

*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*

F BRAY†, S FLAMENT

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

G ABRAMS

Espace muséal d'Andenne, Rue Charles Lapierre 29, 5300, Sclayn, Belgium
Faculty of Archaeology, Leiden University, Van Steenis Building, Einsteinweg 2, 2333 CC
Leiden, The Netherlands

D BONJEAN

Espace muséal d'Andenne, Rue Charles Lapierre 29, 5300, Sclayn, Belgium
Department of Prehistory, University of Liège, Place du XX Août 7, 4000, Liège, Belgium

C ROLANDO

Univ. Lille, CNRS, USR 3290 - MSAP - Miniaturisation pour la Synthèse, l'Analyse et la
Protéomique, F-59000 Lille, France
Shrieking Sixties, 1-3 Allée Lavoisier, 59650 Villeneuve-d'Ascq, France

C TOKARSKI

University of Bordeaux, Proteome Platform, 146 rue Leo Saignat, F-33076 Bordeaux, France

P AUGUSTE

Univ. Lille, CNRS, UMR 8198 - Evo-Eco-Paléo, Unité Evolution, Ecologie et Paléontologie,
F-59000 Lille, France

Corresponding author: fabrice.bray@univ-lille.fr

Content:

Detailed experimental protocols for proteomics analysis.

Faunal assemblage from Scladina Cave

Figure S1.1 – S1.10 pictures of the bones from Scladina Cave.

Materials and methods

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 Sigma-Aldrich (Saint-Louis, Missouri, USA) and used without purification. All solvents were MS analytical grade.

Protein extraction from bones and delipidation.

Bone fragments were mechanically ground in fine powder with a pestle in an agate mortar. 5 mg of bone powder was transferred to a 2 mL Eppendorf® tube (Eppendorf; Westbury, NY, USA) and 1 mL of demineralizing solution (5% trifluoroacetic acid (TFA) in water) was added. The solution was stirred during 24 h at 4 °C and then centrifuged for 10 min at 10,000 g at room temperature. The bone powder and the demineralizing solution were carefully separated. The demineralizing solution was evaporated at room temperature with a SpeedVac™ Concentrator (Eppendorf™ Concentrator Plus, Eppendorf, Hamburg, Germany). 1 mL of pure water was added to the centrifuged bone powder and the solution was stirred for 10 min at 4 °C. The solution was centrifuged for 5 min at 10,000 g, then the supernatant was removed. This step was repeated a second time. The delipidation was performed by adding 100 µL of water to the bone powder, then 900 µL of a chloroform / methanol solution (2/1 v/v) (Folch et al., 1957). The sample was stirred for 1 h at 4 °C and then centrifuged for 10 min at 10,000 g. The lower phase that contains the lipids was removed. The bone powder was washed again twice with water. Residual water in bone powder was then evaporated at room temperature with a SpeedVac™ Concentrator. Then 200 µL of lysis buffer (8 M urea, 4 % SDS, 0.2% DCA, 50 mM DTT, 100 mM ammonium bicarbonate pH 8.8) were added to the bone powder and on the evaporated demineralization solution. The suspensions were incubated overnight with shaking at 4 °C, before the eFASP digestion step.

eFASP digestion

eFASP digestion was used (Erde et al., 2014). Before use, 0.5 mL Amicon® ultra centrifugal filters with a cut-off of 10 kDa (EMD Millipore, Darmstadt, Germany) were freshly incubated overnight with the passivation solution containing 5% (v/v) Tween®-20. Amicon® ultra centrifugal filters were washed in a water bath for 20 min each four times. The bone powder suspension and the demineralized fraction were transferred to separate Amicon® filters, 100 µL of exchange buffer was added (8 M urea, 0.2% DCA (deoxycholic acid) w/w, 100 mM ammonium bicarbonate pH 8.8) and then Amicon® filters centrifuged for 30 min at 10,000 g. The filtrates were discarded then 200 µL of exchange buffer were added again into the Amicon® filters which were centrifuged. This step was repeated twice. The 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® filters were centrifuged for 30 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 30 min at 10,000 g and the filtrate was discarded. 200 µl of digestion buffer (0.2% DCA (deoxycholic acid) w/w, 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® filters were transferred to new 2 mL tubes. 100 µL of digestion buffer and 40 µl of trypsin / LysC (Promega, Madison, USA) were added and incubated with shaking 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 tube by centrifugation for 15 min at 10,000 g. In order to obtain a maximum of peptides, the Amicon® filters were washed twice with 50 µl of ammonium bicarbonate solution (50 mM pH 8.8). The filtrates containing all the peptides were transferred to 1.5 mL Eppendorf® tubes. 200 µL of ethyl acetate and 2.5 µL of TFA were added inducing the appearance of a white precipitate. At once, 800 µL of ethyl acetate were added again and the Eppendorf® tubes were centrifuged for 10 min at 10,000 g. The organic phase was eliminated. This step was repeated twice. The Eppendorf® tube were placed for 5 min at 60 °C in a heating block tubes (SBH130, Stuart, Staffordshire, UK) to remove the remaining ethyl acetate. The samples were evaporated to dryness at room temperature with a SpeedVac™ Concentrator. The wall of the Eppendorf® tubes were rinsed with 100 µl of a methanol / water (50/50 v/v) mixture and again evaporated to dryness. For mass spectrometry analysis, the sample was dissolved again in 10 µl of water solution containing 0.1% of formic acid. 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). The samples were diluted at a concentration of 1 µg/µL before 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) (Helle et al., 2019). 1 µl of the peptide mixture at a concentration of 1 µg/µL was injected with solvent A (5% acetonitrile and 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 in 160 minutes and then 100% of solution B in 5 minutes. The column was washed for 5 minutes with buffer B 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 180 min. The LC runs were acquired in positive ion mode with MS scans from m/z 350 to 1,500 in the Orbitrap mass analyser at 70,000 resolution at m/z 400. The automatic gain control was set at 1E6.

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 5E5, and the normalized collision energy was set to 28 eV. Dynamic exclusion was set at 90 s and ions with 1 and more than 8 charges were excluded.

Data analysis

Proteomics data were processed with Mascot (version 2.5.1, Matrix Science, London, UK) against NCBI database (NCBiProt_20210815) restricted to Mammalian (12,254,278 sequences). Error tolerant searches with up to three missed cleavages, 10 ppm mass error for MS and 15 mmu for MS/MS. Cysteine carbamidomethylation was set as fixed modification. Methionine oxidation and asparagine, glutamine deamidation were selected as variable modifications. The second bioinformatics analysis was performed with PEAKS X (Bioinformatics software, Waterloo, Canada) against a home-made database containing 1,765 collagen sequences extracted from NCBI database (All_Collagen, downloaded 08-2019) restricted to Mammalian. Precursor's mass tolerance was fixed to 10 ppm and fragment ion mass tolerance to 0.02 Da. Three missed cleavages were allowed. The same post-translational modifications (PTMs) above were allowed plus hydroxylation of amino acids (RYFPNKD) 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 $-10\log P \geq 20$, 1% peptide False Discovery Rate (FDR), PTM with Ascore = 20, mutation ion intensity = 5% and *Denovo* ALC $\geq 50\%$. Peptides with amino acids substitutions was filtered with minimal intensity set as $1E+7$. Peptides identified on collagen I alpha 1 and I alpha 2 were aligned against the NCBI non redundant protein sequence (all non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF excluding environmental samples from WGS projects containing 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. The specific peptides were validated with a score of 100% of identity and full query coverage. Fragment ion spectra for specific peptides, deamidated and oxidative peptides and novel amino acid substitution biomarkers were manually examined for quality (Supplementary data 3, 4 and 5). COL1A1 and COL1A2 from *Panthera tigris altaica*, *Panthera pardus* X2, X3, X4 were aligned with MEGAX version 11 (Molecular Evolutionary Genetics Analysis, <https://www.megasoftware.net/home>) (Tamura et al., 2021) with ClustalW algorithm and default parameters. The mass spectrometry proteomics data have been deposited on the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository (Vizcaíno et al., 2012) with the data set identifier PXD031386

Faunal assemblage

The faunal assemblage from Scladina Cave is composed of herbivorous species including horse (*Equus ferus*), red deer (*Cervus elaphus*), fallow deer (*Dama dama*), chamois (*Rupicapra rupicapra*), aurochs or bison (*Bos primigenius* or *Bison priscus*), woolly rhinoceros (*Coelodonta antiquitatis*). Omnivorous species including cave bear (*Ursus spelaeus*), brown bear (*Ursus arctos*), red fox (*Vulpes vulpes*) and arctic fox (*Alopex lagopus*), as well as carnivorous species including cave lion (*Panthera spelaea*), panther (*Panthera pardus*), cave hyaena (*Crocuta crocuta spelaea*) and wolf (*Canis lupus*). Modified bones with cutting marks, filleting marks or breakage damages were also recovered in the cave (Otte et al., 2000).

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



Figure S1.1. Picture of sample Sc-1.



Figure S1.2. Picture of sample Sc-2.

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



Figure S1.3. Picture of sample Sc-3.



Figure S1.4. Picture of sample Sc-4.

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



Figure S1.5. Picture of sample Sc-5.



Figure S1.6. Picture of sample Sc-6.

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



Figure S1.7. Picture of sample Sc-7.



Figure S1.8. Picture of sample Sc-8.

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



Figure S1.9. Picture of sample Sc-9.



Figure S1.10. Picture of sample Sc-10.

Reference

- Erde, J., R. R. Ogorzalek Loo, and J. A. Loo, 2014, Enhanced FASP (eFASP) to increase proteome coverage and sample recovery for quantitative proteomic experiments, *Journal of proteome research*, **13**(4), 1885-1895.
- Folch, J., M. Lees, and G. S. Stanley, 1957, A simple method for the isolation and purification of total lipides from animal tissues, *Journal of biological chemistry*, **226**(1), 497-509.
- Helle, S., F. Bray, J.-L. Putaux, J. Verbeke, S. Flament, C. Rolando, C. D'hulst, and N. Szydlowski, 2019, Intra-sample heterogeneity of potato starch reveals fluctuation of starch-binding proteins according to granule morphology, *Plants*, **8**(9), 324.
- Otte, M., M. Patou-Mathis, and D. Bonjean, 2000, Les espèces chassées et consommées par l'homme en couche 5, in *Bulletin de la Société préhistorique française*, 669-670, A. Tuffreau, ed., Université de Liège Liège.
- Tamura, K., G. Stecher, and S. Kumar, 2021, MEGA11: Molecular Evolutionary Genetics Analysis Version 11, *Mol Biol Evol*, **38**(7), 3022-3027.
- Vizcaíno, J. A., R. G. Côté, A. Csordas, J. A. Dienes, A. Fabregat, J. M. Foster, J. Griss, E. Alpi, M. Birim, and J. Contell, 2012, The PRoteomics IDentifications (PRIDE) database and associated tools: status in 2013, *Nucleic acids research*, **41**(D1), D1063-D1069.