

RESEARCH ARTICLE

Deep palaeoproteomic profiling of archaeological human brains

Alexandra Morton-Hayward^{1,2*}, Sarah Flannery¹, Iolanda Vendrell¹, Roman Fischer¹

1 Department of Earth Sciences, University of Oxford, Oxford, United Kingdom, **2** Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom

* alexandra.morton-hayward@earth.ox.ac.uk



Abstract

Palaeoproteomics leverages the persistence, diversity, and biological import of ancient proteins to explore the past, and answer fundamental questions about phylogeny, environment, diet, and disease. These insights are largely gleaned from hard tissues like bone and teeth, as well-established protocols exist for extracting ancient proteins from mineralised tissues. No such method, however, exists for the soft tissues, which are underexplored in palaeoproteomics given permission for destructive analysis routinely depends on a proven methodology. Considering less than one-tenth of all human proteins are expressed in bone, compared to three-quarters in the internal organs, the amount of biological information presently inaccessible is substantial. We address this omission with an optimised LC-FAIMS-MS/MS workflow yielding the largest, most diverse palaeoproteome yet described. Using archaeological human brains, we test ten protocols with varied chemistries and find that urea lysis effectively disrupts preserved membrane regions to expose low-abundant, intracellular analytes. Further, we show that ion mobility spectrometry improves unique protein identification by as much as 40%, and represents a means of “cleaning” dirty archaeological samples. Our methodology will be useful for improving protein recovery from a range of ancient tissues and depositional environments.

OPEN ACCESS

Citation: Morton-Hayward A, Flannery S, Vendrell I, Fischer R (2025) Deep palaeoproteomic profiling of archaeological human brains. PLoS One 20(5): e0324246. <https://doi.org/10.1371/journal.pone.0324246>

Editor: Qi Wu, Aarhus University, DENMARK

Received: January 10, 2025

Accepted: April 22, 2025

Published: May 28, 2025

Copyright: © 2025 Morton-Hayward et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: All LC-(FAIMS)-MS/MS files are available from the PRIDE database (accession number PXD062254; <https://www.ebi.ac.uk/pride/archive/projects/PXD062254>).

Funding: Morton-Hayward was supported by the UKRI Natural Environment Research Council (NE/S007474/1). The funder had

Introduction

Proteins are long-lived biomolecules able to persist for millions of years in the archaeological record [1–3]. Their longevity derives in part from their structure, comprising long chains of 20 standard amino acids (each of which possesses a chemically distinct side chain) folded into complex secondary, tertiary and quaternary structures, which are compounded by an array of possible post-translational modifications (PTMs). This compact architecture packs the same amount of sequence information as DNA into approximately one-sixth the number of atoms and, with fewer chemical bonds, degrades more slowly than the double helix that encodes it [4]. Ancient protein sequences thus trace the same phylogenetic relationships as DNA,

no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

with the additional advantage that preservation of *in vivo* PTMs yields physiological information on the lived experience of long-dead lineages [5]. That is, where DNA provides the potential, proteins provide the record. Their persistence and biochemical diversity make proteins an ideal vehicle for navigating the recent and deep past [6], and the challenge of palaeoproteomics is to retrieve, identify, and make sense of their remnants as they transition from the biosphere to the lithosphere [4].

To be detected and analysed, proteins must first be extracted. Sample preparation remains the major bottleneck in any bottom-up proteomics workflow given the complexity of biological samples and the multi-stage processes associated with it [7–9], which typically include cell lysis, protein denaturation, disulfide bond reduction, sulfhydryl group alkylation, and digestion of proteins into peptides for efficient separation by liquid chromatography (LC) and detection by mass spectrometry (MS) [10]. This complexity is only exacerbated in an archaeological context, where the action of time has further depleted and degraded the ancient proteome [11]. Nonetheless, ancient proteomes have been retrieved from hard tissues like bone and teeth in diverse taxa [4], and a handful of studies have investigated the efficiency of different methods for extracting ancient proteins from these remains [12–14]. However, archaeological soft tissues remain underexplored as repositories of ancient proteins: those few to-date have been restricted largely to hair [15] and skin [16–18], and no method optimisation has been conducted beyond extraction from keratinised and collagenous tissues. Ancient hair, skin and bone proteomes largely comprise keratins and collagens, and less than 9% of all human proteins are expressed in bone (1,730/20,162) [19]. By contrast, 76% are expressed in the brain (15,331/20,162) [20], and the ancient brain proteome is thus an order of magnitude more diverse, reflective of its functional complexity [21]. The same is likely true of other internal organs preserved in the archaeological record, such as the intestines (75%) [22,23], stomach (72%) [24,25], kidneys (71%) [26,27], lungs (71%) [28,29] and heart (68%) [30,31].

More than 4,400 human brains have been excavated from an array of environments worldwide, some up to 12,000 years old [32]; just two, however, have been probed by LC-MS/MS [21,33]. Soft tissues of the internal organs (and particularly the brain) comprise substantially greater proportions of lipid-linked membrane proteins than hard tissues [20]. Anchored in the lipid bilayer environment, membrane proteins comprise one or more alpha-helical domains [34] whose hydrophobic nature can cause aggregation in aqueous solutions [35], which makes sample preparation challenging. Membrane proteins are often extracted and kept soluble with strong denaturants (e.g., urea) or ionic detergents (e.g., sodium dodecyl sulphate [SDS]) which, when maintained above their critical micelle concentration, effectively disrupt the lipid bilayer [36]. While such agents may facilitate extraction, they can inhibit proteolysis and interfere with separation by LC and detection by MS, hampering protein recovery [9,37]. As such, a range of sample clean-up methods exist at both the protein- and peptide-level (i.e., conducted before or after proteolysis) that aim to remove reagents that would otherwise negatively impact downstream analyses [38]. However, additional steps cause sample losses, which it is crucial to minimise in

cases where low starting material availability must be a consideration [39], such as with archaeological tissues. Moreover, choices made at each stage of the proteomics workflow affect the composition of the recovered proteome: for example, reagents can selectively enrich or deplete functional classes of proteins and cellular components, depending on the underlying chemistry [40–42].

We test ten protocols for ancient protein recovery from archaeological brain tissues using a bottom-up LC-MS/MS approach, comparing the effectiveness of a range of extraction buffers and clean-up strategies (Fig 1). The former include urea, SDS and sodium laurate (SL); while the latter include the in-StageTip method (iST) [43], solid-phase extraction (SPE), suspension trapping (S-Trap) [44], filter-aided sample preparation (FASP) [45], single-pot solid-phase-enhanced sample preparation (SP3) [46], and in-gel purification with SDS-polyacrylamide gel electrophoresis (SDS-PAGE). We assess the use of a high-field asymmetric-waveform ion mobility spectrometry (FAIMS) source at different compensation voltages to reduce chemical noise and improve detection of low-abundant proteins. Finally, the range of laboratory-based options available to the protein chemist is paralleled by the array of analytical software available to the bioinformatician: we test protein identification and relative quantification with DIA-NN [47], FragPipe [48], MaxQuant [49] and PEAKS® Studio (Bioinformatics Solutions, Inc.). We assess the efficacy, efficiency (i.e., performance relative to ease of use), and suitability of all options for broad application in palaeoproteomics, comparing protocols on the basis of numbers of identified precursors, peptides and proteins, intensity-based absolute quantification (iBAQ), protein scoring, hydropathy and isoelectric point, levels of reagent-based contamination, and functional and cellular enrichment biases.

Materials and methods

Archaeological brain material

Approximately 4,500 individuals were excavated from the site of the former Blackberry Hill Hospital in the Fishponds area of Bristol, UK (2018–2022, Cotswold Archaeology), one in ten of which evidenced preserved brain tissues ($n=456$). Originally built as Stapleton Prison (1779–1814) to house prisoners of war from Britain's conflicts with France, Spain, the Netherlands and the USA throughout the late 18th century, the site became a hospital during Bristol's 1832 cholera outbreak, and was converted to the Stapleton Workhouse shortly thereafter (1837–1890). In the 20th century, it evolved into Manor Park Hospital, later Blackberry Hill Hospital, before its closure in 2007.

Immediately post-excavation, brains were refrigerated at 4 °C in airtight containers, without washing or preservative treatment, and acquired by the Duckworth Laboratory (University of Cambridge). One brain (SK 17999) was selected for investigation based on the exceptional quality of its preservation, possessing identifiable cerebral hemispheres with intact sulci and gyri that facilitated confident and reproducible sampling. A single cortical biopsy was taken from the anterior right frontal lobe, which was separated into ten 50 mg samples; these were stored at –80 °C prior to analysis. Permission for destructive sampling was sought and obtained from the Duckworth Laboratory, and ethical approval from the Central University Research Ethics Committee at the University of Oxford (R80617/RE001). All necessary permits were obtained for the described study, which complied with all relevant regulations.

In-solution

PreOmics® BeatBox iST. The sample was processed according to the iST Sample Preparation Kit protocol (v.4.0; PreOmics®) [50]. A 50 mg aliquot of brain was lysed with 50 mg glass beads in 100 µL of LYSE, using BeatBox sonication for 10 cycles of 30 sec. Proteins were further denatured at 95 °C and 1,000 rpm for 10 min in an Eppendorf ThermoMixer® C (Thermo Fisher Scientific). Digestion was performed with 50 µL DIGEST at 37 °C and 500 rpm for 3 hr. The sample was acidified with 100 µL STOP then clarified at 20 °C and 500 rpm for 1 min, followed by centrifugation at 16,000 rcf for 1 min. Peptides were purified using the PreOmics® iST method. Eluted peptides were dried by vacuum centrifugation for storage at –20 °C, then reconstituted in 50 µL LC-LOAD at 20°C and 500 rpm for 5 min prior to LC-MS analysis.

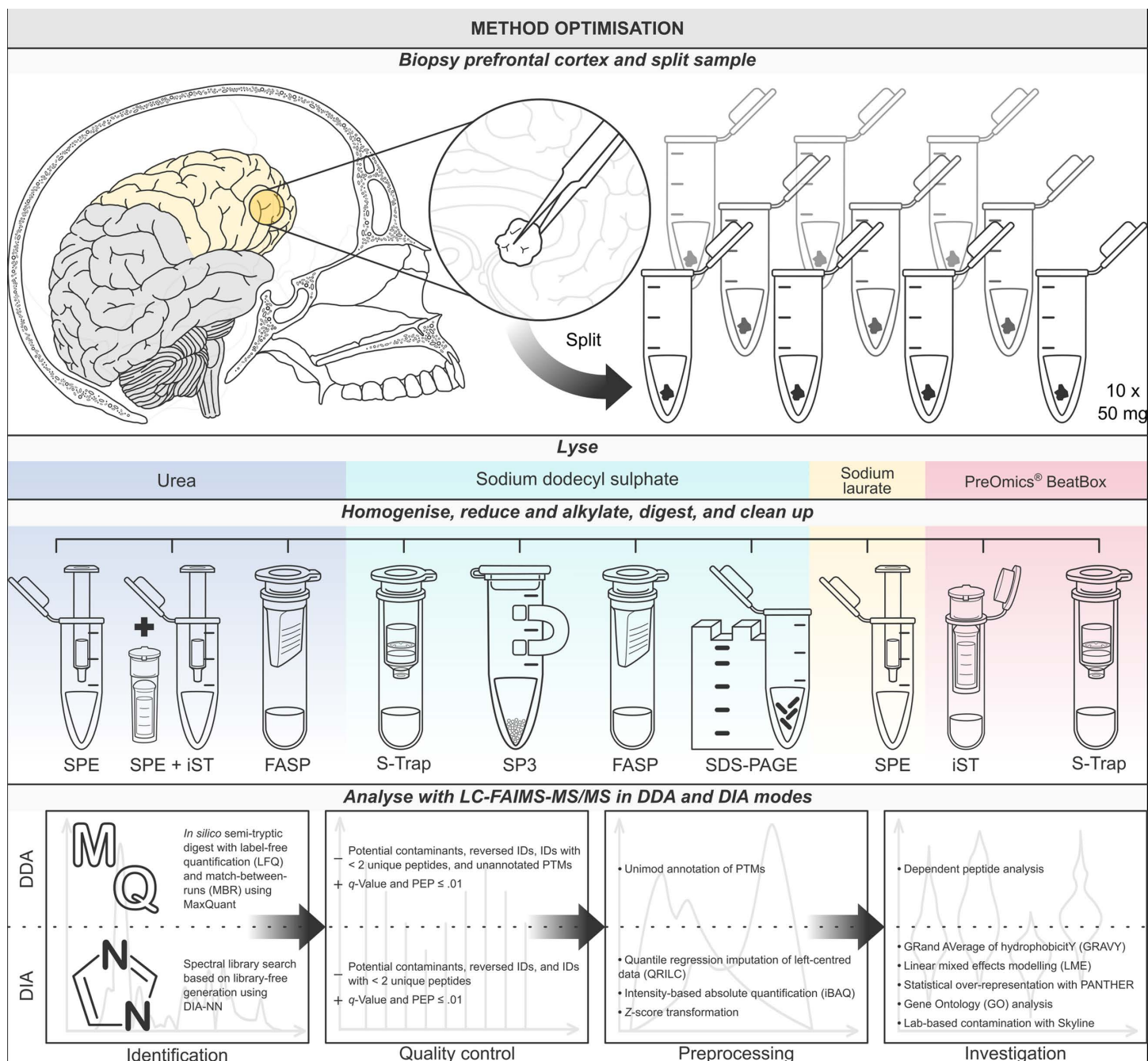


Fig 1. Overview of the study design. A single cortical biopsy was taken from the anterior right frontal lobe and separated into ten 50 mg samples, which were lysed with four different extraction methods (urea, SDS, SL and PreOmics® BeatBox) before homogenisation by bead-beating. After reduction, alkylation, and digestion with trypsin, samples were processed with one of eight clean-up methods either in solution (iST, SPE, SPE + iST), in-gel (SDS-PAGE), on-filter (FASP, S-Trap) or on-bead (SP3). Samples were analysed by LC-FAIMS-MS/MS in DDA and DIA modes and, based on preliminary testing with a range of software (see Table S1 in [S1 File](#)), the final DDA dataset was processed in MaxQuant [49], and the DIA dataset in DIA-NN [47].

<https://doi.org/10.1371/journal.pone.0324246.g001>

Urea with SPE (and iST). A 50 mg aliquot of brain was lysed in 8 M urea, 100 mM triethylammonium bicarbonate (TEAB) pH 8.5, using a Precellys® 24 Touch homogeniser (Bertin Technologies) for three cycles of 30 sec at 5,500 rpm. The sample was placed on ice between cycles to reduce heating. Proteins were further denatured at 20 °C and 1,000 rpm for 30 min in an Eppendorf ThermoMixer® C (Thermo Fisher Scientific). Cysteine reduction and alkylation were performed at 20 °C and 1,000 rpm for 30 min with 10 mM tris(2-carboxyethyl)phosphine (TCEP) and 50 mM iodoacetamide (IAA), respectively. The urea concentration was diluted to 1.5 M with 50 mM TEAB pH 8.5. Digestion was performed with 1 µg trypsin (MS-grade, Promega) overnight at 37 °C and 1,250 rpm. The sample was acidified with TFA to 1% then clarified by centrifugation at 22,000 rcf for 10 min. Peptides were diluted 10-fold in 0.1% trifluoroacetic acid (TFA), 2% acetonitrile (ACN) then purified by SPE (SOLA™ SPE, Thermo Fisher Scientific). The sample was split: subsample A underwent no further clean-up, and eluted peptides were dried by vacuum centrifugation for storage at –20 °C, then reconstituted in 0.1% formic acid (FA), 3% ACN prior to LC-MS analysis. Subsample B was subjected to the PreOmics® iST method as described in the iST Sample Preparation Kit protocol (v.4.0; PreOmics®) [50], and eluted peptides dried by vacuum centrifugation for storage at –20 °C, then reconstituted in 50 µL LC-LOAD at 20 °C and 500 rpm for 5 min prior to LC-MS analysis.

Sodium laurate with SPE. A 50 mg aliquot of brain was lysed in 1% sodium laurate (SL), 100 mM TEAB pH 8.5, using a Precellys® 24 Touch homogeniser (Bertin Technologies) for three cycles of 30 sec at 5,500 rpm. The sample was placed on ice between cycles to reduce heating. Proteins were further denatured in an ultrasonic bath at 4 °C for 10 min. Cysteine reduction and alkylation were performed at 20 °C without shaking for 30 min with 10 mM TCEP and 20 mM IAA, respectively. Digestion was performed with 1 µg trypsin (MS-grade, Promega) at 37 °C and 1,000 rpm for 3 hr. The sample was acidified with TFA to 1% then clarified by centrifugation at 22,000 rcf for 5 min. To remove SL, a volume of ethyl acetate equal to the volume of supernatant was added and the sample centrifuged at 15,700 rcf for 5 min, before removal of the lower phase and repetition of this step. Peptides were diluted 10-fold in 0.1% TFA, 2% ACN then purified by SPE (SOLA™ SPE, Thermo Fisher Scientific). Eluted peptides were dried by vacuum centrifugation for storage at –20 °C, then reconstituted in 0.1% FA, 3% ACN prior to LC-MS analysis.

In-gel

SDS-PAGE. A 50 mg aliquot of brain was lysed in 5% SDS, 100 mM TEAB pH 8.5, using a Precellys® 24 Touch homogeniser (Bertin Technologies) for three cycles of 30 sec at 5,500 rpm. The sample was placed on ice between cycles to reduce heating. NuPAGE™ LDS (4x) sample buffer (Invitrogen™, Thermo Fisher Scientific) was added with 200 mM dithiothreitol (DTT) to a final concentration of 1%, before incubating at 20 °C and 2,000 rpm for 1 hr. The sample was loaded into six lanes of a 1 mm NuPAGE™ Bis-Tris gel, mounted in a Mini Gel Tank and loaded with NuPAGE™ MES SDS (20x) running buffer (all Invitrogen™, Thermo Fisher Scientific). The gel was run at 120 V for 10 min before removal and staining with InstantBlue™ dye (Sigma-Aldrich). Each gel lane was excised into approximately 1 mm² cubes and combined, before submersion in 5% acetic acid at 20 °C and 1,000 rpm overnight in an Eppendorf ThermoMixer® C (Thermo Fisher Scientific). Cysteine reduction was performed with 10 mM DTT at 20 °C without shaking for 30 min, followed by alkylation with 50 mM IAA at 20 °C without shaking for 30 min. The gel pieces were dehydrated with ACN at 20 °C and rehydrated with 100 mM ammonium bicarbonate (ambic) at 20 °C prior to digestion, which was performed with 1 µg trypsin (MS-grade, Promega) in 50 mM ambic overnight at 37 °C without shaking. Peptides were eluted with 5% FA, 50–85% ACN, dried by vacuum centrifugation for storage at –20 °C, then reconstituted in 0.1% FA, 3% ACN prior to LC-MS analysis.

On-filter

PreOmics® BeatBox with S-Trap™ micro. A 50 mg aliquot of brain was lysed in 5% SDS reducing buffer, 10 mM TCEP and 100 mM TEAB pH 8.5, using BeatBox sonication (PreOmics®) for 10 cycles of 30 sec. Cysteine alkylation was performed with 20 mM IAA at 20 °C without shaking for 30 min. The sample was acidified with phosphoric acid to 1.2% concentration, and loaded onto an S-Trap™ micro (Protifi, Biosys Technologies, Inc.). Digestion was performed on-filter

in 50 mM TEAB pH 8.5 with 1 µg trypsin (MS-grade, Promega) overnight at 37 °C without shaking. The sample was eluted sequentially with 50 mM TEAB pH 8.5, 0.2% FA, then 0.2% FA, 50% ACN. Eluted peptides were dried by vacuum centrifugation for storage at −20 °C, then reconstituted in 0.1% FA, 3% ACN prior to LC-MS analysis.

SDS with S-Trap™ micro. A 50 mg aliquot of brain was lysed in 5% SDS reducing buffer, 10 mM TCEP and 100 mM TEAB pH 8.5, using a Precellys® 24 Touch homogeniser (Bertin Technologies) for three cycles of 30 sec at 5,500 rpm. The sample was placed on ice between cycles to reduce heating. Cysteine alkylation was performed with 20 mM IAA at 20 °C without shaking for 30 min. The sample was acidified with phosphoric acid to 1.2% concentration, and loaded onto an S-Trap™ micro (Protifi, Biosys Technologies, Inc.). Digestion was performed on-filter in 50 mM TEAB pH 8.5 with 1 µg trypsin (MS-grade, Promega) overnight at 37 °C without shaking. The sample was eluted sequentially with 50 mM TEAB pH 8.5, 0.2% FA, then 0.2% FA, 50% ACN. Eluted peptides were dried by vacuum centrifugation for storage at −20 °C, then reconstituted in 0.1% FA, 3% ACN prior to LC-MS analysis.

Urea with FASP. A 50 mg aliquot of brain was lysed in 8 M urea, 100 mM TEAB pH 8.5, using a Precellys® 24 Touch homogeniser (Bertin Technologies) for three cycles of 30 sec at 5,500 rpm. The sample was placed on ice between cycles to reduce heating. The sample was filtered with a Sartorius Vivacon™ 500 30 kDa molecular-weight cut-off (MWCO) filter (Thermo Fisher Scientific), centrifuged at 14,300 rcf for 10 min. Cysteine reduction and alkylation were performed with 10 mM TCEP and 50 mM IAA respectively, on-filter in 8 M urea, 100 mM TEAB pH 8.5 at 20 °C without shaking for 30 min. The filter was washed three times with 50 mM TEAB pH 8.5, and digestion performed with 1 µg trypsin (MS-grade, Promega) overnight at 37 °C without shaking. Peptides were eluted with three cycles of centrifugation at 14,300 rcf for 10 min, with dilution in 0.1% TFA, 50% ACN. Eluted peptides were dried by vacuum centrifugation for storage at −20 °C, then reconstituted in 0.1% FA, 3% ACN prior to LC-MS analysis.

SDS with FASP. A 50 mg aliquot of brain was lysed in 5% SDS, 100 mM TEAB pH 8.5, using a Precellys® 24 Touch homogeniser (Bertin Technologies) for three cycles of 30 sec at 5,500 rpm. The sample was placed on ice between cycles to reduce heating. The sample was filtered with a Sartorius Vivacon™ 500 30 kDa MWCO filter (Thermo Fisher Scientific), centrifuged at 14,300 rcf for 10 min. Cysteine reduction and alkylation were performed with 10 mM TCEP and 50 mM IAA respectively, on-filter in 8 M urea, 100 mM TEAB pH 8.5 at 20 °C without shaking for 30 min. The filter was washed three times with 50 mM TEAB pH 8.5, and digestion performed with 1 µg trypsin (MS-grade, Promega) overnight at 37 °C without shaking. Peptides were eluted with three cycles of centrifugation at 14,300 rcf for 10 min, with dilution in 0.1% TFA, 50% ACN. Eluted peptides were dried by vacuum centrifugation for storage at −20 °C, then reconstituted in 0.1% FA, 3% ACN prior to LC-MS analysis.

On-bead

SP3. A 50 mg aliquot of brain was lysed in 0.5% SDS using a Precellys® 24 Touch homogeniser (Bertin Technologies) for three cycles of 30 sec at 5,500 rpm. The sample was placed on ice between cycles to reduce heating. Cysteine reduction was performed with 5 mM DTT at 20 °C without shaking for 30 min, followed by alkylation with 20 mM IAA at 20 °C without shaking for 30 min. The sample was combined with 3 µL of SP3 beads (1:1 hydrophobic:hydrophilic) and ACN added to a final concentration of 70%, before mixing at 20 °C and 1,000 rpm for 18 min. The sample was placed on a magnetic rack and the beads allowed to settle for 2 min, before being washed three times with 100% ACN. Beads were resuspended in 50 mM TEAB pH 8.5 and digestion performed with 1 µg trypsin (MS-grade, Promega) overnight at 37 °C and 1,000 rpm. The sample was acidified with TFA to 1% and eluted peptides were dried by vacuum centrifugation for storage at −20 °C, then reconstituted in 0.1% FA, 3% ACN prior to LC-MS analysis.

Data acquisition

LC-FAIMS-MS/MS was performed on a Vanquish Neo UHPLC coupled to an Orbitrap™ Ascend Tribrid™ mass spectrometer equipped with a FAIMS Pro Duo interface (all Thermo Fisher Scientific). This system was benchmarked against a Hela

cell lysate to ensure optimal performance. The Vanquish Neo was operated in “Trap and Elute” mode, using a PepMap™ Neo Trap Cartridge (5 mm x 300 µm) and EASY-Spray™ PepMap™ Neo UHPLC Column (50 cm x 75 µm, 1500 bar; both Thermo Fisher Scientific). Tryptic peptides were trapped and separated with a 60 min linear gradient; from 2% to 18% buffer B (0.1% FA in ACN) in buffer A (0.1% FA in H₂O) over 40 min, from 18% to 35% B over 20 min, and from 35% to 99% B over 15 min at 300 nL/min flow rate.

In data dependent acquisition (DDA) mode, MS1 spectra were acquired in the Orbitrap™ at 120K resolution between 380 and 1500 *m/z*, with an automatic gain control (AGC) target of 4e5 ions, a maximum injection time of 251 ms, radio frequency (RF) lens array at 30%, and advanced peak determination toggled on. Precursor ions were fragmented using high energy collisional dissociation with 30% normalised collision energy (NCE). MS2 spectra were acquired in the ion trap with rapid scan mode, a quad isolation window of 1.2 *m/z*, an AGC target of 5e3 ions, and a maximum injection time of 50 ms. The FAIMS Pro Duo was operated at standard resolution with a carrier gas flow rate of 3.8 L/min and cycling through two compensation voltages (CV, −40 and −55 V), with one sec per CV. In data independent acquisition (DIA) mode, MS1 spectra were acquired in the Orbitrap™ at 45K resolution between 350 and 1650 *m/z*, with an AGC target of 5e5 ions, a maximum injection time of 91 ms, and RF lens array at 30%. MS2 spectra were acquired using the tMSn scan function at 30K resolution over 40 scan windows (with variable isolation width) covering the full 350–1650 *m/z* range, with an AGC target of 4e6 ions, an automatic maximum injection time, and 30% NCE. The FAIMS source was operated at standard resolution with a carrier gas flow rate of 3.8 L/min, and compensation voltage of −45 V.

Data analysis

Eight searches of all LC-FAIMS-MS/MS raw data files acquired in DDA (*n* = 10) and DIA mode (*n* = 10) were conducted using a range of software packages for protein identification and quantification (see, Table S1 in [S1 File](#)). In all cases, data were searched against the human brain proteome with semi-specific tryptic digestion and up to three missed cleavages allowed [51], with the following modifications: carbamidomethylation of Cys (fixed), N-terminal acetylation (variable), oxidation of Met and Pro (variable), and deamidation of Asn, Gln and Arg (variable). The human brain proteome was retrieved by downloading all protein-coding genes detected in the human brain (*n* = 15,331) from the Human Protein Atlas (v.23.0, accessed 15.02.2024) [20], and converting Ensembl identifiers to UniProt accession numbers using the UniProt ID mapping tool (v. 2024.01) [52]. Having removed duplicates, the remaining proteins (*n* = 15,049) were downloaded as a FASTA file.

A search of DDA data based on *in silico* FASTA digest with label-free quantification and match-between-runs in MaxQuant (v.2.4.3.0) [49] yielded the fewest unique proteins identified at 1% FDR but the most information in terms of protein modification, and all further analyses of DDA data were performed on this dataset using the R programming language (v.2024.04.2). Global peptide and protein lists were removed of potential contaminants, reversed identifications, and identifications with < 2 unique peptides, then filtered for *q*-value and posterior error probability (PEP) ≤ .01. Modified specific and dependent peptide lists were removed of unannotated modifications and those corresponding to < 2 unique peptides, then filtered for *q*-value and PEP ≤ .05.

Spectral library search of DIA data based on library-free generation with DIA-NN (v.1.9) [47] yielded the most confident identifications at 1% FDR (mean global *Q*-value: 2.7×10^{-3}), and all further analyses of DIA data were conducted on this dataset. DIA data were preprocessed using the R package *DIAgui* (v.1.4.2) [53], which implements the *iq* package (v.1.9.12) [54] for calculating proteome-wide, label-free quantification by delayed normalisation and maximal peptide ratio extraction (*aka*. MaxLFQ) [55], and the *imputeLCMD* package for left-centred missing data imputation [56]. Precursors were filtered using the following parameters: precursor *q*-value ≤ .01, protein group *q*-value ≤ .01, and protein and gene-group *q*-values ≤ 1. Peptides and protein groups were filtered with the same parameters. iBAQ was performed on protein groups [57]: data were log₂-transformed to normalise the data range, and missing data imputed using quantile regression (QRILC) [58]. A z-score transformation on samples was subsequently performed to enable comparison of proteomic

profiles across protocols: this approach standardises the data for each protocol by giving each sample a mean of 0 and a standard deviation (SD) of 1 across all proteins [59].

GRand AVerage of hydropathY (GRAVY) scores were calculated using the Kyte-Doolittle hydrophobicity scale [60], and the average isoelectric point (pI) using Isoelectric Point Calculator (IPC) 2.0 [61]. Both measures were calculated at the protein-level for the subset of proteins identified in all protocols ($n=98$), and at the peptide-level for peptides corresponding to these proteins recovered by each protocol. This “common proteome” allowed for a comparison in terms of the relative hydropathy of recovered peptides corresponding to these proteins against a shared pool. Statistical testing was performed to probe significant differences between methods: the data failed a Shapiro-Wilk test for normality ($W[10] = .984$, $p = 2.5 \times 10^{-2}$) and Levene’s test for homogeneity of variance ($F[10] = 6.93$, $p = 7.6 \times 10^{-11}$). Given the data is not independent (all protocols having been conducted on the same biopsy), and given the inequality of proteome size across protocols, a linear mixed-effects model with Satterthwaite t -tests ($n = 10$) and *post-hoc* pairwise comparison using Bonferroni correction at 95% confidence interval (CI) was developed for both GRAVY and pI scores, using the R packages *lme4* (v.1.1.35.5) [62] and *emmeans* (v.1.10.3) [63]. It should be noted that any statistically significant differences must be treated with caution given that, due to the irreplaceable nature of archaeological material, we were unable to perform replicate extractions (see Limitations).

Over- and under-representation (i.e., the difference between the number of observed vs. expected proteins) was investigated at the protein class level using a Fisher’s exact test and Bonferroni correction with the Protein ANalysis THrough Evolutionary Relationships (PANTHER) classification system (v.19.0) [64]. Gene Ontology (GO) functional enrichment analysis was also performed in PANTHER using a Fisher’s exact test and Bonferroni correction, to investigate enrichment in cellular components [65]. Finally, levels of common laboratory contamination were assessed using Skyline (v.24.1) [66] to extract features from raw LC-FAIMS-MS/MS data corresponding to potential targets in the Molecular Contaminant List [67], and extracted ion chromatograms processed using the *HowDirty* R package (v.0.2.1) [68].

Results

Benchmarking protocol performance with ion mobility spectrometry

While the number of unique proteins varied with the search strategy used (Fig S1 in [S1 File](#)), employing a FAIMS source increased protein detection by an average of 39.0% for DIA and 26.4% for DDA regardless of the search strategy, with a compensation voltage of -45 V yielding the most identifications (510 vs. -40 V: 487; Fig S2 in [S1 File](#)). Total numbers of proteotypic precursors, peptides and proteins recovered with each protocol in DIA are summarised in the Supplementary (Table S2 in [S1 File](#)), and the total numbers of unique proteins and overlap in proteomic profiles illustrated in [Fig 2](#). The best-performing protocols in terms of protein recovery were the PreOmics® BeatBox iST (1,205) followed by urea with FASP (1,157), which also yielded the highest numbers of proteins recovered exclusively by one method (111 and 59 respectively). To the best of our knowledge, these represent the largest and most diverse palaeoproteomes identified in any archaeological material investigated to-date. SP3 yielded the fewest total proteins (104), with less than a third as many identifications as SL with SPE (339). Given the overall effectiveness of urea-based methods, we latterly tested SP3 following urea extraction, but this did not improve protein recovery appreciably (139). While more than half of identifications (53.7%, 797/1,483) were common to the best-performing methods, less than one-quarter (23.0%, 121/525) were shared by the worst-performing. Statistical over-representation testing for protein class of this common subset (Table S3 in [S1 File](#)) indicated statistically significant enrichment by almost 30x of tubulins (fold enrichment [FE]: 26.68, $p = .03$) and intermediate filaments (24.19, $p = .003$). GO functional enrichment analysis suggested that the majority of protocols (80.0%, 8/10) enriched first and foremost for cytoskeletal fibres such as these (Data S1 in [S1 Dataset](#)); however, urea-based protocols enriched principally for the mitochondrial ATP synthase complex (e.g., with SPE: 7.11, $p = .008$; with FASP: 6.58, $p = .01$).

Unsupervised hierarchical clustering of iBAQ by protocol ([Fig 3a](#)) illustrated the range of protein recovery; particularly, the extent of missing data with SP3, for which 94.5% (1,973/1,087) of proteins had a z -score < 0 (i.e., below average

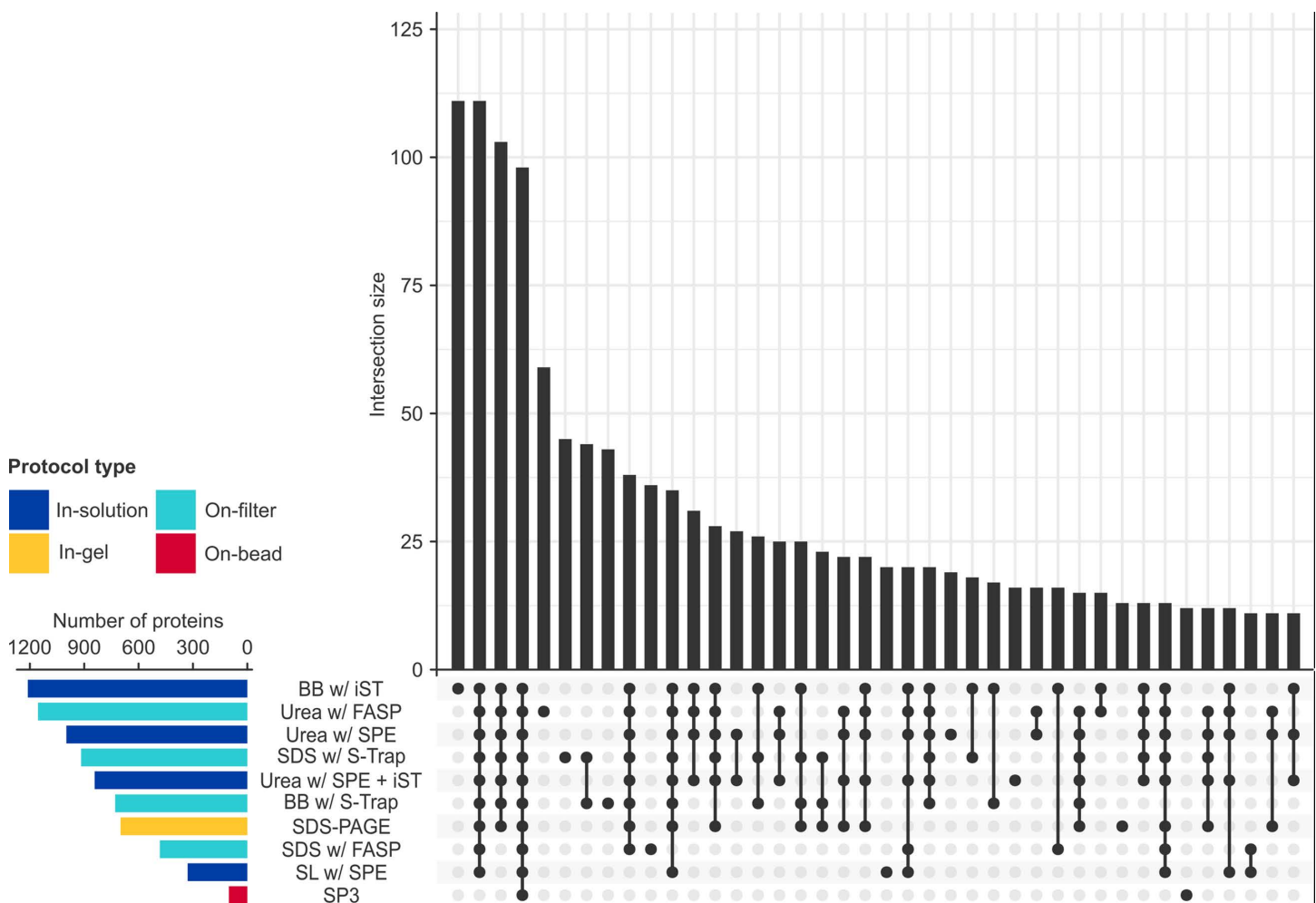


Fig 2. Total numbers of proteins retrieved with each protocol. Upset plot of the number of unique proteins at 1% FDR and q -value $\leq .01$ recovered by each protocol (*bottom left*), and the number (*top*) and nature (*bottom right*) of intersections in the data.

<https://doi.org/10.1371/journal.pone.0324246.g002>

for the dataset), compared with the highest intensities yielded by urea with FASP, for which almost the same proportion (93.2%, 1,945/2,087) had a z -score > 0 (i.e., above average for the dataset). Correlation analysis with hierarchical clustering (Fig 3b) revealed only positive correlations (i.e., $\rho > 0$) between protocols, suggesting the identification and quantification of overlapping proteomic data, albeit with varying degrees of similarity. While direct reproducibility assessments were not performed, the observed correlations suggested that protocols sharing an extraction buffer or clean-up method yield comparable iBAQ intensities, preserving protein relative abundance proportions. For instance, the strongest correlations were observed between urea extraction on-filter and in-solution (FASP with SPE: $\rho = .796$, $p < .001$; FASP with SPE + iST: $\rho = .778$, $p < .001$), followed by clean-up with S-Trap, regardless of extraction buffer ($\rho = 0.744$, $p < .001$). In contrast, weak correlation between divergent protocols (e.g., SL with SPE vs. BeatBox with S-Trap: $\rho = .037$, $p < .005$) suggested disparities in quantification, potentially due to differences in protein denaturation, solubility, digestion or purification. Multi-dimensional scaling (Fig 3c) corroborated these trends with the emergence of two clusters. The first pertains largely to what might be considered “high-performance” protocols ($\geq 1,000$ protein identifications, exceeding the upper quartile for the dataset [984 proteins]), which share either urea extraction or iST clean-up; while the second pertains largely to

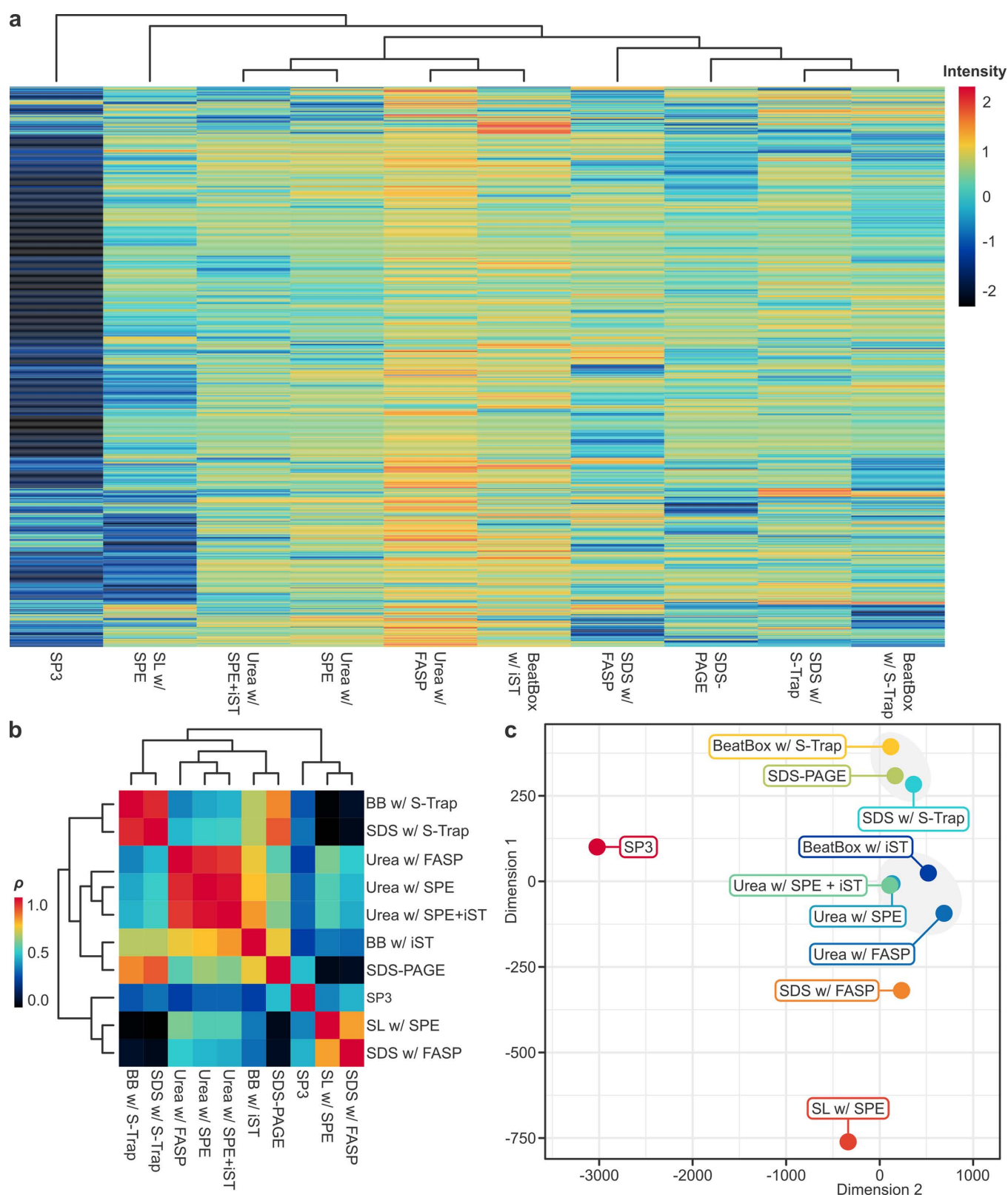


Fig 3. Unsupervised data clustering and multivariate ordination analyses. (a) Heat map with hierarchically clustered dendrograms of iBAQ by protocol. iBAQ data were \log_2 -transformed and missing data imputed using QRILC, before a z-score transformation on samples was applied. Each

row represents a recovered protein ($n=2,087$) clustered by Manhattan distance with complete linkage (dendrograms not shown), and each column a protocol clustered by Euclidean distance. **(b)** Pearson's correlation coefficient (ρ) matrix with hierarchically clustered dendrograms of iBAQ by protocol. As a normalised measure of covariance, ρ has a value between -1 (perfect negative correlation) and 1 (perfect positive correlation); however, note the restricted range illustrated (0.0 – 1.0), reflecting the fact that only positive correlations exist between protocols in our dataset. **(c)** Metric multidimensional scaling plot with Manhattan distance of iBAQ data by protocol. The distance between data points in two-dimensional Cartesian space correlates with the dissimilarity between those points. Grey circles delineate clusters.

<https://doi.org/10.1371/journal.pone.0324246.g003>

“mid-performance” protocols (500–1,000 identifications, encompassing the median for the dataset [790]), which share SDS extraction. “Low-performance” protocols (≤ 500 identifications, below the lower quartile for the dataset [530]) appear as outliers in Cartesian space.

Average missed tryptic cleavages and MS1/MS2 signal ratios may be tentatively used as gauges of digestion and fragmentation efficiency respectively (Fig 4a and 4b). On-filter clean-up methods with S-Trap evidenced among the lowest average missed cleavages (e.g., with SDS: 0.12) but the highest MS1/MS2 signal ratios (with SDS: 1.57), which suggest that, while proteolysis was relatively complete, fragmentation was not. The reverse pattern was observed for on-filter clean-up methods with FASP, which featured among the highest missed cleavages (e.g., with SDS: 0.18) and the lowest MS1/MS2 signal ratios (with SDS: 1.15). Mean sequence coverage across all protocols was relatively low ($< 5\%$; Fig 4c), as expected for ancient proteins. The highest median coverage was obtained by extraction using urea (with FASP: 4.20%; with SPE + iST: 4.05%). BeatBox with S-Trap, although a mid-performance protocol in terms of protein recovery (735 identifications), yielded the lowest median (0.00%) and mean (2.22%) coverage, alongside SP3 (median: 0.00%; mean: 2.27%).

Identifying biases in modification, hydropathy, isoelectric point and contamination

During diagenesis (the period after death and burial), ancient proteins are progressively degraded by hydrolysis and modified by deamidation (so-called, “diagenetic modifications”). DDA data was analysed to assess the extent of tryptic vs. non-tryptic peptide identification (i.e., the extent of diagenetic peptide bond cleavage), and the nature and extent of protein modification. Approximately two-thirds of the dataset comprised tryptic peptides, being cleaved C-terminal to either Lys (33.4%, 194/580) or Arg (32.8%, 190/580); of the remaining non-tryptic third, the most common cleavage sites were C-terminal to aliphatic Leu (5.2%, 30/580) and aromatic Phe (4.5%, 26/580). More than half of all non-tryptic peptides (53.6%, 105/196) corresponded to myelin proteolipid protein (Fig S3 in S1 File), with the myelin protein class more than 40-fold enriched (FE: 43.85, $p=7.7 \times 10^{-3}$). SDS with S-Trap and urea with FASP yielded the highest numbers of non-tryptic peptides (73 and 67 respectively), while BeatBox with iST yielded the fewest (29) after SP3 (21), despite being the best-performing protocol in terms of protein recovery.

The nature and extent of deamidation – a common diagenetic modification to ancient proteins, but also potentially induced during sample preparation – was assessed for both endogenous and contaminant peptides (Fig 5). In contrast to high- and mid-performance protocols, low-performance protocols evidenced greater proportions of deamidated contaminants than endogenous peptides (e.g., SP3: 42.4% vs. 15.5%). On average, approximately half of deamidated sites in endogenous peptides affected Asn (51.4%), approximately one-third Gln (30.2%), and approximately one-fifth Arg (18.3%). Functional enrichment analysis suggested that deamidation of Arg disproportionately affected cellular components associated with the myelin sheath (strength: 2.0; $p=4.5 \times 10^{-6}$). In addition to modifications explicitly searched for, a dependent peptide search identified >200 annotated modifications to ≥ 2 peptides (Fig 6 and Table S4 in S1 File), of which the most frequent were loss of water ($n=14$), acetaldehyde adduct formation ($n=11$), and metal-catalysed oxidation ($n=8$).

Statistical testing revealed significant differences in peptide hydropathy between protocols (Figs 7 and 8). GRAVY scores were calculated for the entire brain proteome ($n=15,049$; mean [SD]: -0.382 [0.392]) and for all peptides retrieved

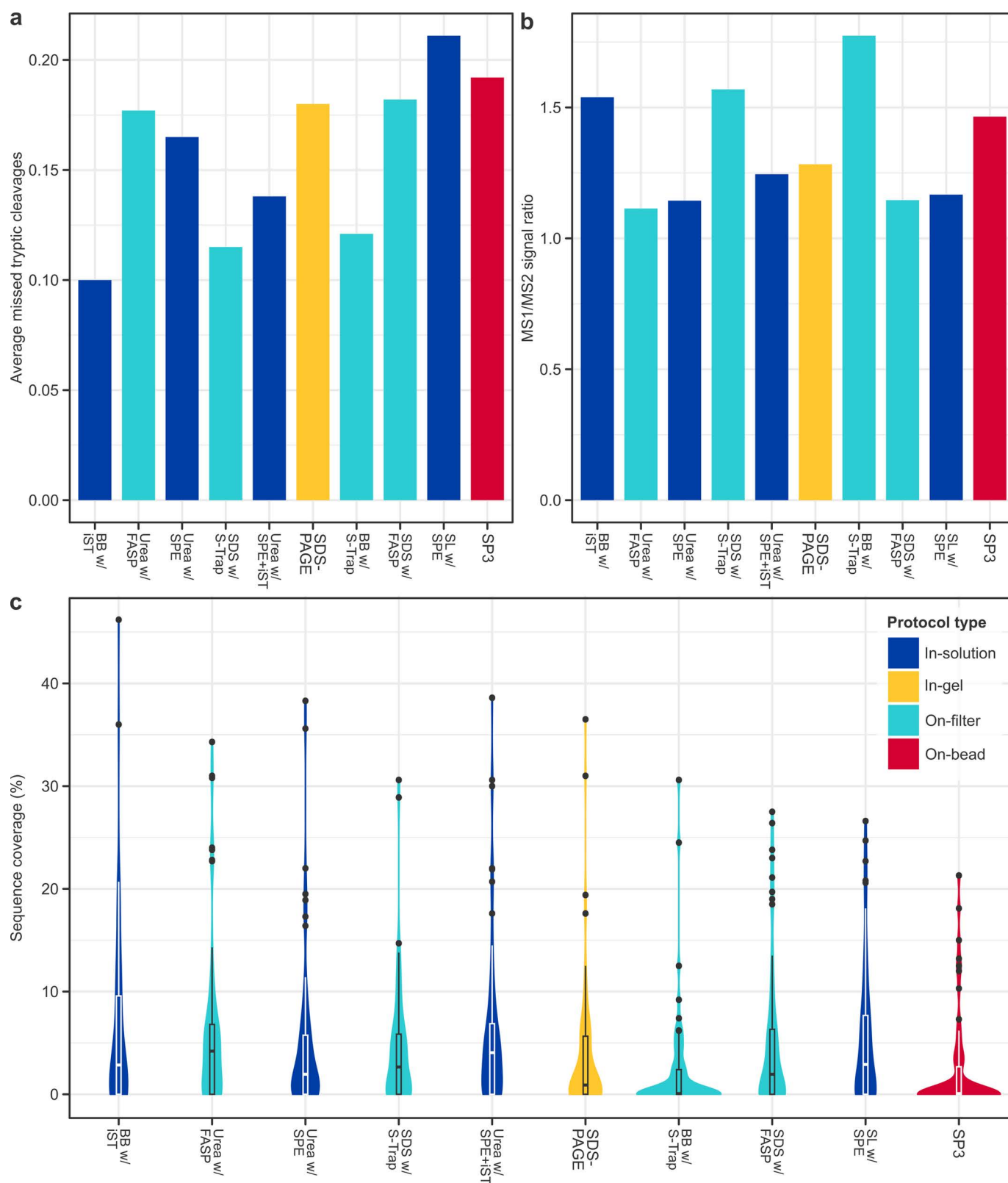


Fig 4. Measures of extraction efficiency. Bar charts of the (a) average missed tryptic cleavages and (b) MS1/MS2 signal ratios by protocol, gauges of digestion and fragmentation, respectively. (c) Violin plots of the range of sequence coverage of unique proteins identified at 1% FDR and q -value $\leq .01$,

retrieved by each protocol. Boxplots indicate the median and upper and lower quartiles, and black dots represent statistical outliers. Colours are consistent across panels, and in each panel protocols are ordered by, from left to right, highest to lowest total proteins.

<https://doi.org/10.1371/journal.pone.0324246.g004>

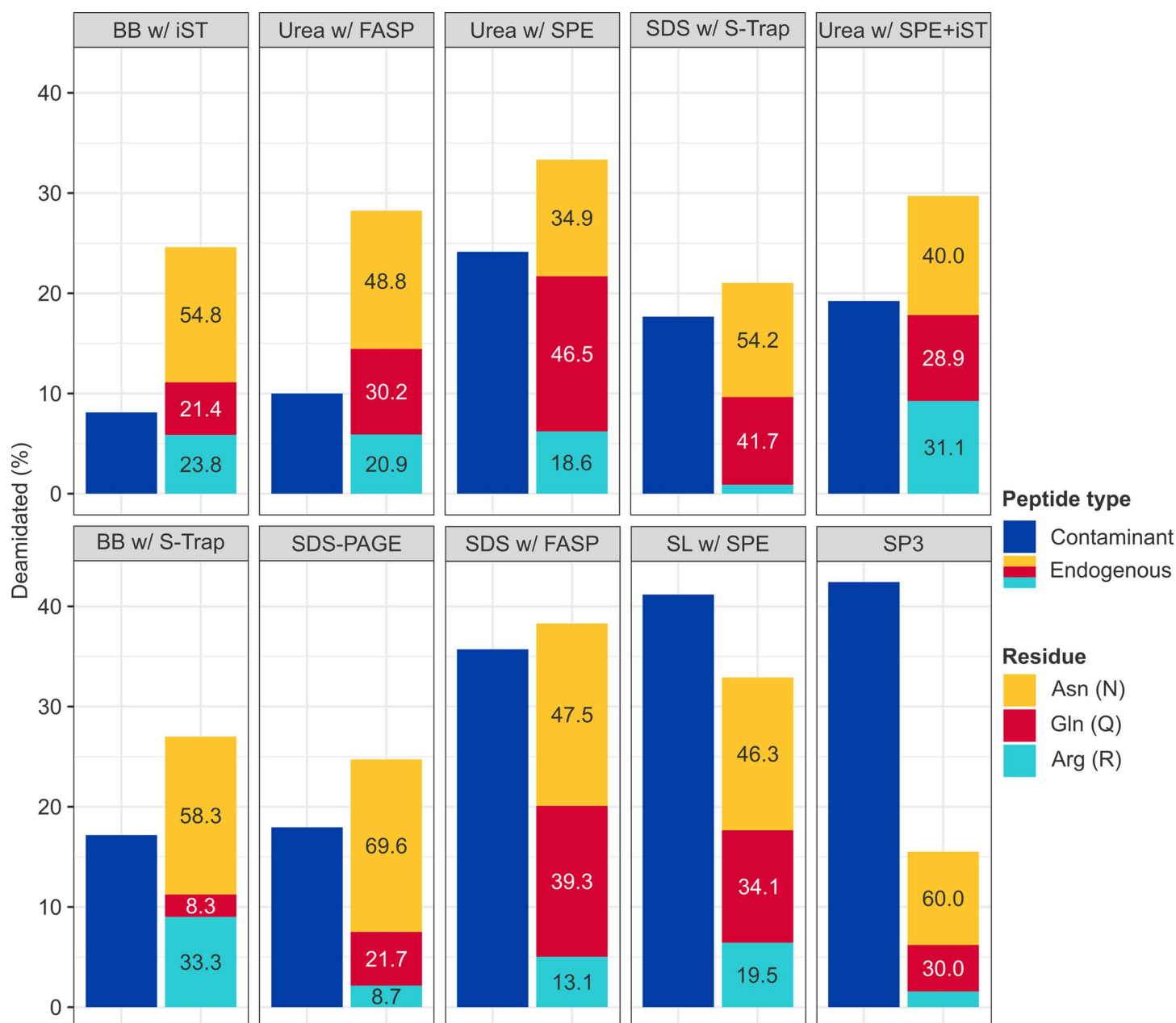


Fig 5. Nature and extent of deamidation. Percentage of deamidated contaminant (*left*) and endogenous (*right*) peptides at 1% FDR and q -value $\leq .05$ by protocol (from left to right, highest to lowest total proteins). Endogenous peptides are coloured by the affected amino acid, with numbers indicating the percentage each type of deamidation represents of the total.

<https://doi.org/10.1371/journal.pone.0324246.g005>

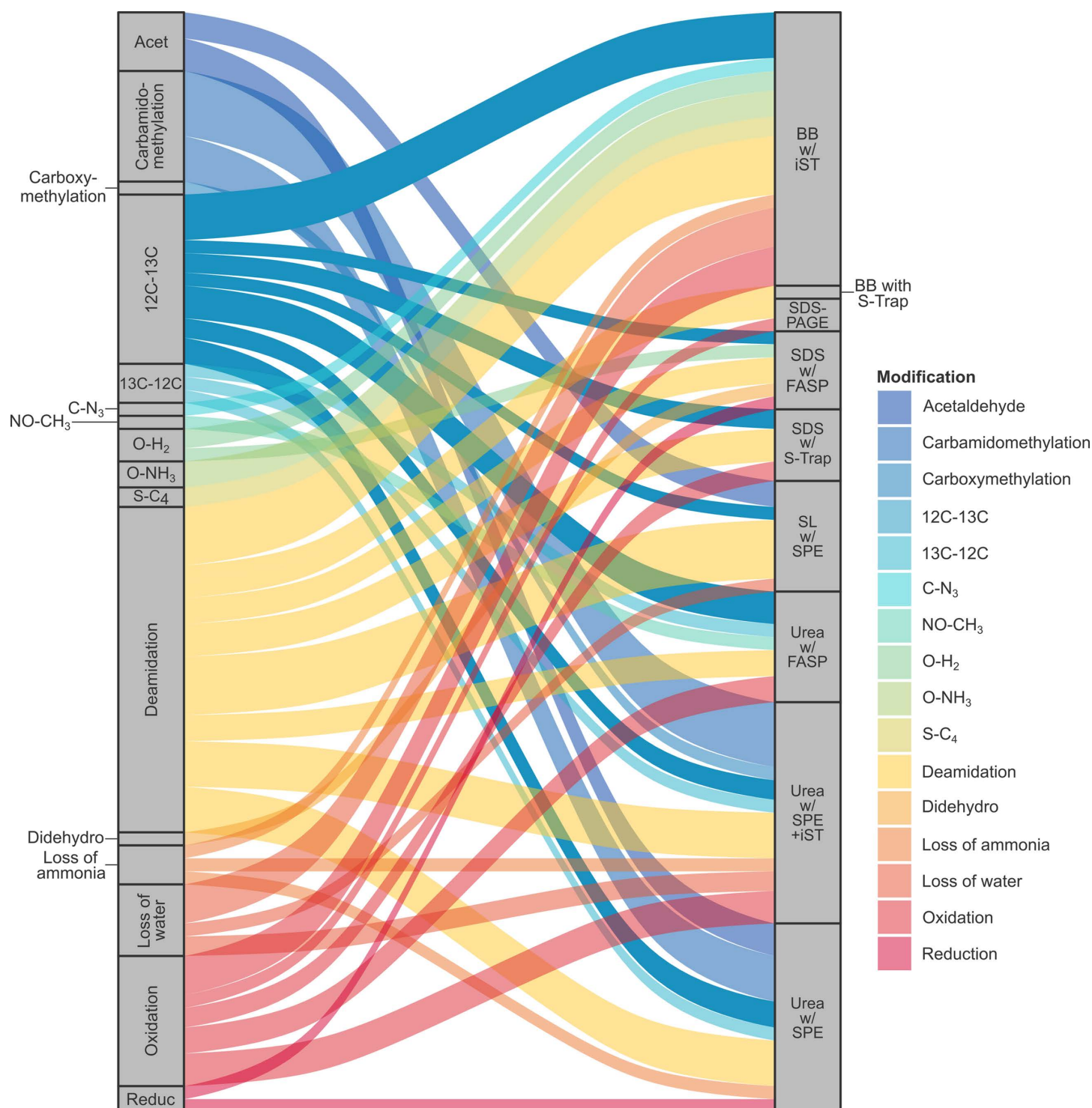


Fig 6. Discovery of covalent modifications. Alluvial diagram of modifications (*left*) to ≥ 2 unique peptides per protocol (*right*), identified by dependent peptide search. SP3 is not shown, since no identifications met the ≥ 2 unique peptides threshold.

<https://doi.org/10.1371/journal.pone.0324246.g006>

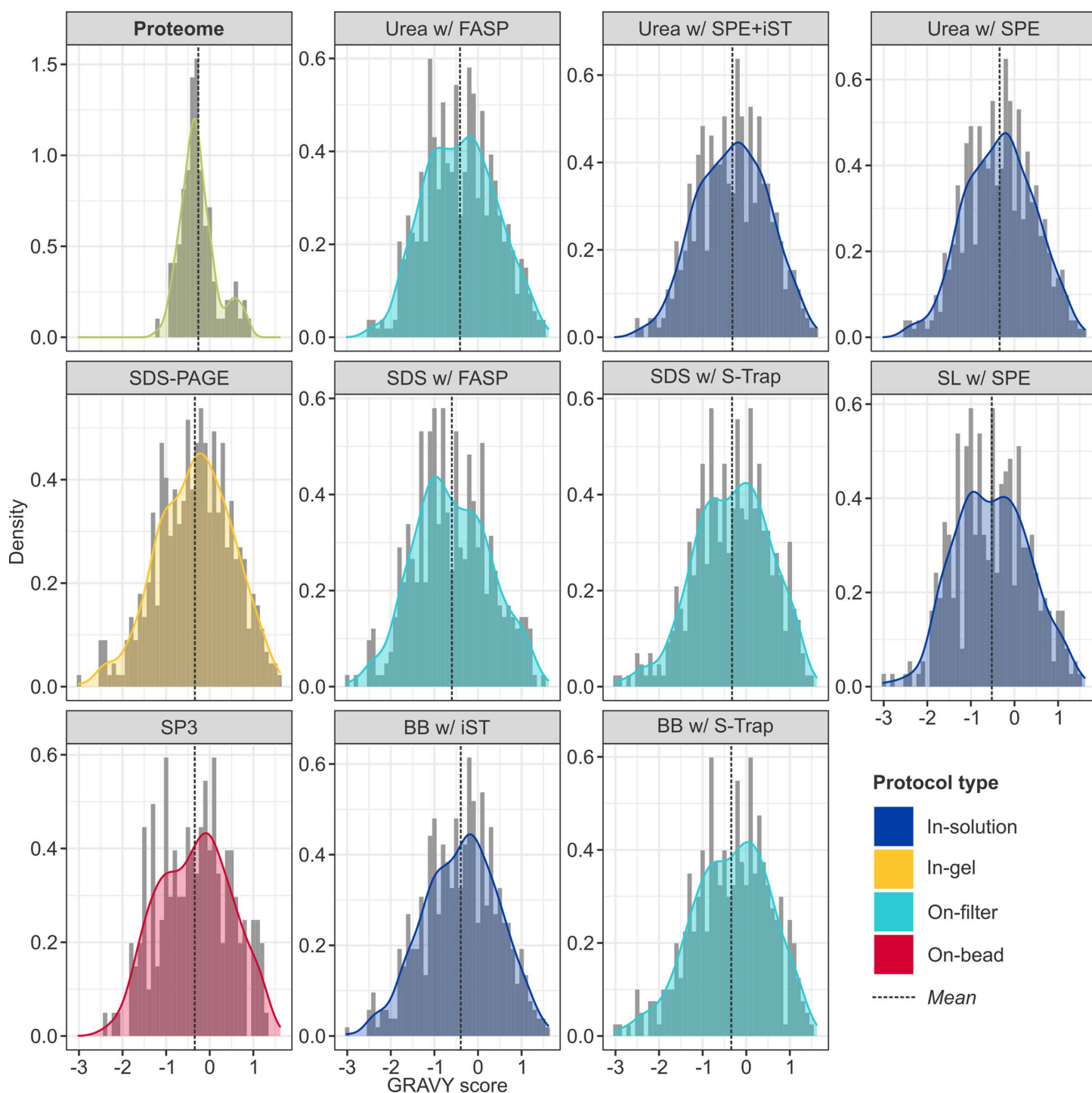


Fig 7. Distribution of GRAVY scores by protocol. Distribution of GRAVY scores calculated at the protein-level for proteins retrieved in all protocols (the “common proteome”, green), and at the peptide-level for proteins recovered by each protocol, with smoothed density curves. Given the variability in sample sizes, data has been normalised to reflect relative frequency rather than absolute counts (note the differing y-axes) and facilitate visual pairwise comparison of extraction buffers (broadly, horizontal rows; e.g., urea, top) and clean-up strategies (broadly, vertical columns; e.g., SPE, right). Dashed lines reflect the mean GRAVY score: a higher score indicates more hydrophobic peptides, and a lower score more hydrophilic peptides.

<https://doi.org/10.1371/journal.pone.0324246.g007>

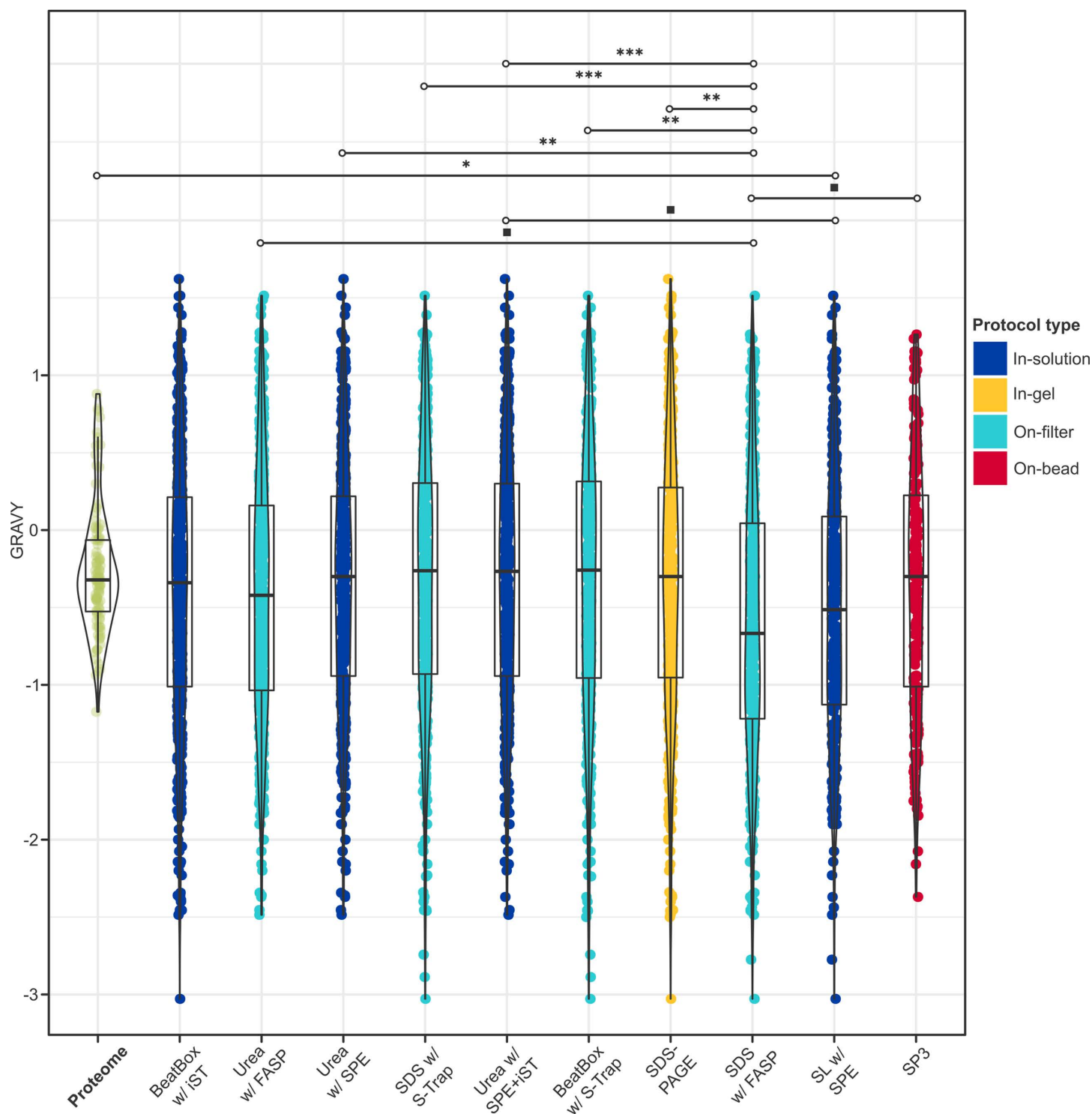


Fig 8. Hydropathic profiles retrieved with each protocol. Results of a linear mixed effects model with Satterthwaite t -tests ($n_{\text{obs}} = 4,383$) and *post-hoc* pairwise comparisons ($n = 55$) with Bonferroni correction at 95% CI with respect to GRAVY score, calculated at the protein-level for proteins retrieved in all protocols (the “common proteome”, green), and at the peptide-level for proteins recovered by each protocol (from left to right, highest to lowest total proteins). Only statistically significant comparisons are shown (for all other comparisons, see Table S6 in [S1 File](#)), and indicated by the following significance codes: *** = p -value < .0001; ** = p -value \leq .001; * = p -value \leq .01; ■ = p -value \leq .05.

<https://doi.org/10.1371/journal.pone.0324246.g008>

across the experiment. While only 13.0% of the total brain proteome is hydrophobic, an average of 39.6% [0.06] of recovered peptides were hydrophobic, suggesting a preferential preservation of hydrophobic domains. In contrast, peptides corresponding to proteins identified by all protocols exhibited lower-than-expected mean hydropathy values, indicating that more hydrophilic domains were preferentially retrieved (Table S5 in [S1 File](#)). However, only the low-performance protocol, SDS with FASP, showed a statistically significant deviation (-0.600 [0.863], $p = 8.1 \times 10^{-3}$; Table S6 in [S1 File](#)), and 90% (9/10) of significant pairwise comparisons involved SDS with FASP, suggesting it was an outlier in this regard. Meanwhile, urea-based in-solution methods exhibited the most positive estimated effects (E) on GRAVY scores, indicating that they extracted relatively more hydrophobic peptides than other reagents or clean-up strategies (e.g., urea with SPE + iST vs. SL: $E = 0.196$, $p = 3.4 \times 10^{-2}$).

Analogous statistical testing of the average pI revealed that, like hydropathy, pI of recovered peptides was lower than expected for the common pool ([Figs 9 and 10](#)), suggesting that more acidic peptides were preferentially retrieved (Table S7 in [S1 File](#)). This relationship was statistically significant for all in-solution protocols (e.g., urea with SPE + iST: $p = 3.0 \times 10^{-4}$), and for on-filter methods with FASP (e.g., urea with FASP: $p = 8.1 \times 10^{-3}$; Tables S8 in [S1 File](#)). Although the common proteome exhibited the expected bimodal distribution in pI for a eukaryotic proteome [69], a shift toward lower values was apparent: while the major peak associated with cytoplasmic proteins occurred as expected at ~ 5.0 , the major peak corresponding to integral membrane proteins was shifted from ~ 8.5 to ~ 7.5 [70]. By contrast, all protocols evidenced a distinct trimodal distribution, with major peaks at ~ 4.5 and 6.5 , and a minor peak ~ 9.0 , in descending order of density (with the exception of BeatBox with S-Trap, whose peak at ~ 6.5 marginally exceeded that at ~ 4.5 ; 0.266 vs. 0.248). Corresponding to nuclear proteins [71], this minor peak evidenced the greatest density in urea-based protocols (e.g., urea with SPE: 0.080), corroborating the observation of GO functional enrichment for the mitochondrial compartment with this lysis buffer.

Common laboratory contaminants identified with associated risk levels are illustrated in [Fig 11a](#). Global risk level was very low (1) in nine of ten protocols, and low (2) in BeatBox with S-Trap. Kit-based protocols (i.e., those involving BeatBox and/or iST technology) evidenced a higher abundance of more diverse contaminants, particularly anionic surfactants such as Triton, Tween and IGEPAL-630/NP40. Polyethylene glycol (PEG) was the most abundant contaminant detected at high or very high risk levels (≥ 4) in all samples, posing the greatest risk in the SDS with FASP and urea with SPE + iST protocols. However, no patterns could be discerned in terms of the abundance of PEG contamination and the use of particular extraction buffers or clean-up methods ([Fig 11b](#)); nor with the frequency of heating, vortexing, shaking, centrifugation, sonication, or homogenisation steps.

Limitations

Our study was conducted on ten equal-sized subsamples from a single brain biopsy. Given the destructive nature of the analysis and irreplaceable nature of the material, we elected not to perform replicate extractions: this decision is in line with ethical considerations regarding the treatment of human remains in archaeology, which stipulate a reduced program of sampling for proof-of-concept studies where the feasibility and/or applicability of a method may be questionable [72,73]. Our approach minimised biological variability and provided the internal consistency necessary to ensure that we could effectively compare the efficacy and efficiency of the tested protocols; however, we were unable to assess technical variability. Building on this proof-of-concept study, further work should incorporate both biological and technical replicates to test the generalisability of our findings to a broader population and context.

Discussion

Only a fraction of the *in vivo* proteome is expected to persist in the archaeological record, and those proteins that survive diagenesis to be extracted, analysed and identified are generally mineralised, highly abundant, and/or possess particular structural features [74]. In this work, we address the absence of a sensitive and specific protocol in palaeoproteomics for non-mineralised tissues. We assessed the efficacy, efficiency and suitability of extraction buffers and clean-up strategies

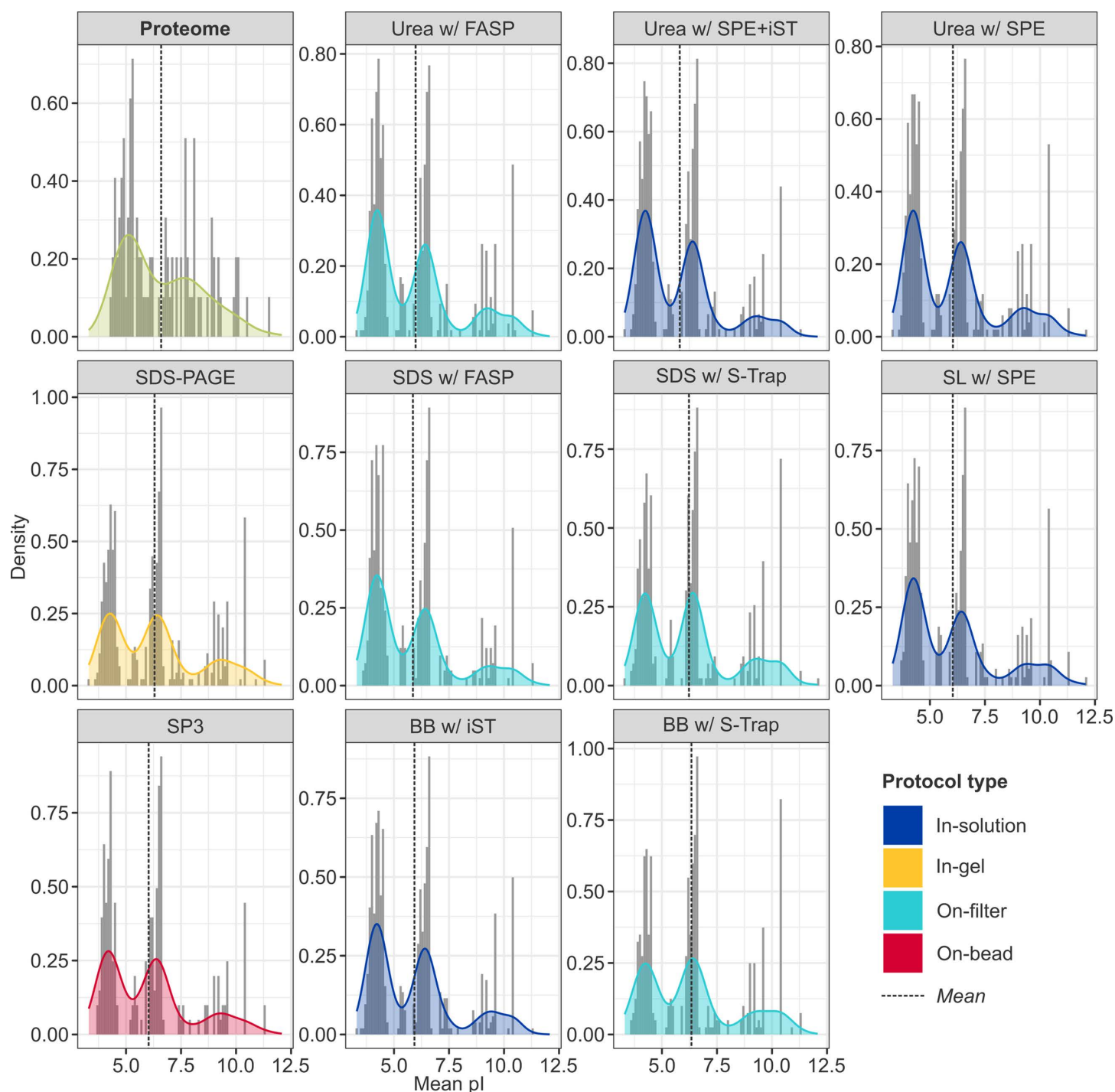


Fig 9. Distribution of isoelectric point by protocol. Distribution of mean pI calculated at the protein-level for proteins retrieved in all protocols (the "common proteome", green), and at the peptide-level for proteins recovered by each protocol, with smoothed density curves. Given the variability in sample sizes, data has been normalised to reflect relative frequency rather than absolute counts (note the differing y-axes) and facilitate visual pairwise comparison of extraction buffers (broadly, horizontal rows; e.g., urea, top) and clean-up strategies (broadly, vertical columns; e.g., SPE, right). Dashed lines reflect the mean pI: a higher value indicates more basic peptides, and a lower value more acidic peptides.

<https://doi.org/10.1371/journal.pone.0324246.g009>

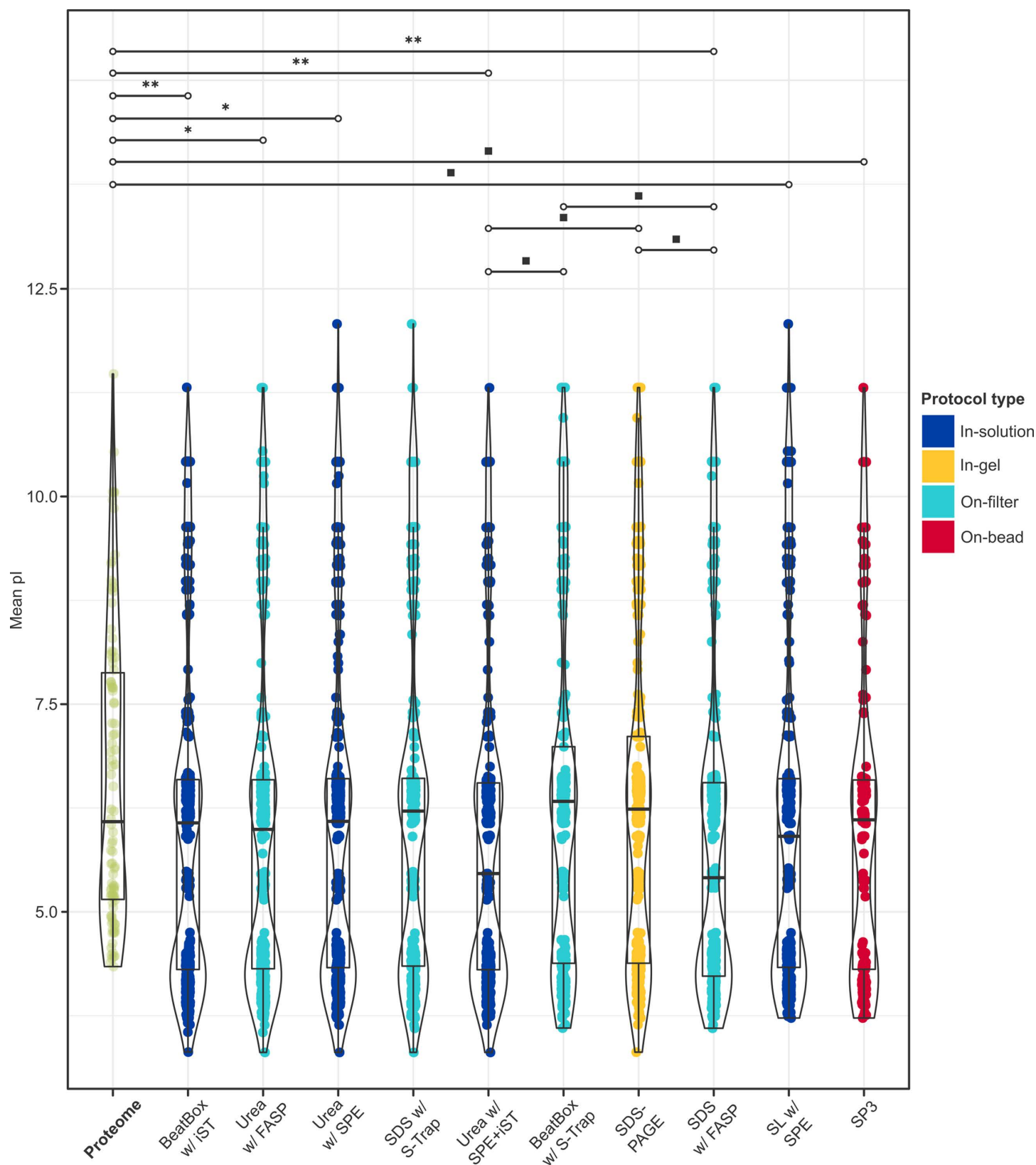


Fig 10. Isoelectric profiles retrieved with each protocol. Results of a linear mixed effects model with Satterthwaite t -tests ($n_{\text{obs}} = 4,383$) and *post-hoc* pairwise comparisons ($n = 55$) with Bonferroni correction at 95% CI with respect to mean pI, calculated at the protein-level for proteins retrieved in

all protocols (the “common proteome”, *green*), and at the peptide-level for proteins recovered by each protocol (from left to right, highest to lowest total proteins). Only statistically significant comparisons are shown (for all other comparisons, see Table S8 in [S1 File](#)), and indicated by the following significance codes: *** = p -value < .0001; ** = p -value ≤ .001; * = p -value ≤ .01; ■ = p -value ≤ .05.

<https://doi.org/10.1371/journal.pone.0324246.g010>

with varied chemistries using archaeological human brain tissues, and describe an optimised LC-FAIMS-MS/MS workflow yielding the largest, most diverse palaeoproteome recovered from any archaeological or fossil material to-date. Further, although 50 mg starting material was employed for extraction, just 5% was injected into the LC-MS platform; demonstrating that even smaller amounts of tissue could be sufficient for deep palaeoproteomic analysis.

We found that urea-based methods (both in-solution and on-filter) proved the most effective and efficient overall, evidencing low contamination and high numbers of protein identifications with relatively high intensities and sequence coverage. In addition, our results suggested that urea-based methods facilitated the recovery of a broader range of peptides, thereby providing greater access to the full diversity of the palaeoproteome: specifically, a greater proportion of non-tryptic and hydrophobic peptides corresponding to membrane proteins, which are more abundant in soft than hard tissues [20]. A chaotropic agent, urea denatures proteins by disrupting non-covalent interactions (such as hydrogen bonds and van der Waals forces), decreasing the net hydrophobic effect of hydrophobic regions and forcing unfolded proteins into solution. This same action destabilises lipid bilayers (compromising membrane integrity and liberating intracellular components) and protein aggregates, which accumulate in aged organs like the brain [75], heart [76], and kidneys [77], and which have previously been identified in preserved soft tissues [21,78]. Indeed, with the exception of SDS-PAGE, urea-based methods were the only to enrich primarily for recovery of the intracellular mitochondrial ATP synthase complex, a critical component of energy production in soft tissues [79], which possess metabolic rates many times that of mineralised tissues [80]. While further work will be required to confirm its suitability for a wider range of samples (see Limitations), urea lysis is likely to be more effective for archaeological soft tissues than those used conventionally in palaeoproteomics for mineralised tissues, such as weak acids or chelating agents [13].

Across the experiment, we identified preferential preservation of hydrophobic proteins, but retrieval of relatively acidic (low-pI) peptides. Protein extraction is strongly influenced by hydrophobic and isoelectric behaviour, both of which govern solubility [81]; it is likely that molecular decomposition follows similar principles. During soft tissue decay, the breakdown of proteins and other biomolecules generates acidic byproducts – such as lactic acid, fatty acids, and negatively charged amino acids like aspartate and glutamate – which lower the pH of the surrounding tissue [82]. Under these conditions, hydrophobic proteins, being largely insoluble, are more likely to persist than their hydrophilic counterparts, as their tendency to precipitate and aggregate dominates in all but the most alkaline environments [83]. Consequently, hydrophobic components are notoriously challenging to solubilise and extract [84,85]. While this alone may explain the observed bias toward more soluble, low-pI peptides in our dataset, our pI analyses suggest an additional influence. Proteome-wide distribution studies of pI [69,70,86] reveal an “acidic shoulder” (~ pI 4.75) adjacent to the major acidic peak, which is attributed to an over-representation of low-pI cytoplasmic and extracellular proteins [69]. Cytoskeletal proteins, in particular, are extremely acid-biased, with a clear low-pI shift and trimodal distribution, in contrast to the bimodal distribution observed in cytoplasmic proteins from the same subcellular compartment [69]. This difference is thought to arise from the polymeric nature of cytoskeletal proteins, which precludes true solubility [69] (although progressive changes in solubility have been noted *postmortem* [87,88]). In our dataset, which features an overrepresentation of cytoskeletal proteins, the major peak aligns with this acidic shoulder, suggesting that these recalcitrant fibres contribute to both the preservation and retrieval biases observed.

In step with current best practices in palaeoproteomics [4,6], our study employed stringent search parameters – such as requiring >2 unique peptides for an identification – to ensure the authenticity of ancient proteins. This stipulation works well for archaeological tissues like bone, skin and hair, where contamination with modern collagens and keratins (which it

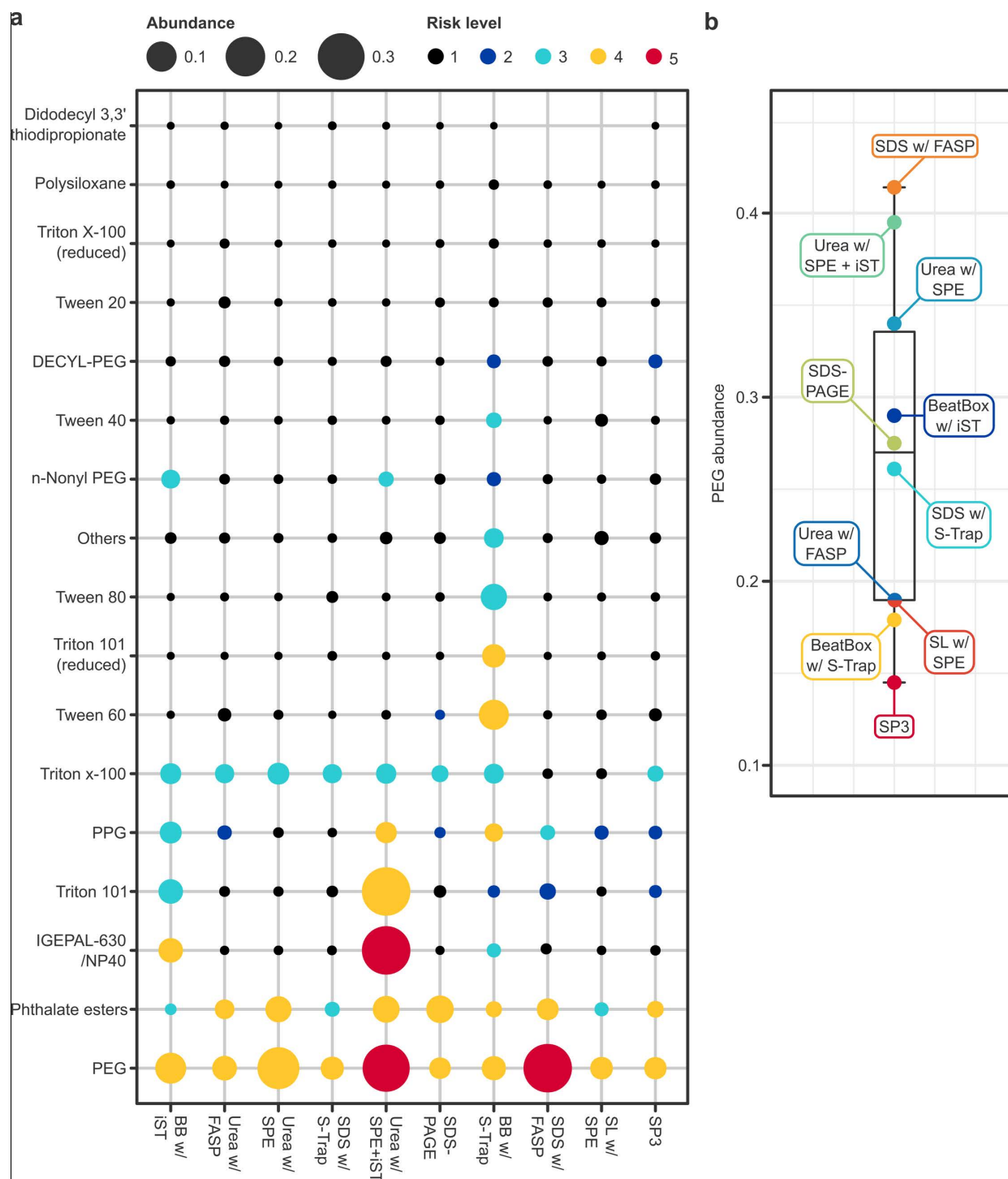


Fig 11. Abundance of common laboratory contaminants. (a) Dotplot of contaminants (y-axis) identified by protocol (x-axis; from left to right, highest to lowest total proteins). Point size indicates total abundance and point colour indicates risk level, which are designated *very high* (5), *high* (4), *medium* (3), *low* (2) and *very low* (1). (b) Boxplot of PEG abundance by protocol.

<https://doi.org/10.1371/journal.pone.0324246.g011>

is essential to minimise but difficult to eliminate [6]) risks misclassifying exogenous peptides as ancient [89]. However, it is likely that for archaeological internal organs, where the risk of a false-positive of this nature is much lower, this stringency leads to more false negatives: that is, discarding genuine identifications that may be better authenticated by other means (e.g., use of post-processing classifiers).

Beyond the risk of false negatives, a greater concern is that genuine identifications may be missed altogether if (semi-) tryptic peptides from digested proteins compete with highly degraded, non-tryptic peptides or contaminants for fragmentation and ionisation. In this respect, the sensitivity of the MS system employed is crucial: the Orbitrap™ Ascend features a number of improvements to the Tribrid™ architecture (including a second ion-routing multipole, which permits parallel ion injection) that together increase the scan speed and sensitivity of the instrument, producing a greater number of higher-quality tandem mass spectra [90]. However, this enhanced sensitivity could also elevate chemical noise, particularly if smaller and/or degraded molecules – such as lipids, metabolites, or other non-peptide contaminants – are transmitted more efficiently into the mass spectrometer.

To address these challenges, we found that LC-FAIMS-MS/MS improves unique protein identification by up to 40%, and represents an effective means of “cleaning” dirty archaeological samples, regardless of the extraction buffer or clean-up strategy used upstream. By filtering out singly charged ions associated with chemical noise, FAIMS both improved the signal-to-noise ratio in complex samples and enabled the detection of low-abundant analytes that would otherwise have been masked, overlapping in mass-to-charge with interfering ions. Moreover, we demonstrate that FAIMS is a highly tuneable technique: given compensation voltages can be optimised to select for ions with specific mobility characteristics, it is likely to be broadly useful for improving protein recovery not only from a range of ancient tissues, but even from a range of depositional environments. Given >90% of fragment ion spectra from ancient samples fail to be assigned [91], further work in optimising collision energies – or indeed, precursory separation by LC – is expected to help tackle this problem for palaeoproteomics, particularly if coupled with FAIMS.

In sum, this study aimed to identify a suitable method for extracting ancient proteins from soft tissues, maximising the amount of biological information recoverable, and laying the groundwork for future research in non-mineralised bioarchaeological remains. As this method generates increasingly large datasets, optimising their management and analysis presents a new and exciting challenge for palaeoproteomics.

Supporting information

S1 File. Supporting information, including methodological detail, and additional figures and tables.
(DOCX)

S1 Dataset. Supporting data and analyses.
(XLSX)

Acknowledgments

We are grateful to Sharon Clough and Trish Biers for access to samples. We thank members of the Discovery Proteomics Facility, and the Kessler and Fischer groups at the Target Discovery Institute for useful discussions and support.

Author contributions

Conceptualization: Alexandra Morton-Hayward, Sarah Flannery, Roman Fischer.

Data curation: Alexandra Morton-Hayward, Sarah Flannery, Iolanda Vendrell.

Formal analysis: Alexandra Morton-Hayward, Iolanda Vendrell, Roman Fischer.

Funding acquisition: Alexandra Morton-Hayward, Roman Fischer.

Investigation: Alexandra Morton-Hayward, Sarah Flannery, Roman Fischer.

Methodology: Alexandra Morton-Hayward, Sarah Flannery, Iolanda Vendrell, Roman Fischer.

Project administration: Alexandra Morton-Hayward.

Resources: Alexandra Morton-Hayward, Iolanda Vendrell, Roman Fischer.

Supervision: Roman Fischer.

Validation: Sarah Flannery, Iolanda Vendrell, Roman Fischer.

Visualization: Alexandra Morton-Hayward.

Writing – original draft: Alexandra Morton-Hayward.

Writing – review & editing: Alexandra Morton-Hayward, Sarah Flannery, Iolanda Vendrell, Roman Fischer.

References

- Demarchi B, Hall S, Roncal-Herrero T, Freeman CL, Woolley J, Crisp MK, et al. Protein sequences bound to mineral surfaces persist into deep time. *Elife*. 2016;5:e17092. <https://doi.org/10.7554/eLife.17092> PMID: 27668515
- Dutta S, Kumar S, Singh H, Khan M, Barai A, Tewari A, et al. Chemical evidence of preserved collagen in 54-million-year-old fish vertebrae. *Palaeontology*. 2020;63:195–202.
- Schmidt-Schultz TH, Reich M, Schultz M. Exceptionally preserved extracellular bone matrix proteins from the late Neogene proboscidean *Anancus* (Mammalia: Proboscidea). *PalZ*. 2021;95(4):757–65. <https://doi.org/10.1007/s12542-021-00566-7>
- Warinner C, KorzowRichter K, Collins M. Paleoproteomics. *Chem Rev*. 2022;122:13401–46.
- Welker F. Palaeoproteomics for human evolution studies. *Quaternary Science Reviews*. 2018;190:137–47. <https://doi.org/10.1016/j.quascirev.2018.04.033>
- Hendy J, Welker F, Demarchi B, Speller C, Warinner C, Collins MJ. A guide to ancient protein studies. *Nat Ecol Evol*. 2018;2(5):791–9. <https://doi.org/10.1038/s41559-018-0510-x> PMID: 29581591
- Weston LA, Bauer KM, Hummon AB. Comparison of bottom-up proteomic approaches for LC-MS analysis of complex proteomes. *Anal Methods*. 2013;5(18):10.1039/C3AY40853A. <https://doi.org/10.1039/C3AY40853A> PMID: 24288579
- León IR, Schwämmle V, Jensen ON, Sprenger RR. Quantitative assessment of in-solution digestion efficiency identifies optimal protocols for unbiased protein analysis. *Mol Cell Proteomics*. 2013;12(10):2992–3005. <https://doi.org/10.1074/mcp.M112.025585> PMID: 23792921
- Brough Z, Zhao Z, Duong van Hoa F. From bottom-up to cell surface proteomics: detergents or no detergents, that is the question. *Biochem Soc Trans*. 2024;52(3):1253–63. <https://doi.org/10.1042/BST20231020> PMID: 38666604
- Duong V-A, Lee H. Bottom-Up Proteomics: Advancements in Sample Preparation. *Int J Mol Sci*. 2023;24(6):5350. <https://doi.org/10.3390/ijms24065350> PMID: 36982423
- Umamaheswaran R, Dutta S. Preservation of proteins in the geosphere. *Nat Ecol Evol*. 2024;8(5):858–65. <https://doi.org/10.1038/s41559-024-02366-z> PMID: 38472431
- Lanigan LT, Mackie M, Feine S, Hublin J-J, Schmitz RW, Wilcke A, et al. Multi-protease analysis of Pleistocene bone proteomes. *J Proteomics*. 2020;228:103889. <https://doi.org/10.1016/j.jprot.2020.103889> PMID: 32652221
- Mylopotamitaki D, Harking FS, Taurozzi AJ, Fagnäs Z, Godinho RM, Smith GM, et al. Comparing extraction method efficiency for high-throughput palaeoproteomic bone species identification. *Sci Rep*. 2023;13(1):18345. <https://doi.org/10.1038/s41598-023-44885-y> PMID: 37884544
- Wilkin S, Lanigan LT, Montes N, Sharma M, Avanzi C, Sejdiu D, et al. Sequential trypsin and ProAlanase digestions unearth immunological protein biomarkers shrouded by skeletal collagen. *iScience*. 2024;27(5):109663. <https://doi.org/10.1016/j.isci.2024.109663> PMID: 38655200
- Brandt LØ, Taurozzi AJ, Mackie M, Sinding M-HS, Vieira FG, Schmidt AL, et al. Palaeoproteomics identifies beaver fur in Danish high-status Viking Age burials - direct evidence of fur trade. *PLoS One*. 2022;17(7):e0270040. <https://doi.org/10.1371/journal.pone.0270040> PMID: 35895633
- Demarchi B, Boano R, Ceron A, Bello FD, Favero-Longo SE, Fiddymment S, et al. Never boring: Non-invasive palaeoproteomics of mummified human skin. *J Archaeol Sci*. 2020;119:105145.
- Brandt LØ, Mackie M, Daragan M, Collins MJ, Gleba M. Human and animal skin identified by palaeoproteomics in Scythian leather objects from Ukraine. *PLoS One*. 2023;18(12):e0294129. <https://doi.org/10.1371/journal.pone.0294129> PMID: 38091270
- Ebsen JA, Haase K, Larsen R, Sommer DVP, Brandt LØ. Identifying archaeological leather – discussing the potential of grain pattern analysis and zooarchaeology by mass spectrometry (ZooMS) through a case study involving medieval shoe parts from Denmark. *J Cult Herit*. 2019;39:21–31.
- Fretwurst T, Tritschler I, Rothweiler R, Nahles S, Altmann B, Schilling O, et al. Proteomic profiling of human bone from different anatomical sites - A pilot study. *Proteomics Clin Appl*. 2022;16(5):e2100049. <https://doi.org/10.1002/prca.202100049> PMID: 35462455

20. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. *Science*. 2015;347(6220):1260419. <https://doi.org/10.1126/science.1260419> PMID: 25613900
21. Petzold A, Lu C-H, Groves M, Gobom J, Zetterberg H, Shaw G, et al. Protein aggregate formation permits millennium-old brain preservation. *J R Soc Interface*. 2020;17(162):20190775. <https://doi.org/10.1098/rsif.2019.0775> PMID: 31910770
22. Öhrström LM, Seiler R, Böni T, Aali A, Stöllner T, Rühli FJ. Radiological findings in an ancient Iranian salt mummy (Chehrābād ca. 410–350 BC). *Skeletal Radiol*. 2015;44(6):811–21. <https://doi.org/10.1007/s00256-015-2103-y> PMID: 25662179
23. Chen D, Huang J-F, Chen J-M, You Z-Q, Wang H, Wang X-S, et al. Autopsy and forensic study on a rare human corpse preserved over two thousand years: The Mawangdui ancient cadaver. *Biopreserv Biobank*. 2019;17:105–12.
24. Bianucci R, Loynes RD, Sutherland ML, Lallo R, Kay GL, Froesch P, et al. Forensic Analysis Reveals Acute Decompensation of Chronic Heart Failure in a 3500-Year-Old Egyptian Dignitary. *J Forensic Sci*. 2016;61(5):1378–81. <https://doi.org/10.1111/1556-4029.13138> PMID: 27362779
25. Panzer S, Wittig H, Zesch S, Rosendahl W, Blache S, Müller-Gerbl M, et al. Evidence of neurofibromatosis type 1 in a multi-morbid Inca child mummy: A paleoradiological investigation using computed tomography. *PLoS One*. 2017;12(4):e0175000. <https://doi.org/10.1371/journal.pone.0175000> PMID: 28403237
26. Panzer S, Zink AR, Piombino-Mascoli D. Scenes from the past: radiologic evidence of anthropogenic mummification in the Capuchin Catacombs of Palermo, Sicily. *Radiographics*. 2010;30(4):1123–32. <https://doi.org/10.1148/rq.304095174> PMID: 20631372
27. Väre T, Junno J-A, Niinimäki J, Niskanen M, Niinimäki S, Núñez M, et al. Computed tomography of mummified human remains in old Finnish churches, a case study: the mummified remains of a 17th-century vicar revisited. *Postmediev Archaeol*. 2016;50:368–79.
28. Kim MJ, Park SS, Bok GD, Choi YH, Lee IS, Shin KJ, et al. Medieval mummy from Yangju. *ArcheolEthnol Anthropol Eurasia*. 2006;28(1):122–9. <https://doi.org/10.1134/s156301100604013x>
29. Lee S, Lee E, In K, Yoo S, Woo O, Rhee Y, et al. Paleopathological analysis on bronchial washing fluid from a mummy: The “General Hak-Bong” mummy, Korea. *Basic and Applied Pathology*. 2008;1(4):164–7. <https://doi.org/10.1111/j.1755-9294.2008.00039.x>
30. Panzer S, Peschel O, Haas-Gebhard B, Bachmeier BE, Pusch CM, Nerlich AG. Reconstructing the life of an unknown (ca. 500 years-old South American Inca) mummy—multidisciplinary study of a Peruvian Inca mummy suggests severe Chagas disease and ritual homicide. *PLoS One*. 2014;9(2):e89528. <https://doi.org/10.1371/journal.pone.0089528> PMID: 24586848
31. Piombino-Mascoli D, Jankauskas R, Tamošiūnas A, Valančius R, Thompson RC, Panzer S. Atherosclerosis in mummified human remains from Vilnius, Lithuania (18th–19th centuries AD): a computed tomographic investigation: Atherosclerosis in Lithuanian Mummies. *Am J Hum Biol*. 2014;26(5):676–81. <https://doi.org/10.1002/ajhb.22578> PMID: 24948424
32. Morton-Hayward AL, Anderson RP, Saupe EE, Larson G, Cosmidis JG. Human brains preserve in diverse environments for at least 12 000 years. *Proc Biol Sci*. 2024;291(2019):20232606. <https://doi.org/10.1098/rspb.2023.2606> PMID: 38503334
33. Maixner F, Overath T, Linke D, Janko M, Guerriero G, van den Berg BHJ, et al. Paleoproteomic study of the Iceman’s brain tissue. *Cell Mol Life Sci*. 2013;70(19):3709–22. <https://doi.org/10.1007/s00018-013-1360-y> PMID: 23739949
34. Vit O, Petrak J. Integral membrane proteins in proteomics. How to break open the black box? *J Proteomics*. 2017;153: 8–20.
35. Vuckovic D, Dagley LF, Purcell AW, Emili A. Membrane proteomics by high performance liquid chromatography-tandem mass spectrometry: Analytical approaches and challenges. *Proteomics*. 2013;13(3–4):404–23. <https://doi.org/10.1002/pmic.201200340> PMID: 23125154
36. Garavito RM, Ferguson-Miller S. Detergents as tools in membrane biochemistry. *J Biol Chem*. 2001;276(35):32403–6. <https://doi.org/10.1074/jbc.R100031200> PMID: 11432878
37. Zhang X. Detergents: Friends not foes for high-performance membrane proteomics toward precision medicine. *Proteomics*. 2017;17(3–4):10.1002/pmic.201600209. <https://doi.org/10.1002/pmic.201600209> PMID: 27633951
38. Rey M, Mrázek H, Pompach P, Novák P, Pelosi L, Brandolin G, et al. Effective removal of nonionic detergents in protein mass spectrometry, hydrogen/deuterium exchange, and proteomics. *Anal Chem*. 2010;82(12):5107–16. <https://doi.org/10.1021/ac100171m> PMID: 20507168
39. Feist P, Hummon AB. Proteomic challenges: sample preparation techniques for microgram-quantity protein analysis from biological samples. *Int J Mol Sci*. 2015;16(2):3537–63. <https://doi.org/10.3390/ijms16023537> PMID: 25664860
40. Peng M, Taouatas N, Cappadona S, van Breukelen B, Mohammed S, Scholten A, et al. Protease bias in absolute protein quantitation. *Nat Methods*. 2012;9(6):524–5. <https://doi.org/10.1038/nmeth.2031> PMID: 22669647
41. Glatter T, Ahn E, Schmidt A. Comparison of Different Sample Preparation Protocols Reveals Lysis Buffer-Specific Extraction Biases in Gram-Negative Bacteria and Human Cells. *J Proteome Res*. 2015;14(11):4472–85. <https://doi.org/10.1021/acs.jproteome.5b00654> PMID: 26412744
42. Klont F, Bras L, Wolters JC, Ongay S, Bischoff R, Halmos GB, et al. Assessment of Sample Preparation Bias in Mass Spectrometry-Based Proteomics. *Anal Chem*. 2018;90(8):5405–13. <https://doi.org/10.1021/acs.analchem.8b00600> PMID: 29608294
43. Kulak NA, Pichler G, Paron I, Nagaraj N, Mann M. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat Methods*. 2014;11(3):319–24. <https://doi.org/10.1038/nmeth.2834> PMID: 24487582
44. Zougman A, Selby PJ, Banks RE. Suspension trapping (STrap) sample preparation method for bottom-up proteomics analysis. *Proteomics*. 2014;14(9):1006–0. <https://doi.org/10.1002/pmic.201300553> PMID: 24678027
45. Wiśniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. *Nat Methods*. 2009;6(5):359–62. <https://doi.org/10.1038/nmeth.1322> PMID: 19377485

46. Hughes CS, Moggridge S, Müller T, Sorensen PH, Morin GB, Krijgsveld J. Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat Protoc.* 2019;14(1):68–85. <https://doi.org/10.1038/s41596-018-0082-x> PMID: [30464214](#)
47. Demichev V, Messner CB, Vernardis SI, Lilley KS, Ralser M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods.* 2020;17(1):41–4. <https://doi.org/10.1038/s41592-019-0638-x> PMID: [31768060](#)
48. Yu F, Teo G, Kong A, Fröhlich K, Li G, Demichev V. Analysis of DIA proteomics data using MSFragger-DIA and FragPipe computational platform. *Nat Commun.* 2023;14:4154.
49. Tyanova S, Temu T, Cox J. The maxquant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc.* 2016;11:2301–19.
50. PreOmics. iST Sample Preparation Kit 8x. Available from: preomics.com/resources/kit
51. Hildonen S, Halvorsen TG, Reubsaet L. Why less is more when generating tryptic peptides in bottom-up proteomics. *Proteomics.* 2014;14(17–18):2031–41. <https://doi.org/10.1002/pmic.201300479> PMID: [25044798](#)
52. UniProt Consortium. UniProt: the universal protein knowledgebase in 2023. *Nucleic Acids Res.* 2023;51:D523–31.
53. Gerault M-A, Camoin L, Granjeaud S. DIAgui: a Shiny application to process the output from DIA-NN. *Bioinform Adv.* 2024;4(1):vbae001. <https://doi.org/10.1093/bioadv/vbae001> PMID: [38249340](#)
54. Pham TV, Henneman AA, Jimenez CR. iq: an R package to estimate relative protein abundances from ion quantification in DIA-MS-based proteomics. *Bioinformatics.* 2020;36(8):2611–3. <https://doi.org/10.1093/bioinformatics/btz961> PMID: [31909781](#)
55. Cox J, Hein MY, Lubner CA, Paron I, Nagaraj N, Mann M. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics.* 2014;13(9):2513–26. <https://doi.org/10.1074/mcp.M113.031591> PMID: [24942700](#)
56. Liu M, Dongre A. Proper imputation of missing values in proteomics datasets for differential expression analysis. *Brief Bioinform.* 2021;22(3):bbaa112. <https://doi.org/10.1093/bib/bbaa112> PMID: [32520347](#)
57. Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, et al. Global quantification of mammalian gene expression control. *Nature.* 2011;473(7347):337–42. <https://doi.org/10.1038/nature10098> PMID: [21593866](#)
58. Lazar C, Gatto L, Ferro M, Bruley C, Burger T. Accounting for the Multiple Natures of Missing Values in Label-Free Quantitative Proteomics Data Sets to Compare Imputation Strategies. *J Proteome Res.* 2016;15(4):1116–25. <https://doi.org/10.1021/acs.jproteome.5b00981> PMID: [26906401](#)
59. Carvalho LB, Teigas-Campos PAD, Jorge S, Protti M, Mercolini L, Dhir R, et al. Normalization methods in mass spectrometry-based analytical proteomics: A case study based on renal cell carcinoma datasets. *Talanta.* 2024;266(Pt 1):124953. <https://doi.org/10.1016/j.talanta.2023.124953> PMID: [37490822](#)
60. Kyte J, Doolittle R. A simple method for displaying the hydropathic character of a protein. *J Mol Biol.* 1982;157:105–32.
61. Kozłowski LP. IPC 2.0: prediction of isoelectric point and pKa dissociation constants. *Nucleic Acids Res.* 2021;49(W1):W285–92. <https://doi.org/10.1093/nar/gkab295> PMID: [33905510](#)
62. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw.* 2015;67:1–48.
63. Lenth RV. Emmeans: estimated marginal means, aka least-squares means. 2024. Available from: <https://rvinlenth.github.io/emmeans/>
64. Thomas PD, Ebert D, Muruganujan A, Mushayahama T, Albu L-P, Mi H. PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Sci.* 2022;31(1):8–22. <https://doi.org/10.1002/pro.4218> PMID: [34717010](#)
65. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet.* 2000;25(1):25–9. <https://doi.org/10.1038/75556> PMID: [10802651](#)
66. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics.* 2010;26(7):966–8. <https://doi.org/10.1093/bioinformatics/btq054> PMID: [20147306](#)
67. Rardin MJ. Rapid Assessment of Contaminants and Interferences in Mass Spectrometry Data Using Skyline. *J Am Soc Mass Spectrom.* 2018;29(6):1327–30. <https://doi.org/10.1007/s13361-018-1940-z> PMID: [29667163](#)
68. Gomez-Zepeda D, Michna T, Ziesmann T, Distler U, Tenzer S. HowDirty: An R package to evaluate molecular contaminants in LC-MS experiments. *Proteomics.* 2024;24(8):e2300134. <https://doi.org/10.1002/pmic.202300134> PMID: [37679057](#)
69. Kurotani A, Tokmakov AA, Sato K-I, Stefanov VE, Yamada Y, Sakurai T. Localization-specific distributions of protein pI in human proteome are governed by local pH and membrane charge. *BMC Mol Cell Biol.* 2019;20(1):36. <https://doi.org/10.1186/s12860-019-0221-4> PMID: [31429701](#)
70. Tokmakov AA, Kurotani A, Sato K-I. Protein pI and Intracellular Localization. *Front Mol Biosci.* 2021;8:775736. <https://doi.org/10.3389/fmolb.2021.775736> PMID: [34912847](#)
71. Schwartz R, Ting CS, King J. Whole proteome pI values correlate with subcellular localizations of proteins for organisms within the three domains of life. *Genome Res.* 2001;11(5):703–9. <https://doi.org/10.1101/gr-1587r> PMID: [11337469](#)
72. Mays S, Edlers J, Humphrey L, White W, Marshall P. Science and the dead a guideline for the destructive sampling of archaeological human remains for scientific analysis. 2013.
73. Squires K, Booth T, Roberts C. The ethics of sampling human skeletal remains for destructive analyses. Ethical approaches to human remains. Cham: Springer International Publishing. 2019. p. 265–97.
74. Wadsworth C, Buckley M. Proteome degradation in fossils: investigating the longevity of protein survival in ancient bone: Proteome degradation in fossils. *Rapid Commun Mass Spectrom.* 2014;28(6):605–15. <https://doi.org/10.1002/rcm.6821> PMID: [24519823](#)

75. Harel I, Chen YR, Ziv I, Singh PP, Heinzer D, Navarro Negredo P, et al. Identification of protein aggregates in the aging vertebrate brain with prion-like and phase-separation properties. *Cell Rep.* 2024;43(6):112787. <https://doi.org/10.1016/j.celrep.2023.112787> PMID: [38810650](#)
76. Ghosh R, Vinod V, Symons JD, Boudina S. Protein and Mitochondria Quality Control Mechanisms and Cardiac Aging. *Cells.* 2020;9(4):933. <https://doi.org/10.3390/cells9040933> PMID: [32290135](#)
77. Yamamoto T, Isaka Y. Pathological mechanisms of kidney disease in ageing. *Nat Rev Nephrol.* 2024;20(9):603–15. <https://doi.org/10.1038/s41581-024-00868-4> PMID: [39025993](#)
78. Wiemann J, Fabbri M, Yang T-R, Stein K, Sander PM, Norell MA, et al. Fossilization transforms vertebrate hard tissue proteins into N-heterocyclic polymers. *Nat Commun.* 2018;9(1):4741. <https://doi.org/10.1038/s41467-018-07013-3> PMID: [30413693](#)
79. McLaughlin KL, Hagen JT, Coalson HS, Nelson MAM, Kew KA, Wooten AR, et al. Novel approach to quantify mitochondrial content and intrinsic bioenergetic efficiency across organs. *Sci Rep.* 2020;10(1):17599. <https://doi.org/10.1038/s41598-020-74718-1> PMID: [33077793](#)
80. Wang Z, Ying Z, Bosy-Westphal A, Zhang J, Schautz B, Later W, et al. Specific metabolic rates of major organs and tissues across adulthood: evaluation by mechanistic model of resting energy expenditure. *Am J Clin Nutr.* 2010;92:1369–77.
81. Arakawa T, Timasheff SN. Theory of protein solubility. *Methods Enzymol.* 1985;114:49–77. [https://doi.org/10.1016/0076-6879\(85\)14005-x](https://doi.org/10.1016/0076-6879(85)14005-x) PMID: [4079776](#)
82. Powers RH. The decomposition of human remains: a biochemical perspective. *Forensic Medicine of the Lower Extremity.* Totowa, NJ: Humana Press; 2005. pp. 3–15.
83. Dill KA. Dominant forces in protein folding. *Biochemistry.* 1990;29(31):7133–55. <https://doi.org/10.1021/bi00483a001> PMID: [2207096](#)
84. Helbig AO, Heck AJR, Slijper M. Exploring the membrane proteome—challenges and analytical strategies. *J Proteomics.* 2010;73(5):868–78. <https://doi.org/10.1016/j.jprot.2010.01.005> PMID: [20096812](#)
85. Eichacker LA, Granvogl B, Mirus O, Müller BC, Miess C, Schleiff E. Hiding behind hydrophobicity. Transmembrane segments in mass spectrometry. *J Biol Chem.* 2004;279(49):50915–22. <https://doi.org/10.1074/jbc.M405875200> PMID: [15452135](#)
86. Wu S, Wan P, Li J, Li D, Zhu Y, He F. Multi-modality of pI distribution in whole proteome. *Proteomics.* 2006;6(2):449–55. <https://doi.org/10.1002/pmic.200500221> PMID: [16317776](#)
87. ElHajj Z, Cachot A, Müller T, Riederer IM, Riederer BM. Effects of postmortem delays on protein composition and oxidation. *Brain Res Bull.* 2016;121:98–104. <https://doi.org/10.1016/j.brainresbull.2016.01.005> PMID: [26791740](#)
88. Krassner MM, Kauffman J, Sowa A, Cialowicz K, Walsh S, Farrell K, et al. Postmortem changes in brain cell structure: a review. *Free Neuropathol.* 2023;4:10. <https://doi.org/10.17879/freeneuropathology-2023-4790> PMID: [37384330](#)
89. Ramsøe A, Crispin M, Mackie M, McGrath K, Fischer R, Demarchi B, et al. Assessing the degradation of ancient milk proteins through site-specific deamidation patterns. *Sci Rep.* 2021;11: 7795.
90. He Y, Shishkova E, Peters-Clarke T, Brademan D, Westphal M, Bergen D, et al. Evaluation of the orbitrap ascend tribrid mass spectrometer for shotgun proteomics. *Anal Chem.* 2023;95:10655–63.
91. Chiang Y, Welker F, Collins MJ. Spectra without stories: reporting 94% dark and unidentified ancient proteomes. *Open Res Eur.* 2024;4:71. <https://doi.org/10.12688/openreseurope.17225.1> PMID: [38903702](#)