chemistry* **2012,** *84* (21), 9041-8.
66. Schroeter, E. R.; Cleland, T. P., Glutamine deamidation: an indicator of antiquity, or preservational quality? *Rapid Communications in Mass Spectrometry* **2016,** *30* (2), 251-5.
67. Pal Chowdhury, M.; Wogelius, R.; Manning, P. L.; Metz, L.; Slimak, L.; Buckley, M., Collagen deamidation in archaeological bone as an assessment for relative decay rates. *Archaeometry* **2019,** *61* (6), 1382-1398.
68. Leo, G.; Bonaduce, I.; Andreotti, A.; Marino, G.; Pucci, P.; Colombini, M. P.; Birolo, L., Deamidation at asparagine and glutamine as a major modification upon deterioration/aging of proteinaceous binders in mural paintings. *Analytical Chemistry* **2011,** *83* (6), 2056-64.
69. Simpson, J. P.; Penkman, K. E. H.; Demarchi, B.; Koon, H.; Collins, M. J.; Thomas-Oates, J.; Shapiro, B.; Stark, M.; Wilson, J., The effects of demineralisation and sampling point

variability on the measurement of glutamine deamidation in type I collagen extracted from bone. *Journal of Archaeological Science* **2016,** *69*, 29-38.
70. Welker, F.; Soressi, M. A.; Roussel, M.; van Riemsdijk, I.; Hublin, J.-J.; Collins, M. J., Variations in glutamine deamidation for a Châtelperronian bone assemblage as measured by peptide mass fingerprinting of collagen. *STAR: Science & Technology of Archaeological Research* **2016,** *3* (1), 15-27.
71. Boudier-Lemosquet, A.; Mahler, A.; Bobo, C.; Dufossee, M.; Priault, M., Introducing protein deamidation: Landmark discoveries, societal outreach, and tentative priming workflow to address deamidation. *Methods* **2022,** *200*, 3-14.
72. Ramsøe, A.; van Heekeren, V.; Ponce, P.; Fischer, R.; Barnes, I.; Speller, C.; Collins, M. J., DeamiDATE 1.0: Site-specific deamidation as a tool to assess authenticity of members of ancient proteomes. *Journal of Archaeological Science* **2020,** *115*, 105080.
73. Josse, J.; Harrington, W. F., Role of pyrrolidine residues in the structure and stabilization of collagen. *Journal of Molecular Biology* **1964,** *9* (2), 269-287.
74. Hall, D.; Reed, R., Hydroxyproline and thermal stability of collagen. *Nature* **1957,** *180* (4579), 243-243.
75. Ntasi, G.; Palomo, I. R.; Marino, G.; Piaz, F. D.; Cappellini, E.; Birolo, L.; Petrone, P., Molecular signatures written in bone proteins of 79 AD victims from Herculaneum and Pompeii. *Scientific Reports* **2022,** *12* (1), 1-15.
76. Li, P.; Wu, G., Roles of dietary glycine, proline, and hydroxyproline in collagen synthesis and animal growth. *Amino acids* **2018,** *50*, 29-38.

For TOC Only