

This is a repository copy of *The impact of cooking and burial on proteins:a characterisation of experimental foodcrusts and ceramics*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/219044/>

Version: Published Version

Article:

Evans, Miranda, Hagan, Richard, Boyd, Oliver J et al. (4 more authors) (2024) The impact of cooking and burial on proteins:a characterisation of experimental foodcrusts and ceramics. Royal Society Open Science. 240610. ISSN 2054-5703

<https://doi.org/10.1098/rsos.240610>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



Research



Check for updates

The impact of cooking and burial on proteins: a characterisation of experimental foodcrusts and ceramics

Cite this article: Evans M, Hagan R, Boyd OJ, Bondetti M, Craig OE, Collins MJ, Hendy J. 2024 The impact of cooking and burial on proteins: a characterisation of experimental foodcrusts and ceramics. *R. Soc. Open Sci.* **11**: 240610. <https://doi.org/10.1098/rsos.240610>

Received: 12 April 2024

Accepted: 1 August 2024

Subject Category:

Biochemistry, cellular and molecular biology

Subject Areas:

biochemistry

Keywords:

palaeoproteomics, ceramics, foodcrusts, experimental archaeology

Author for correspondence:

Miranda Evans

e-mail: mae52@cam.ac.uk

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.7431942>.

Miranda Evans^{1,2}, Richard Hagan², Oliver J. Boyd³, Manon Bondetti², Oliver E. Craig², Matthew J. Collins^{1,4} and Jessica Hendy²

¹McDonald Institute for Archaeological Research, University of Cambridge, Cambridge CB2 3ER, UK

²BioArCh, Department of Archaeology, University of York, York, UK

³Independent

⁴The GLOBE Institute, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

ME, 0000-0002-9284-9268; RH, 0000-0003-1270-100X; OJB, 0009-0008-8688-0067; MB, 0000-0002-9667-600X; OEC, 0000-0002-4296-8402; MJC, 0000-0003-4226-5501; JH, 0000-0002-3718-1058

Foodcrusts have received relatively little attention in the burgeoning field of proteomic analysis of ancient cuisine. We remain ignorant of how cooking and burial impact protein survival, and crucially, the extent to which the extractome reflects the composition of input ingredients. Therefore, through experimental analogues, we explore the extent of protein survival in unburied and buried foodcrusts and ceramics using 'typical' Mesolithic ingredients (red deer, Atlantic salmon and sweet chestnut). We then explore a number of physicochemical properties theorised to aid protein preservation. The results reveal that proteins were much more likely to be detected in foodcrusts than ceramics using the methodology employed, that input ingredient strongly influences protein preservation, and that degradation is not universal nor linear between proteins, indicating that multiple protein physicochemical properties are at play. While certain properties such as hydrophobicity apparently aid protein preservation, none single-handedly explain why particular proteins/peptides survive in buried foodcrusts: this complex interplay requires further investigation. The findings

demonstrate that proteins indicative of the input ingredient can be identifiable in foodcrust, but that the full proteome is unlikely to preserve. While this shows promise for the survival of proteins in archaeological foodcrust, further research is needed to accurately interpret foodcrust extractomes.

1. Introduction

Proteomics has become a valuable tool for identifying food and vessel use in archaeological samples particularly from exceptionally well-preserved remains from frozen [1,2], desiccated [3–5] or waterlogged contexts [6,7], as well as residues in metal vessels [8–10], and calcified residues such as limescale on ceramics [11,12] and dental calculus [13–16]. Due to the tissue and taxonomically specific sequence information proteins can hold, proteomics is particularly useful for the detection of ingredients and cuisine. However, exceptionally well-preserved remains are rare, rendering studies with statistically significant sample sizes and regional comparisons difficult. While protein analysis of human dental calculus can be immensely valuable for understanding consumed diets, it does not necessarily give clear insights into food preparation or links with culinary material culture.

Ceramics associated with food preparation and consumption would be an ideal target sample for proteomics as they are ubiquitous in many contexts. However, the detection of proteins from ceramics themselves has proved challenging, due to either strong binding, or degradation during burial. Proteins have been found to bind strongly to the mineral matrix of ceramic vessels, which likely results in good protein preservation yet renders their extraction challenging [17] without the use of harsh solvents [18]. Conversely, Barker *et al.* [19] concluded that protein content decayed rapidly upon burial, although they did not measure the initial protein content prior to burying their samples, and thus the rapidity of protein loss may be difficult to estimate. Food proteins have been reported from archaeological ceramics [11,20–22] and modern replicas [23,24]. However, there are potential factors aiding the detection of proteins in archaeological cases, such as the inclusion of remnant encrustations [20], or the sampling of ceramic from immediately beneath a limescale deposit [11], which may have provided protection from diagenesis. In the case of Solazzo *et al.* [21], the sherd was from relatively cold conditions in the Arctic coast of Alaska, and contained lipid-rich foods including whale and seal meat, both factors which may have improved protein preservation, although we note that food proteins were not detected in similar ceramics in a later study [25].

1.1. Biomolecular analyses of foodcrusts

Given the challenges in extracting proteins from ceramics themselves, foodcrusts may offer a good alternative target sample for proteomic analysis. Foodcrust, sometimes referred to as ‘carbonised residue’ and ‘char’ is broadly defined as ‘amorphous charred or burnt deposits adhering to the surface of containers associated with heating organic matter’ [26]. The prevalence of foodcrusts varies considerably; however, they are particularly abundant in Mesolithic and Neolithic contexts in northern Europe and Eurasia where they are sometimes found on the majority of ceramics within assemblages [27,28]. Examples of archaeological foodcrusts can be viewed in fig. 2 in [29], fig. 1 in [30] and fig. 3 in [31].

Lipid analysis of foodcrusts has considerably improved our understanding of ancient diet and particularly of marine resource utilisation. While frequently applied to ceramics themselves, lipid analysis has also been applied to foodcrusts to detect food and vessel use in assemblages across a vast geography spanning Europe and northern Asia [27,32–38], East Asia [28,39–43] and the Americas [25,44], and to select samples that do not contain aquatic resources for use in carbon dating, which are thus unhindered by the reservoir effect [26]. The formation of foodcrusts is a topic of ongoing research. They are often presumed to be formed by cooking food, although they can also result from the use of fuel for illumination [45] or the production of sealants, moisturisers, adhesives or glues [26,46,47]. A possible correlation exists between foodcrust formation and the processing of aquatic resources, or alternatively these particular lipids may preserve better in foodcrusts than ceramics [27].

Proteomic analysis has recently been applied to archaeological foodcrusts, demonstrating the viability of the technique [29,48], but also has generated questions around protein survival and biases [29,48]. Results so far appear congruent with the association between aquatic resource processing and foodcrusts. Shevchenko *et al.* [29] performed proteomic analysis on four Mesolithic–Neolithic

foodcrusts from the site of Friesack 4, Germany. Their results revealed the presence of deamidated fish vitellogenins and parvalbumins in one foodcrust and Suidae collagens detected in another. Lyu *et al.* [48] analysed 21 foodcrusts from the site of Xiawan in southeast China for both lipids and proteins. Their results revealed the presence of potential dietary proteins in five samples, including myosin from large yellow croaker and mandarin fish, and collagen from Caprinae and potentially other mammals [48]. In both of these cases, a low proportion of samples analysed produced dietary results, and the number of dietary proteins and peptides was also low. Despite this initial headway, many questions remain outstanding concerning the survival of proteins in ancient foodcrusts.

1.2. Potential preservation biases

The key question is simply the degree to which the proteins identified in ceramics and their residues reflect the original food processed in the vessel. Although proteins indeed become altered through different cooking processes, we remain ignorant of the degree to which those changes impact the detection of proteins in ceramics and foodcrust residues. Similarly, we are also unaware of the impact of burial on the survival of food proteins in these samples.

Compared with proteomics, there is a diversity of published experiments exploring how lipids derived from different ingredients respond to a range of cooking and deposition practises (e.g. [49–55]). For instance, Miller *et al.* [49] demonstrate that absorbed lipids extracted from ceramics represent a long period of use, while surface deposits represent the most recent cooking events. However, as a much younger discipline, such studies are rarer in palaeoproteomics with most experimental studies focused on understanding if proteins survive at all in ceramics, or optimising extraction protocols [18,19,24,56,57] rather than investigating the range of cooking and deposition variables which may impact them (although see [3,12]).

In this study, we characterise the impact of cooking and entrapment in foodcrusts and ceramics, followed by burial on the identification of proteins and peptides from three common Mesolithic foods: *Cervus elaphus* (red deer), *Salmo salar* (Atlantic salmon) and *Castanea sativa* (sweet chestnut), to anticipate results and consider expectations for archaeological interpretations of diet and food preparation practises derived from foodcrusts and ceramics. We aim to identify proteins that persist throughout the cooking process and become embedded in foodcrusts and ceramics, as well as those that persist through burial in soil for six months. Specifically, we examine metrics including peptide and protein count, concentration of different amino acids, peptide length, peptide hydrophobicity, peptide isoelectric point, protein thermal stability, protein secondary structure, protein disorder, protein amyloid propensity and relative solvent accessibility (RSA) of the detected peptides in order to identify characteristics of their survival, and we also compare protein with lipid content. Given that proteomics is capable of providing highly specific information concerning ancient ingredients and cuisine, understanding the impact of cooking and burial on the protein content of ceramics and foodcrusts is crucial to accurately interpreting proteomic results of ancient samples. These results provide a maximum baseline for protein recovery from ceramics and foodcrusts under similar conditions.

2. Material and methods

2.1. Sample creation

Experimental samples included deer meat, salmon meat and chestnut flour, each individually cooked in replica ceramic vessels over the same open fire to induce foodcrust formation, before being split and one half buried. The experiment was repeated in triplicate (figure 1; see also electronic supplementary material, S1 and S2). These were originally created and described previously by Bondetti *et al.* [52] to investigate the formation and diagnostic value of ω -(*o*-alkylphenyl)alkanoic acids (APAAs). After lipid extraction, samples were stored at 4°C until protein analysis was performed in summer 2020. Full details of the methodology used for sample creation, protein extraction, and machine analysis can be found in electronic supplementary material S1.

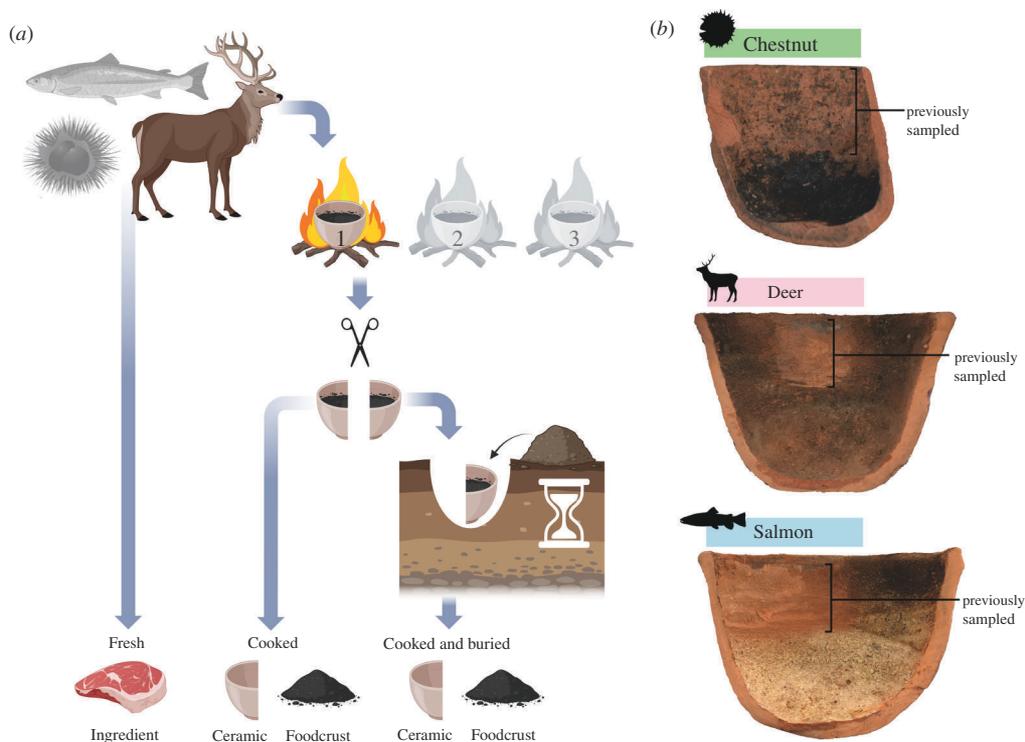


Figure 1. (a) Sample creation process (image created with BioRender.com). (b) Example of cooked foodcrust samples.

2.2. Protein extraction

All samples and machine washes were analysed following an SP3 protein extraction protocol [58,59] adapted for ancient samples [60,61] which can be found on protocols.io (<https://doi.org/10.17504/protocols.io.bfgrjjv6>), and is routinely applied to archaeological samples [61–63].

2.3. Mass spectrometry analysis

The samples were analysed by liquid chromatography tandem mass spectrometry (LC-MS/MS) on an Orbitrap Fusion at the Centre for Excellence in Mass Spectrometry at the University of York. Blank machine washes were run between each sample injection in order to examine and reduce the degree of carry-over between samples.

2.4. Data analysis

Samples and machine washes were analysed using MAXQUANT (v. 2.1.0.0). Peptides were searched allowing for tryptic cleavage, up to two missed cleavages, minimum length of seven amino acids, with both a protein and peptide target false-discovery rate of 1%. Variable modifications included oxidation (M), acetylation (protein N-term), deamidation (NQ), glutamine to pyroglutamic acid and the fixed modification of carbamidomethyl (C) was specified.

All samples were searched against a combined database which included a European red deer proteome (UP000242450), an Atlantic salmon proteome (UP000087266), a Chinese chestnut proteome (UP000737018) and 'cRAP', a database of common laboratory contaminants. *Castanea mollissima* (Chinese chestnut) was chosen as a reference database as there was no proteome for *C. sativa* (sweet chestnut) on Uniprot at the time of analysis. To investigate any potential cross-contamination during the outdoor cooking experiment, in the laboratory and due to carry-over in the LC-MS/MS, all samples were searched against all databases, to establish a baseline of cross-contamination. The 'match between runs' option was not allowed, given the varying proteomes present as match between runs has been found to falsely inflate peptide identifications [64]. Lowest common ancestor (LCA) was generated for peptides where possible using Unipept Desktop (v. 2.0.0). The data were filtered to remove potential machine carry-over and laboratory contaminants. Full details of this process can be found in electronic

supplementary material, S1, and the resulting data can be seen in electronic supplementary material, S3 and S4. It became apparent that cross-contamination occurred during field experiments, which is particularly evident in low protein samples such as ceramic extracts (electronic supplementary material, S5). To minimise the impact of cross-contamination on protein characterisation, only samples with known cross-contaminant peptide concentration $\leq 2\%$ of the total peptide count were included in the analysis of protein properties (electronic supplementary material S5).

3. Results and discussion

3.1. Impact of cooking and burial on protein and peptide detection in ceramics and foodcrust

3.1.1. Do proteins preferentially survive cooking and burial in ceramics or foodcrust?

It is immediately apparent that foodcrusts are more likely to harbour preserved proteins than ceramics using the extraction methodology utilised here. Overall, the peptide and protein count for each food was high in the fresh ingredient, reduced slightly in the cooked foodcrust and further reduced (yet still appreciable) in the buried foodcrust samples, with some variation depending on ingredient (figure 2; see also electronic supplementary material, S6) (i.e. <4 – 8 proteins in chestnut foodcrusts, <24 – 27 in salmon foodcrusts, <16 – 33 in deer foodcrusts, supported by >1 peptide spectral match). Statistical testing supports that protein count is similar in fresh ingredients and cooked foodcrust samples (no difference observed at $p < 0.05$, Wilcoxon rank sum test), while a difference in protein count was observed between unburied and buried foodcrusts (Wilcoxon rank sum test, $p < 0.001$; see electronic supplementary material, S7, for all statistical test results). In contrast to the foodcrust, protein and peptide counts were much lower in the ceramic samples, even prior to burial, and remained extremely low after burial (figure 2). This finding was supported by a Wilcoxon rank sum test which did not support a difference in protein count between unburied and buried ceramic samples at $p < 0.05$. This indicates that small but appreciable numbers of food proteins may be detectable in foodcrusts in archaeological samples of similar ingredients and conditions but in contrast, given that few positive protein identifications could be made from buried ceramics, ceramic samples should not be expected to result in positive protein results if a similar protocol is followed.

3.1.2. Why are so few proteins detected in ceramics?

Euclidean clustering revealed that the cooked ceramic samples clustered together with the buried cooked ceramic samples for all food types (figure 3a), and statistical testing did not support a difference in peptide count between unburied and buried ceramic samples at $p < 0.05$ (Wilcoxon rank sum test), indicating little or no change. This likely demonstrates that the proteins were too strongly bound to the ceramic matrix to be extracted using the protocol here employed, or alternatively, that even prior to burial, protein identification was already very low in ceramic samples, for instance if protein content had not yet impregnated the ceramic walls or if cooking had rapidly degraded any protein content present in the ceramic. This sharp reduction of both protein and peptide counts in the ceramic samples compared to the fresh ingredients is unsurprising in light of existing published research indicating similar findings [17], and that positive results from ceramics lacking encrustations and under normal preservation conditions have rarely been reported (although see [21–23]). Further work is necessary to devise optimal extraction methods for ceramic-bound protein.

3.1.3. Which proteins survive cooking and burial?

A key aim of this study was to identify the proteins that persist throughout the cooking process and become embedded in foodcrusts and ceramics, as well as those that persist after burial for six months. The most abundant proteins detected in foodcrusts and ceramics (by peptide count) can be seen in table 1. The most abundant proteins (by peptide count) detected in foodcrusts surviving cooking and burial for six months can be seen in table 2. We note that the highest peptide count proteins detected in ceramic samples include several probable contaminants such as keratins, while the highest peptide count proteins in foodcrust tend to contain more genuine ingredient matches. We also note that despite filtering described above, some cross-contamination is observable especially in chestnut samples, likely

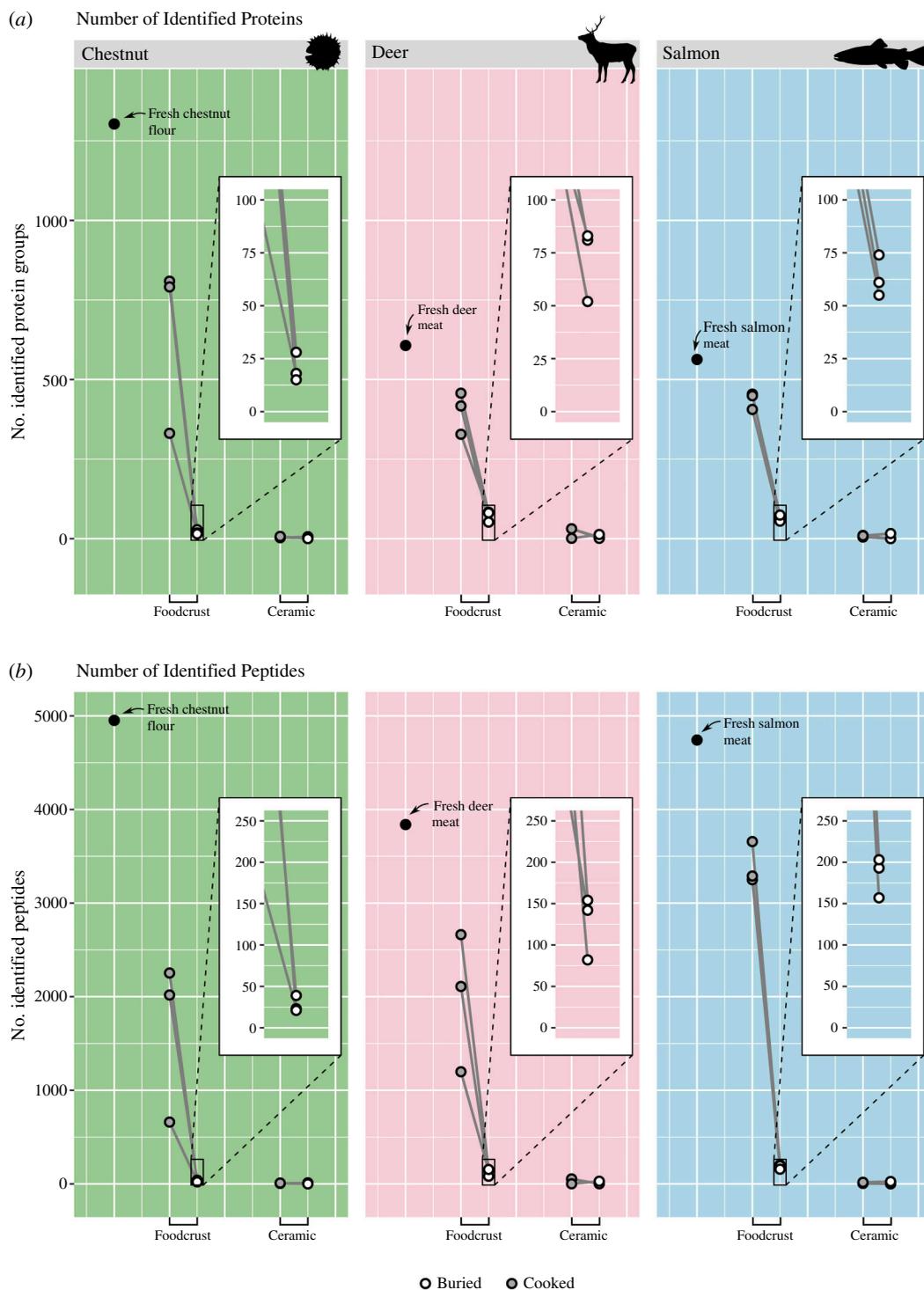


Figure 2. The number of identified proteins (a) and peptides (b) in foodcrust and ceramic samples cooked with chestnut, deer and salmon. Counts include all peptides/proteins not excluded by filtering steps described above.

derived from the field experiments, when samples of different ingredients were cooked over the same open fire.

3.1.4. Do all proteins have an equal chance of survival?

A central aim of this study was to investigate whether there is a bias towards or against the detection of certain food proteins. We sought to investigate if all proteins followed the same decay trend, i.e. highly numerous in the fresh food, then reducing in number when cooked (and entrapped in foodcrust) and

Table 1. The top five most abundant proteins (by peptide count) preserved in foodcrust and ceramic samples after cooking. Data from replicates have been merged. Probable protein source is noted, including where LCA matches to input ingredient, where protein LCA is non-specific and could potentially belong to either input ingredient or another taxonomy, and where protein LCA could not possibly match to the input ingredient and is therefore derived from contamination. This table is based on proteins with the highest peptide count regardless of taxonomic specificity. We note that many muscle proteins are highly conserved, therefore the Unipept LCA for all peptides for which it was available can be viewed in electronic supplementary material, S3. Note: while uncharacterised proteins were the most abundant class, they have been excluded from this table because they represent many different proteins.

		protein	peptide count	unipept LCA	probable protein source
chestnut	foodcrust	cupin type-1 domain-containing protein	218	<i>Castanea crenata</i> and <i>Quercus lobata</i> ^a	input ingredient
		heat shock protein 70	97	<i>Prunus avium</i> , <i>Juglans regia</i> and <i>Q. lobata</i> ^a	non-specific
		starch synthase, chloroplastic/amyloplastic	90	<i>C. mollissima</i>	input ingredient
		SHSP domain-containing protein	81	<i>C. sativa</i> and <i>Q. lobata</i> ^a	input ingredient
		chitin-binding type-1 domain-containing protein	67	<i>C. sativa</i> , <i>Q. lobata</i> and <i>Salix viminalis</i> ^a	input ingredient
	ceramic	IF rod domain-containing protein	5	Mammalia	contaminant
		keratin, type II cytoskeletal 5	4	Mammalia	contaminant
		keratin, type II cytoskeletal 75	3	Sarcopterygii	contaminant
		glyceraldehyde-3-phosphate dehydrogenase	2	Craniata	contaminant
		glycogenin-1	2	Artiodactyla	contaminant
deer	foodcrust	phosphopyruvate hydratase (2-phospho-D-glycerate hydrolyase)	238	Mammalia ^b	non-specific
		troponin T, fast skeletal muscle	235	<i>C. elaphus hippelaphus</i>	input ingredient
		myosin-1	230	<i>C. elaphus hippelaphus</i>	input ingredient
		fructose-bisphosphate aldolase	221	Artiodactyla	non-specific
		alpha-1,4 glucan phosphorylase	196	Mammalia ^c	non-specific
	ceramic	globin family profile domain-containing protein	10	Cervidae	input ingredient
		IF rod domain-containing protein	10	Mammalia	non-specific
		keratin, type II cytoskeletal 5	8	Mammalia	non-specific
		myoglobin	5	Mammalia	non-specific
		glyceraldehyde-3-phosphate dehydrogenase	4	root	non-specific
	ceramic	haemoglobin subunit alpha	4	<i>C. elaphus hippelaphus</i>	input ingredient
		phosphopyruvate hydratase (2-phospho-D-glycerate hydrolyase)	4	Mammalia	non-specific
		superoxide dismutase [Cu-Zn]	4	not available	
		glyceraldehyde-3-phosphate dehydrogenase	930	Salmoninae	input ingredient

(Continued.)

Table 1. (Continued.)

	protein	peptide count	unipept LCA	probable protein source
	phosphopyruvate hydratase (2-phospho-D-glycerate hydrolyase)	820	Salmonidae	input ingredient
	myosin heavy chain, fast skeletal muscle-like	814	<i>S. salar</i>	input ingredient
	creatine kinase	496	<i>S. salar</i> and <i>Oncorhynchus kisutch</i>	input ingredient
	fructose-bisphosphate aldolase	390	<i>Salmo</i> ^d	input ingredient
ceramic	keratin, type II cytoskeletal 5	7	Mammalia	contaminant
	IF rod domain-containing protein	5	Mammalia	contaminant
	collagen alpha-2(I) chain isoform X3	4	Salmoninae	input ingredient
	collagen alpha-1(I) chain	3	<i>Salmo</i>	input ingredient
	keratin, type II cytoskeletal 75	3	Sarcopterygii	contaminant

^aPlant Unipept LCA analysis yielded multiple non-target plant taxa, possibly due to database absences.

^bOne Salmonidae peptide also detected.

^cOne Neopterygii peptide also detected.

^dThree Temnothorax peptides also detected.

then further reduced when buried. To explore this, hierarchical cluster plots based on highest peptide count grouped by leading razor protein were created (figure 4). For this, only proteins present in all three buried foodcrust replicates with >1 peptide matches were explored, as they were considered to consistently preserve at a quantity appreciable in general palaeoproteomic analysis. The cluster plots revealed that preservation varies by specific protein, and that not all proteins follow the same trend. One notable observation is an increased number of peptide matches to particular proteins in cooked samples when compared with the fresh ingredient, including: Troponin T3, Fast Skeletal Type (TNNT3), myoglobin (MB) and globulin-family profile domain-containing protein (haemoglobin) in deer, and phosphopyruvate hydratase (enolase) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH), and fructose-bisphosphate aldolase in salmon (figure 4, see also figure 3). For example, haemoglobin had a peptide count of 15 in the fresh ingredient, and an average peptide count of 37 in the unburied foodcrust replicates. One explanation for this phenomenon could be the role played by heat in denaturing proteins, partly degrading them and opening them up so that enzymatic cleavage is more efficient. In the buried foodcrusts, while the overall number of proteins is lower than unburied samples, some proteins retained relatively high peptide count compared with others. For example, in the deer samples haemoglobin had an average peptide count of 37 in the unburied foodcrust and 8 in the buried foodcrust, which was relatively high compared with most other proteins. Another example is ACTC protein (actin) which was also relatively abundant (by peptide count) in all three buried foodcrust replicates, and in the salmon samples the comparatively high peptide count proteins included myosin heavy chain fast skeletal muscle-like, GAPDH and fast myotomal muscle tropomyosin. In contrast, while phosphopyruvate hydratase (enolase) is the most abundant protein in terms of peptide count in fresh and cooked deer replicates, it was found in relatively low peptide count in buried foodcrust samples. This reveals that protein preservation is variable: certain proteins persist particularly well in buried foodcrusts while others do not.

3.1.5. Do buried foodcrust results reflect initial ingredient input?

The buried foodcrust samples were intended as analogues of archaeological foodcrusts, and thus to provide a baseline for the extractome that might be expected from archaeological foodcrusts. We investigated the extent to which buried cooked foodcrusts resemble the input protein composition.

Table 2. The top five most abundant proteins (by peptide count) preserved in foodcrust and ceramic samples after cooking and burial for six months. Data from replicates have been merged. Probable protein source is noted, including where LCA matches to input ingredient, where protein LCA is non-specific and could potentially belong to either input ingredient or another taxonomy, and where protein LCA could not possibly match to the input ingredient and is therefore derived from contamination. This table is based on proteins with the highest peptide count regardless of taxonomic specificity. We note that many muscle proteins are highly conserved; therefore, in electronic supplementary material, S8, we have noted the most abundant proteins (by peptide count) which match to taxonomic family (Salmonidae, Fagaceae or Cervidae) or lower detected in buried foodcrusts.

		protein	peptide count	unipect LCA	probable protein source		
chestnut	foodcrust	TATA box binding protein-associated factor (TAF) histone-like fold domain-containing protein	8	root	non-specific		
		actin, alpha cardiac muscle 1	6	Eukaryota	non-specific		
		cupin type-1 domain-containing protein	6	Fagaceae and <i>Q. lobata</i>	input ingredient		
		heat shock protein 70	5	root	non-specific		
		IF rod domain-containing protein	5	Mammalia	contaminant		
	ceramic	fast myotomal muscle troponin-T	5	Neopterygii	contaminant		
		IF rod domain-containing protein	3	Mammalia	contaminant		
		keratin, type II cytoskeletal 5	2	Mammalia	contaminant		
		deer	foodcrust	actin, alpha cardiac muscle 1	29	Eukaryota	non-specific
				globin family profile domain-containing protein	25	Cervidae	input ingredient
TATA box binding protein associated factor (TAF) histone-like fold domain-containing protein	12			root	non-specific		
elongation factor 1-alpha	11			root	non-specific		
myoglobin	11			Pecora	input ingredient		
	ceramic	myosin-1	6	Craniata	non-specific		
		actin, alpha cardiac muscle 1	5	root	non-specific		
		pyruvate kinase	4	Mammalia	non-specific		
		creatine kinase	2	Craniata	non-specific		
		filamin-C	2	Craniata	non-specific		
		L-lactate dehydrogenase	2	Neopterygii	contaminant		
		myosin binding protein C1	2	Pecora	input ingredient		
		non-selective voltage-gated ion channel VDAC1	2	Mammalia	non-specific		
		salmon	foodcrust	glyceraldehyde-3-phosphate dehydrogenase	71	Salmonidae	input ingredient
				myosin heavy chain, fast skeletal muscle-like	56	Salmoninae	input ingredient
fast myotomal muscle tropomyosin (Tropomyosin alpha-1 chain)	46			Salmonidae	input ingredient		

(Continued.)

Table 2. (Continued.)

	protein	peptide count	unipept LCA	probable protein source
	phosphopyruvate hydratase (2-phospho-D-glycerate hydrolyase)	46	Salmonidae	input ingredient
	fructose-bisphosphate aldolase	32	Salmonidae	input ingredient
ceramic	IF rod domain-containing protein	6	Mammalia	contaminant
	collagen alpha-2(I) chain	3	Artiodactyla	contaminant
	keratin, type II cytoskeletal 5	3	Sarcopterygii	non-specific
	keratin, type II cytoskeletal 6A	2	Sarcopterygii	non-specific

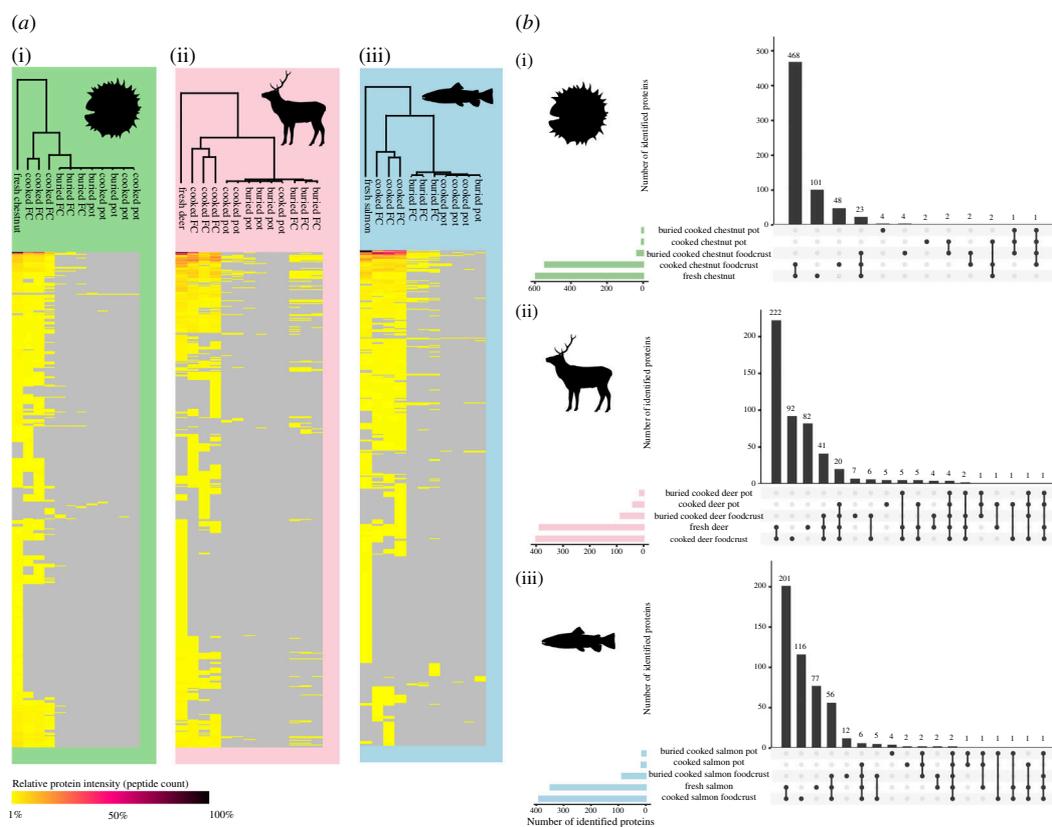


Figure 3. (a) Hierarchical cluster analysis of proteins identified in all chestnut (i), deer (ii) and salmon (iii) foodcrust and ceramic samples (Euclidean correlation), created in Perseus v. 1.6.14. (b) Upset plot displaying intersection of proteins observed in ceramic, foodcrust and fresh sample categories for chestnut (i), deer (ii) and salmon (iii) samples. Created using UpsetR [65].

‘Upset diagrams’, an alternative to Venn diagrams [66], were used to investigate the proteins shared by each sample type (figure 3b). They revealed that for all ingredients, the fresh and cooked foodcrust samples contained the most common shared proteins of any set of samples, indicating that they are compositionally most similar. The largest overlapping group following this was fresh ingredients, cooked foodcrusts and buried foodcrusts in all cases, indicating that despite the reduction of proteins in buried samples, they still somewhat reflect the initial input composition.

To understand whether buried samples revealed the input ingredient taxonomy, Unipept Desktop was used to assign LCA for each peptide where possible (see electronic supplementary material, S3, for peptide Unipept LCAs). Data were then filtered following commonly used proteomic standards (greater than two PSMs to support a protein). In contrast to the buried ceramic samples which harboured no species-specific protein results, all buried deer and salmon foodcrust replicates produced sufficient proteomic evidence to identify the specific input ingredient to a species level in at least

Similarly, this has been observed in the analysis of ancient dental calculus, where a bias towards the detection of milk proteins over other dietary derived proteins has been reported [15]. This has obvious implications on the interpretation of archaeological results, for example rendering plants less visible compared with other ingredients.

Furthermore, in this study, the lack of annotated proteins from some plant species has become starkly apparent. We note that a large number of the peptides identified in the chestnut samples matched to uncharacterised proteins, rendering their analysis difficult. Moreover, the absence of a *C. sativa* proteome in Uniprot at the time of analysis necessitated the use of *C. mollissima* in this investigation—which may impact identifications. This concurs with Hendy *et al.*'s previous comment on the dependence of shotgun proteomics on available databases, and its impact on plant identification [11]. Plants which have much larger proteomes are often absent, particularly for species that are not of current commercial relevance, such as heirloom cultivars. Database absence likely contributes to the lower detection rate of plants in archaeological samples or their detection at higher taxonomic specificity.

3.2. Exploration of characteristics enabling protein survival in buried foodcrust samples

3.2.1. Why do particular proteins survive cooking and burial?

Having identified which proteins persist after cooking, foodcrust formation and burial, as well as the overall trend in the number of proteins preserved, we now explore whether these proteins harbour particular characteristics which may facilitate their survival. In this study, it is apparent that degradation is not universal nor linear between different proteins (in contrast to [19]). As discussed above, certain proteins persist particularly well in buried samples while others do not (figure 4), leading us to hypothesise that individual protein properties aid in their preservation. Previously, particular characteristics have been hypothesised to impact protein preservation in or on pottery and other mineral surfaces [17–19,57,67–69]. We wished to explore if there were particular characteristics on either a peptide or protein level that may be impacting the potential preservation of proteins in buried foodcrust samples.

A range of peptide and protein characteristics were investigated. These included concentration of different amino acids, peptide length, peptide hydropathicity, peptide isoelectric point, the sample's lipid content, protein melting temperature, disorder prediction, amyloid propensity, protein secondary structure and RSA at a given peptide (figure 5). The characteristics of bulk amino acid concentration, peptide length, peptide hydropathicity and bulk deamidation were calculated manually, while various tools were used to calculate the other characteristics. These included: protein thermal stability: DeepSTABp [70]; amyloid propensity: AMYPred-FRL [71]; disorder prediction: IUPred [72]; isoelectric point: IPC [73]; and protein secondary structure and RSA: 'Predict_Property', a standalone, offline version of RaptorX Web server [74,75] (https://github.com/realbigws/Predict_Property). A script was written to extract the secondary structure, RSA and disorder prediction for each peptide in the dataset (https://github.com/miranda-e/peptide_property_analyser). All characteristic results were then compiled and displayed using an R script (figure 5). The full details of data analysis are present in electronic supplementary material, S10. While we initially aimed to explore the relationship between protein tertiary structure and preservation, this proved challenging due to the paucity of tools for this analysis, and the poorly annotated or modelled nature of many target proteins. As an alternative way of investigating structural characteristics, we include secondary structure, disorder prediction and amyloid prediction. The characteristics of protein function and cell location were initially investigated through Gene Ontology (GO) terms; however, these explorations were limited by relatively low levels of annotation in the target species' databases (electronic supplementary material, S11), so this approach was not pursued. Factors other than cooking and burial such as the extraction protocol and data analysis parameters will have also impacted the composition of the extractome; however, all samples have experienced the same extraction and search protocol. Buried chestnut foodcrust data were removed from analysis due to very low peptide counts, and high levels of field cross-contamination (electronic supplementary material, S5), rendering any interpretation with statistical weight challenging.

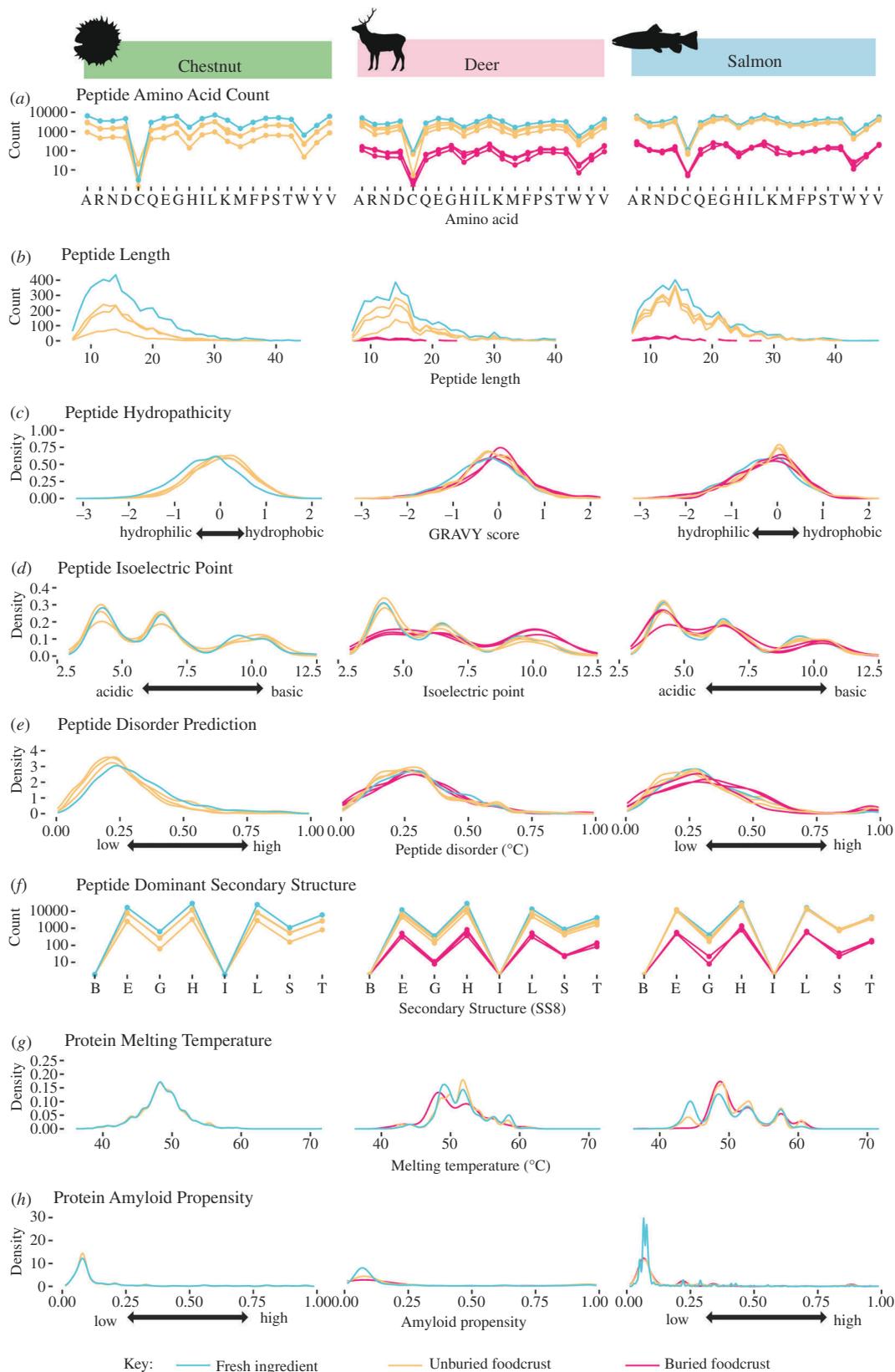


Figure 5. Protein and peptide characteristics for fresh (blue), cooked foodcrusts (orange) and buried foodcrusts (pink) extractomes of chestnut, deer and salmon, including: (a) peptide amino acid count (points connected for visibility), (b) peptide length, (c) peptide hydropathicity, (d) peptide isoelectric point, (e) peptide disorder prediction, (f) peptide dominant secondary structure (points connected for visibility), (g) protein melting temperature and (h) protein amyloid propensity. Buried chestnut replicates removed due to insufficient sample size and field contamination.

3.2.2. Proteins that survive in cooked and buried foodcrusts do not have particular amino acid compositions or secondary structures

Previously, the impact of reactive amino acid content has been noted as a factor influencing protein survival [19], as has the impact of higher order structure and the location of a peptide within the structure, which may protect or expose particular peptides [19,68]. Amino acid sequence also has been reported as a driving factor in protein abundance by determining conformational stability and reducing synthesis cost [76]. Bulk amino acid count (figure 5a) and peptide secondary structure (figure 5f) demonstrated no global change between their fresh, cooked and buried state for any ingredient, indicating that they likely did not play a substantial global role in peptide preservation. Secondary structure is innately linked to a protein's function and stability, with certain structures being more stable. Secondary structure was collected using the following categories: G: 310 helix; H: alpha-helix; I: pi-helix; E: beta-strand; B: beta-bridge; T: beta-turn; S: high curvature loop; L: irregular (figure 5f). Secondary structures have been reported to be distributed across all proteins in the following ratio: alpha-helix, beta-strand, irregular, beta-turn, high curvature loop, 310 helix, beta-bridge, pi-helix = 34:21:20:11:9:4:1:0 [77]. Similar ratios in line with the background distribution were observed in all samples, with only minor variation between input ingredients—meaning that secondary structure does not seem to be a factor in determining which peptides we detect in the extractome. Moreover, as secondary structure and bulk amino acid distribution were no different in cooked or buried samples for any ingredient, it appears that these characteristics do not impact protein preservation.

3.2.3. More hydrophobic peptides are slightly more likely to survive cooking and burial

Previously, the potential role of protein hydrophobicity in protein preservation has been hypothesised, whereby hydrophilic proteins leach from ceramics during washing and/or burial [18,19,57]. In this study, we investigated whether peptide hydrophobicity (solubility) correlated with peptide count in fresh, cooked or buried samples, with the hypothesis that less water-soluble (hydrophobic) peptides might preferentially survive in buried samples. Grand average of hydropathy (GRAVY) score, a standard measure of protein polarity, was calculated on a peptide level for fresh, cooked foodcrust and buried foodcrust samples (figure 5c). Statistical testing supported a difference in GRAVY scores generated across fresh, unburied and buried sample types for the pooled ingredients (Kruskal–Wallis test, $p < 0.001$). Subsequent statistical testing with a Wilcoxon rank sum test supported a difference in GRAVY values between fresh ingredients and their cooked foodcrust counterparts ($p < 0.001$) and between unburied and buried deer foodcrusts ($p < 0.02$). For both salmon and deer, GRAVY score appeared to increase slightly in cooked and buried foodcrusts compared with fresh samples (figure 5c), indicating that peptides were generally slightly more hydrophobic in cooked and buried samples than in fresh samples. This result supports previous hypotheses that water leaching over time may reduce soluble protein content in pottery [19,57], leaving slightly more hydrophobic proteins to be detected in higher peptide counts in buried samples, although we note that an increase in hydrophobic peptides was also observed after cooking, indicating that hydrophilic peptides may also be less likely to be entrapped in foodcrust.

3.2.4. Cooking may liberate peptides located deep within the protein's three-dimensional structure

The potential role of higher order structure and the location of a peptide within the structure, which may protect or expose particular peptides, has been noted [19,68]. RSA is a measure of the exposure of an amino acid within its tertiary structure, and therefore how accessible that residue is to solvents (i.e. amino acids located deeper within the three-dimensional structure are less accessible to solvents). The results revealed that cooked foodcrusts had a higher proportion of peptides with deep RSA (i.e. peptides with amino acids located deep within the tertiary protein structure), than did fresh ingredients, which is most marked in chestnut. One explanation for this is that as tertiary protein structures unfold during denaturation during cooking, peptides which are located deep within the protein structure become more accessible to extraction than they are in uncooked ingredients. This means that in uncooked archaeological samples, we are probably less likely to see deep peptides than in their cooked counterparts. In buried foodcrust samples, the proportion of 'deep' RSA peptides continued to decrease in deer samples but increased in salmon samples, providing inconsistent results.

3.2.5. Protein thermal stability may impact peptide survival

Thermal stability is the ability of proteins to resist changes in structure caused by heating. We investigated this characteristic with the hypothesis that thermostable peptides would persist through the cooking, entrapment and burial process. Melting temperature (T_m) is often used as a measure of protein thermal stability. Proteins surviving in buried salmon foodcrusts appeared to be slightly more thermostable than the fresh ingredient and unburied foodcrust, with fewer peptides from proteins of low thermal stability ($T_m < 45^\circ\text{C}$) surviving in buried foodcrust replicates (figure 5g). This may indicate that thermally stable peptides were more likely to survive in buried salmon than proteins with lower thermal stability. Conversely in buried deer foodcrust, statistical testing supported a difference in T_m in unburied and buried deer foodcrusts at $p < 0.05$ (Wilcoxon rank sum test) with peptides from less thermally stable proteins appearing to survive better than peptides with higher thermal stability (figure 5g), further demonstrating the varied behaviour of different ingredients.

3.2.6. Certain properties may aid in the preservation of particular ingredients or proteins

Some characteristics demonstrated changes after cooking and burial only for particular ingredients. Amyloid propensity and disorder prediction showed changes primarily for salmon, but not in deer samples. Previously, Collins *et al.* [69] speculated that entropic effects would promote survival of flexible molecules that could adapt and bind to the mineral surface. Demarchi *et al.* provided evidence that mineral binding of a small flexible acid rich region was responsible for the persistence of a peptide form of a c-type lectin of African ostrich eggshell into deep time [78]. Most recently, Scott [79] proposed that the robust nature of amyloid fibrils and other factors contributing to protein aggregation may explain the presence of particular proteins and peptides in the archaeological record, noting that dietary proteins persisting in ancient dental calculus are often amyloidogenic. The analysis of intrinsically disordered proteins revealed that in the case of salmon, the proportion of peptides with high disorder prediction slightly increased following cooking and remained high during burial meaning that more flexible peptides become relatively more representative than inflexible ones (figure 5e). Deer samples did not display this trend. Statistical testing supported a difference in disorder propensity between all fresh ingredients and unburied foodcrusts at $p < 0.001$ (Wilcoxon rank sum test), while a difference was only observed between unburied and buried foodcrusts in the case of salmon ($p < 0.05$). Similarly, in the analysis of amyloid propensity, while statistical testing supported a difference in amyloid propensity between both deer and salmon unburied and buried foodcrusts at $p < 0.05$ (Wilcoxon rank sum test), visual scrutiny of figure 5h only revealed a notable difference for salmon samples. This revealed that peptides which could readily stack were more likely to be detected in buried salmon samples than fresh or cooked (figure 5h), potentially indicating that this characteristic plays a role in the preservation of some salmon proteins. In the case of deer, the density of peptides with low isoelectric points (i.e. acidic, water soluble peptides) decreased in the buried samples compared with the fresh ingredient and unburied foodcrust, with a statistical analysis supporting a difference in pI between fresh deer and unburied foodcrust and unburied and buried foodcrust at $p < 0.001$, while the density of peptides with high pIs (basic peptides) increased (figure 5d). This demonstrates that acidic peptides survived poorly in buried deer foodcrust, a change not observed in salmon or chestnut. This indicates that certain characteristics may aid in the preservation of particular ingredients or proteins.

While the buried chestnut samples were not included in the broader characterisation analysis due to small sample size, we note that the only two chestnut-specific proteins to survive in buried foodcrust samples (Cupin type-1 domain-containing protein and Chitin-binding type-1 domain-containing protein) are both allergenic. Allergenic proteins are often characterised by their stability, in terms of either their tertiary structure (such as the beta-barrel observed in Cupin-type proteins), or resistance to heat or digestive degradation [80], and have previously been observed to preferentially preserve in ancient dental calculus [81].

The potential role of lipids in creating water-impermeable barriers which may shelter proteins in pottery from forces of degradation has previously been noted [19], although lipids may complicate protein extraction [18]. A vast body of work has explored the impact of organic content on protein and nitrogen preservation in sediments [82–87]. Previously, some of the buried foodcrust replicates were analysed by flame ionisation detector [52], facilitating an opportunity to examine any correlations between protein and lipid content in the same vessel. The total peptide and protein count was compared with these lipid quantities generated by Bondetti *et al.* [52] for each sample. We note that

there was a surprising level of variation in lipid concentration between replicates of identical input ingredients and weight. This revealed that in addition to the impact of cooking practise and frequency of use [55], even identically processed foodcrusts are not homogeneous. However, due to the small sample size, it was not possible to reveal the impact of lipid content on protein detection (electronic supplementary material, S12). Future controlled dosing studies would further address the impact of lipid content on protein preservations in ceramics and their residues.

3.2.7. Markers of diagenesis

Peptide length was investigated as a potential indicator of diagenesis from the cooking and entrapment process and/or from burial. Differences in peptide length were observed between fresh ingredients, cooked foodcrusts and buried foodcrusts. The distribution generally indicated that shorter peptides were more likely to be detected in buried foodcrust samples (figure 5b). Statistical analysis supports that there is a difference in peptide length between fresh ingredients and buried and unburied foodcrusts (Kruskal–Wallis test, $p < 0.001$), in particular between both fresh chestnut and salmon and their unburied foodcrusts at $p < 0.05$, and between both deer and salmon unburied and buried foodcrusts at $p < 0.001$ (Wilcoxon rank sum tests). Peptide length generally decreased upon cooking, with chestnut providing the most marked reduction. It is notable that long peptides (above ~20 amino acids) were very rarely detected in buried samples of any ingredient. This may demonstrate that long peptides are rarely preserved in buried samples, or alternatively, as long peptides are always less numerous, that the smaller sample size of buried peptides reduces the probability of their detection. We note that this observation, that more degraded samples generally yield shorted peptides, is in line with previous work [88]. As peptide length decreases, it seems likely that the probability of identifying a tissue and taxonomically specific sequence would also decrease. This analysis further revealed differences in the behaviour of the different ingredients.

Deamidation of asparagine and glutamine was explored as a marker of preservational quality [89]. In certain proteins, deamidation has been reported to change protein tertiary structure [90]. In the case of deer and salmon, the highest proportion of peptides bearing deamidation were the unburied foodcrust samples, demonstrating that heating plays a role in deamidation. In the case of salmon and deer, the proportion of deamidated peptides fell after burial (figure 6). Notably, the proportion of deamidated peptides does not continue to increase in the buried state considering that deamidation has been used as an indicator of protein preservation quality, perhaps because the majority of possible deamidations have already occurred during heating, and timescales of chemically mediated deamidation are slow at burial temperatures [91]. This demonstrates that deamidation may not be a good indicator of peptide authenticity where modern contaminants may have undergone heating. Unfortunately, due to the necessary peptide filtration steps discussed in electronic supplementary material, S1, it was not possible to apply deamiDATE here to assess site-specific deamidation. Further investigation of site-specific PTMs may reveal evidence for aspects of food preparation and taphonomy.

3.2.8. Case study: haemoglobin

While no single physicochemical characteristic investigated appears to explain why some proteins survive better than others in buried foodcrusts, it is apparent that these properties may play a role for individual ingredients and proteins. One protein that preserves relatively well in the foodcrust samples is globin-domain containing protein (haemoglobin), so we investigated this protein in greater detail as a case study to understand, on an individual protein level, if particular characteristics may be aiding in its survival. These are likely too complex and varied to be detected broadly, across the whole sample set.

GRAVY score, amyloid propensity, secondary structure and RSA were overlaid on top of a peptide count heatmap (figure 7), illustrating that in the case of the globin family protein, the peptides which preserve most commonly in the buried sample are located on the most hydrophobic parts of the protein, and often in regions with high amyloid propensity (figure 7). This demonstrates in the case of globin that hydrophobic (insoluble) peptides appear to be more likely to persist in our buried samples than hydrophilic peptides. However, this means that a complex array of mechanisms, interactions and protein compositions must be impacting the survival of proteins and peptides in buried foodcrusts.



Figure 6. Proportion of deamidated peptides. Note: cooked and buried foodcrust replicates = 3, fresh ingredient replicates = 1.

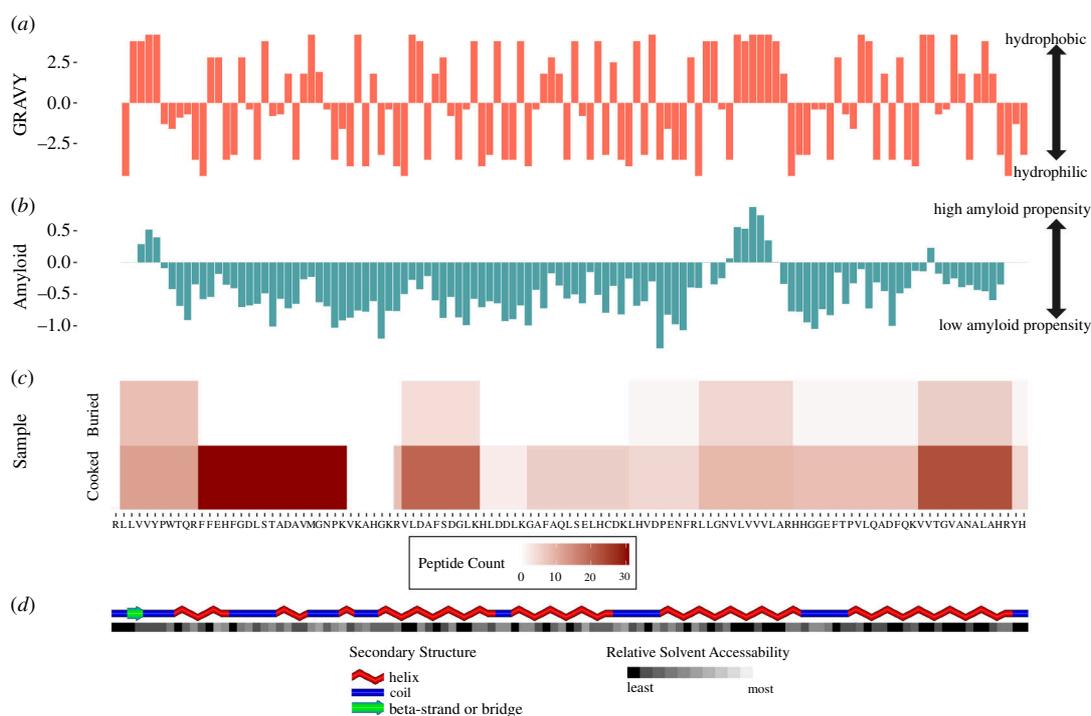


Figure 7. Globin characteristic: hydrophaticity (a), amyloid prediction (b) with heatmaps of peptide counts for each region (c) of buried (top) and cooked (bottom) deer foodcrust samples. (d) Secondary structure of sequence: blue = coil, red = helix; and RSA: black = completely buried, white = fully exposed (created in POLYVIEW-2D [92]).

3.2.9. No property single-handedly explains why particular proteins survive cooking and burial

In this study, many protein characteristics were investigated to understand potential explanations for why certain proteins were detected in buried foodcrusts. We observe that the input ingredient was the most influential factor in protein survival, and individual proteins follow different preservation trends. For example, in the foodcrusts created by cooking deer meat, basic peptides seem to preferentially survive the cooking and burial environment. But this trend was not observed for proteins detected from salmon foodcrust. Similarly, proteins with higher thermal stability, disorder prediction and amyloid propensity appear to preferentially preserve in buried salmon foodcrusts, yet not in deer. Some trends were observed across all food types including that slightly hydrophobic peptides are more likely to preserve after cooking and/or burial, that the recovery of longer peptides tended

to decrease after cooking and burial, and that deamidation increased following cooking, but not necessarily burial. However, no trend could single-handedly explain why particular peptides and proteins survive cooking and burial. While few trends in protein characteristic were observed across all ingredients, when viewed on an individual protein level, haemoglobin peptides that were slightly hydrophobic and had high amyloid propensity were more likely to preserve in buried foodcrust samples, even though they were not particularly numerous in the unburied sample. This indicates that particular protein or peptide characteristics are indeed involved in determining which proteins/peptides survive in buried samples, but that these complex interactions are likely to vary based on a myriad of factors, and may be obscured when the data are viewed more broadly.

3.3. Challenges and future directions

A key goal of this study is to consider expectations for the survival of proteins in foodcrusts and ceramics in archaeological contexts. While the buried foodcrusts in this investigation revealed a number of protein identifications with an informative level of protein and taxa specificity, it is important to note that the samples were buried for only six months and in a temperate climate, and therefore may not be comparable to samples of much older temporal and thermal age. While Barker *et al.* [19] reported that protein content dropped rapidly upon burial and at a slower rate thereafter, the exact rate of dietary protein decay over longer periods of time and in other climatic conditions is poorly understood. The results of this study should therefore be seen as an upper limit to protein preservation, as it may well be that samples of considerable age or from hot climates may not yield confidently preserved proteins, necessitating a cautious approach to the destruction of valuable samples.

In this study, we did not analyse a control sample of the soil in which experimental samples were buried. This was because this experiment was conceived of after the samples used in a different experiment [52] had been excavated, in an effort to maximise the use of and apply comparative approaches to existing material. Steps were taken to mitigate against soil contamination, including removing the most external layer of foodcrust prior to sampling, but we acknowledge that the lack of a soil sample means that it is not possible to filter for any endogenous protein content which may have been present in the soil. However, we note that while endogenous soil proteins may have increased the level of non-specific protein content across all samples, it is unlikely that salmon, deer or chestnut proteins, and even non-taxonomically specific proteins from specific tissues such as muscle proteins, would be present in the soil yet contaminate only certain samples. In future work, the use of a soil sample would enable the testing of background contaminants present in soil.

There are a myriad of factors which could not be investigated in this study, yet may impact protein entrapment and preservation in ancient foodcrusts and ceramics, and thus warrant further investigation. One such factor is the intensity of thermal exposure, for instance with foodcrusts which experienced lower thermal intensity potentially yielding higher protein recovery than those processed at higher thermal intensity. Future work focusing on the impact of increasing thermal exposure to dietary protein detection in ancient foodcrusts may allow more targeted sample strategies. Given the exceptional protein recovery from calcified pottery residues [11,12,93], another factor which will likely impact the preservation of protein in foodcrusts is the level of calcium in the water used during cooking. Here, Yorkshire municipal tap water classified as 'Hard' (average calcium content 107.2 mg L^{-1} [94]) was used in the sample creation process, which may have aided in protein entrapment or preservation. Therefore, investigating the impact of a range of calcium concentrations on protein entrapment is a valuable avenue for future research.

In this experiment, the input ingredient itself appeared to play a strong role in the difference in protein preservation/detection between samples. This has implications on ingredient visibility in archaeological interpretations on foodcrust results, particularly for plants. Future investigation of the impact of ingredient mixing on protein preservation will be necessary to understand the full implications of this. Furthermore, further investigation into the impact of protein interactions with other macronutrients such as lipids and carbohydrates on protein preservation is required to fully understand this issue.

We are aware that extraction and LC-MS/MS protocols will have impacted the peptides which are detected. As a result, we are viewing the detected proteins and peptides through a lens of the analytical processes they have been through. Some of these may be taphonomic or cultural such as cooking and burial, but others are inflicted by post-depositional processes, storage, handling and

extraction. Untangling the role of each will be imperative in understanding the impact of any one variable. The extraction of proteins from ceramics continues to be a challenge, with ongoing optimisation [17–19,24,56,68,95,96]. The protocol used in this analysis is not optimised for extracting protein from ceramic samples, which may explain their poor recovery here. Ongoing protocol development will hopefully improve the extraction of dietary proteins from ceramics (e.g. [24]). While peptide count and protein count are important metrics used in protein identification in ancient protein studies, in this study, we intended to use Label Free Quantification to quantify differences in protein content between samples; however, unfortunately, this was not possible due to the necessary filtering steps outlined in electronic supplementary material, S1. Future studies would benefit from protein quantification in order to more informatively assess changes in relative and total protein or peptide abundance.

4. Conclusion

In this investigation, we sought to explore the utility of proteomic analysis of foodcrusts and ceramics in understanding ancient food preparation, by examining the extent to which the buried foodcrust extractome reflects input ingredients. The results revealed that foodcrusts harboured more preserved proteins than ceramics, and we note, in line with previous studies, that proteins are not easily extracted from ceramic matrices [17,18]. Sufficient taxonomic and tissue-specific identifications were made to detect the relevant ingredient in buried foodcrusts created by cooking deer and salmon meat, and sometimes chestnut flour, demonstrating that ancient foodcrusts may be a viable matrix from which to extract dietary proteins, but that not all ingredients were equal in protein retrieval revealing that input ingredient biases protein recovery. Preservation was not universal between proteins and peptides; those which were most numerous in the fresh meat and flour were not necessarily the most frequently identified in buried samples. These biases have implications on archaeological interpretations of ancient foodcrusts and ceramics, namely that certain ingredients and proteins will be unlikely to be detected in ancient samples. While this study has focused on pottery and foodcrust residues, several of the observations made here regarding protein preservation will likely apply to other areas within the field of palaeoproteomics. It is clear further work is necessary to understand the biases of input ingredients, particularly when they are mixed, and over longer periods of vessel use and burial. Secondly, we attempted to characterise the proteins and peptides retrieved from buried foodcrust samples, to understand physicochemical factors influencing their preservation. While we observed that more hydrophobic peptides were slightly more likely to survive cooking and burial, no property was seen to single-handedly explain why particular proteins/peptides survive in buried foodcrusts, with results indicating that certain properties act on protein preservation in complex ways requiring further investigation, or that characteristics not investigated here may play a role. This study demonstrates the value of experimental analyses to anticipate a maximum baseline of protein results from archaeological samples.

Ethics. This work did not require ethical approval from a human subject or animal welfare committee.

Data accessibility. The mass spectrometry proteomics data, databases used and parameters file have been deposited to the ProteomeXchange Consortium via the PRIDE [97] partner repository with the dataset identifier PXD050001 and 10.6019/PXD050001.

Supplementary material is available online [98].

Declaration of AI use. AI-assisted technology was used to support code writing and debugging of the 'peptide_property_analyser.py' python script created for this paper. The script was then checked by a software developer and tested extensively with test datasets of known output to validate it. AI-assisted technologies were not used in any other capacity in this paper.

Authors' contributions. M.E.: conceptualization, data curation, formal analysis, investigation, methodology, software, visualization, writing—original draft, writing—review and editing; R.H.: investigation, writing—review and editing; O.J.B.: software; M.B.: conceptualization, methodology, resources; O.E.C.: conceptualization, methodology, resources, writing—review and editing; M.J.C.: investigation, supervision, writing—review and editing; J.H.: conceptualization, methodology, resources, supervision, visualization, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This project has received funding from the European Union's EU Framework Programme for Research and Innovation Horizon 2020 under grant agreement no. 722606: TEMPERA. This work was supported by the Cambridge Trust and Newnham College through an international studentship awarded to M.E. and a Philip Leverhulme Prize awarded to J.H. We thank the ERC Advanced Grant INDUCE (Innovation, Dispersal and Use of Ceramics in NE Europe, ERC-ADG-2015 number 695539), which funded the creation of experimental samples.

Acknowledgements. We gratefully acknowledge the use of the Thermo Scientific Orbitrap Fusion™ Tribrid™ in the York Centre of Excellence in Mass Spectrometry. The centre was created thanks to a major capital investment through Science City York, supported by Yorkshire Forward with funds from the Northern Way Initiative, and subsequent support from EPSRC (EP/K039660/1; EP/M028127/1). We thank Niklas Hausmann for assistance with R, Alexandre Lucquin for assistance locating lipid data files, and Samantha Leggett for answering statistical questions. We would like to thank the two anonymous reviewers for their helpful feedback.

References

1. Runge AKW *et al.* 2021 Palaeoproteomic analyses of dog palaeofaeces reveal a preserved dietary and host digestive proteome. *Proc. R. Soc. B* **288**, 20210020. (doi:10.1098/rspb.2021.0020)
2. Colonese AC *et al.* 2017 New criteria for the molecular identification of cereal grains associated with archaeological artefacts. *Sci. Rep.* **7**, 6633. (doi:10.1038/s41598-017-06390-x)
3. Yang Y, Shevchenko A, Knaust A, Abuduresule I, Li W, Hu X, Wang C, Shevchenko A. 2014 Proteomics evidence for kefir dairy in early bronze age china. *J. Archaeol. Sci.* **45**, 178–186. (doi:10.1016/j.jas.2014.02.005)
4. Hong C, Jiang H, Lü E, Wu Y, Guo L, Xie Y, Wang C, Yang Y. 2012 Identification of milk component in ancient food residue by proteomics. *PLoS One* **7**, e37053. (doi:10.1371/journal.pone.0037053)
5. Xie M, Shevchenko A, Wang B, Shevchenko A, Wang C, Yang Y. 2016 Identification of a dairy product in the grass woven basket from Gumugou cemetery (3800BP, Northwestern China). *Quat. Int.* **426**, 158–165. (doi:10.1016/j.quaint.2016.04.015)
6. Buckley M, Melton ND, Montgomery J. 2013 Proteomics analysis of ancient food vessel stitching reveals >4000-year-old milk protein. *Rapid Commun. Mass Spectrom.* **27**, 531–538. (doi:10.1002/rcm.6481)
7. Cappellini E *et al.* 2010 A multidisciplinary study of archaeological grape seeds. *Naturwissenschaften* **97**, 205–217. (doi:10.1007/s00114-009-0629-3)
8. Wilkin S, Hommel P, Ventresca Miller A, Boivin N, Pedergrana A, Shishlina N, Trifonov V. 2023 Curated cauldrons: preserved proteins from early copper-alloy vessels illuminate feasting practices in the Caucasian steppe. *iScience* **26**, 107482. (doi:10.1016/j.isci.2023.107482)
9. Wilkin S, Bayarsaikhan J, Ganbold A, Batsuuri A, Ishtseren L, Nakamura D, Eregzen G, Ventresca-Miller A, Miller BK. 2024 Cauldrons of Bronze Age nomads reveals 2700 year old yak milk and the deep antiquity of food preparation techniques. *Sci. Rep.* **14**, 11625. (doi:10.1038/s41598-024-60607-4)
10. Tan L, Fan X, Rao H, Zhang H, Yang Y. 2024 The scientific analysis of the bronze mous excavated from Wushan, Chongqing, China: new perspectives from alloy composition analysis and proteomic method. *Herit. Sci.* **12**, 88. (doi:10.1186/s40494-024-01203-7)
11. Hendy J *et al.* 2018 Ancient proteins from ceramic vessels at Çatalhöyük West reveal the hidden cuisine of early farmers. *Nat. Commun.* **9**, 4064. (doi:10.1038/s41467-018-06335-6)
12. Evans M *et al.* 2023 Detection of dairy products from multiple taxa in late neolithic pottery from Poland: an integrated biomolecular approach. *R. Soc. Open Sci.* **10**, 230124. (doi:10.1098/rsos.230124)
13. Jeong C *et al.* 2018 Bronze Age population dynamics and the rise of dairy pastoralism on the eastern Eurasian steppe. *Proc. Natl Acad. Sci. USA* **115**, E11248–E11255. (doi:10.1073/pnas.1813608115)
14. Wilkin S *et al.* 2021 Dairying enabled early Bronze Age Yamnaya steppe expansions. *Nature* **598**, 629–633. (doi:10.1038/s41586-021-03798-4)
15. Hendy J *et al.* 2018 Proteomic evidence of dietary sources in ancient dental calculus. *Proc. R. Soc. B* **285**, 20180977. (doi:10.1098/rspb.2018.0977)
16. Geber J, Tromp M, Scott A, Bouwman A, Nanni P, Grossmann J, Hendy J, Warinner C. 2019 Relief food subsistence revealed by microparticle and proteomic analyses of dental calculus from victims of the Great Irish Famine. *Proc. Natl Acad. Sci. USA* **116**, 19380–19385. (doi:10.1073/pnas.1908839116)
17. Craig OE, Collins MJ. 2002 The removal of protein from mineral surfaces: implications for residue analysis of archaeological materials. *J. Archaeol. Sci.* **29**, 1077–1082. (doi:10.1006/jasc.2001.0757)
18. Craig OE, Collins MJ. 2000 An improved method for the immunological detection of mineral bound protein using hydrofluoric acid and direct capture. *J. Immunol. Methods* **236**, 89–97. (doi:10.1016/s0022-1759(99)00242-2)
19. Barker A, Dombrosky J, Venables B, Wolverson S. 2018 Taphonomy and negative results: an integrated approach to ceramic-bound protein residue analysis. *J. Archaeol. Sci.* **94**, 32–43. (doi:10.1016/j.jas.2018.03.004)
20. Tanasi D, Cucina A, Cunsolo V, Saletti R, Di Francesco A, Greco E, Foti S. 2021 Paleoproteomic profiling of organic residues on prehistoric pottery from Malta. *Amino Acids* **53**, 295–312. (doi:10.1007/s00726-021-02946-4)
21. Solazzo C, Fitzhugh WW, Rolando C, Tokarski C. 2008 Identification of protein remains in archaeological potsherds by proteomics. *Anal. Chem.* **80**, 4590–4597. (doi:10.1021/ac800515v)
22. Chowdhury MP, Campbell S, Buckley M. 2021 Proteomic analysis of archaeological ceramics from Tell Khaiber, southern Iraq. *J. Archaeol. Sci.* **132**, 105414. (doi:10.1016/j.jas.2021.105414)
23. Dallongeville S, Garnier N, Casasola DB, Bonifay M, Rolando C, Tokarski C. 2011 Dealing with the identification of protein species in ancient amphorae. *Anal. Bioanal. Chem.* **399**, 3053–3063. (doi:10.1007/s00216-010-4218-2)

24. Pal Chowdhury M, Makarewicz C, Piezonka H, Buckley M. 2022 Novel deep eutectic solvent-based protein extraction method for pottery residues and archaeological implications. *J. Proteome Res.* **21**, 2619–2634. (doi:10.1021/acs.jproteome.2c00340)
25. Admiraal M *et al.* 2023 The role of salmon fishing in the adoption of pottery technology in subarctic Alaska. *J. Archaeol. Sci.* **157**, 105824. (doi:10.1016/j.jas.2023.105824)
26. Heron C, Craig OE. 2015 Aquatic resources in foodcrusts: identification and implication. *Radiocarbon* **57**, 707–719. (doi:10.2458/azu_rc.57.18454)
27. Courel B *et al.* 2021 The use of early pottery by hunter-gatherers of the Eastern European forest-steppe. *Quat. Sci. Rev.* **269**, 107143. (doi:10.1016/j.quascirev.2021.107143)
28. Lucquin A *et al.* 2016 Ancient lipids document continuity in the use of early hunter-gatherer pottery through 9,000 years of Japanese prehistory. *Proc. Natl Acad. Sci. USA* **113**, 3991–3996. (doi:10.1073/pnas.1522908113)
29. Shevchenko A, Schuhmann A, Thomas H, Wetzel G. 2018 Fine endmesolithic fish caviar meal discovered by proteomics in foodcrusts from archaeological site Friesack 4 (Brandenburg, Germany). *PLoS One* **13**, e0206483. (doi:10.1371/journal.pone.0206483)
30. Roffet-Salque M, Dunne J, Alftoft DT, Casanova E, Cramp LJE, Smyth J, Whelton HL, Evershed RP. 2017 From the inside out: upscaling organic residue analyses of archaeological ceramics. *J. Archaeol. Sci.* **16**, 627–640. (doi:10.1016/j.jasrep.2016.04.005)
31. Robson HK, Saul H, Steele VJ, Meadows J, Otto Nielsen P, Fischer A, Heron CP, Craig OE. 2021 Organic residue analysis of early neolithic 'bog pots' from Denmark demonstrates the processing of wild and domestic foodstuffs. *J. Archaeol. Sci.* **36**, 102829. (doi:10.1016/j.jasrep.2021.102829)
32. Craig OE. 2004 Organic analysis of 'food crusts' from sites in the Schelde valley, Belgium: a preliminary evaluation. *Notae Praehist.* **24**, 209–217.
33. Heron C, Craig OE, Lucquin A, Steele VJ, Thompson A, Piličiauskas G. 2015 Cooking fish and drinking milk? Patterns in pottery use in the southeastern Baltic, 3300–2400 cal BC. *J. Archaeol. Sci.* **63**, 33–43. (doi:10.1016/j.jas.2015.08.002)
34. Gunnarssone A, Oras E, Talbot HM, Ilves K, Legzdina D. 2020 Cooking for the living and the dead: lipid analyses of Rauši settlement and cemetery pottery from the 11th–13th century. *Estonian J. Archaeol.* **24**, 45. (doi:10.3176/arch.2020.1.02)
35. Bondetti M *et al.* 2021 Neolithic farmers or neolithic foragers? Organic residue analysis of early pottery from Rakushechny Yar on the lower Don (Russia). *Archaeol. Anthropol. Sci.* **13**, 141. (doi:10.1007/s12520-021-01412-2)
36. Gibbs K *et al.* 2017 Exploring the emergence of an 'aquatic' neolithic in the Russian far east: organic residue analysis of early hunter-gatherer pottery from Sakhalin Island. *Antiquity* **91**, 1484–1500. (doi:10.15184/aqy.2017.183)
37. Bondetti M, Scott S, Lucquin A, Meadows J, Lozovskaya O, Dolbunova E, Jordan P, Craig OE. 2020 Fruits, fish and the introduction of pottery in the Eastern European plain: lipid residue analysis of ceramic vessels from Zamostje 2. *Quat. Int.* **541**, 104–114. (doi:10.1016/j.quaint.2019.05.008)
38. Lucquin A *et al.* 2023 The impact of farming on prehistoric culinary practices throughout Northern Europe. *Proc. Natl Acad. Sci. USA* **120**, e2310138120. (doi:10.1073/pnas.2310138120)
39. Taché K, Jaffe Y, Craig OE, Lucquin A, Zhou J, Wang H, Jiang S, Standall E, Flad RK. 2021 What do 'barbarians' eat? Integrating ceramic use-wear and residue analysis in the study of food and society at the margins of Bronze Age China. *PLoS One* **16**, e0250819. (doi:10.1371/journal.pone.0250819)
40. Shoda S, Lucquin A, Sou CI, Nishida Y, Sun G, Kitano H, Son J ho, Nakamura S, Craig OE. 2018 Molecular and isotopic evidence for the processing of starchy plants in early neolithic pottery from China. *Sci. Rep.* **8**, 17044. (doi:10.1038/s41598-018-35227-4)
41. Yoneda M, Kisida K, Gakuhari T, Omori T, Abe Y. 2019 Interpretation of bulk nitrogen and carbon isotopes in archaeological foodcrusts on potsherds. *Rapid Commun. Mass Spectrom.* **33**, 1097–1106. (doi:10.1002/rcm.8446)
42. Craig OE *et al.* 2013 Earliest evidence for the use of pottery. *Nature* **496**, 351–354. (doi:10.1038/nature12109)
43. Shoda S, Lucquin A, Ahn J ho, Hwang C joo, Craig OE. 2017 Pottery use by early Holocene hunter-gatherers of the Korean peninsula closely linked with the exploitation of marine resources. *Quat. Sci. Rev.* **170**, 164–173. (doi:10.1016/j.quascirev.2017.06.032)
44. Admiraal M *et al.* 2023 Chemical analysis of pottery reveals the transition from a maritime to a plant-based economy in pre-colonial coastal Brazil. *Sci. Rep.* **13**, 16771. (doi:10.1038/s41598-023-42662-5)
45. Robson HK *et al.* 2022 Light production by ceramic using hunter-gatherer-fishers of the circum-Baltic. *Proc. Prehist. Soc.* **88**, 25–52. (doi:10.1017/ppr.2022.12)
46. Oras E, Lucquin A, Lõugas L, Tõrv M, Kriiska A, Craig OE. 2017 The adoption of pottery by north-east European hunter-gatherers: evidence from lipid residue analysis. *J. Archaeol. Sci.* **78**, 112–119. (doi:10.1016/j.jas.2016.11.010)
47. Regert M, Vacher S, Moulherat C, Decavallas O. 2003 Adhesive production and pottery function during the Iron Age at the site of Grand Unay (Sarthe, France). *Archaeometry* **45**, 101–120. (doi:10.1111/1475-4754.00098)
48. Lyu N, Yan L, Wang T, Lin L, Rao H, Yang Y. 2022 Characterization of pottery foodcrusts through lipid and proteomic analyses: a case study from the Xiawan site in Yixing city, East China. *J. Archaeol. Sci.* (doi:10.2139/ssrn.4216384)
49. Miller MJ *et al.* 2020 Interpreting ancient food practices: stable isotope and molecular analyses of visible and absorbed residues from a year-long cooking experiment. *Sci. Rep.* **10**, 13704. (doi:10.1038/s41598-020-70109-8)
50. Evershed RP, Copley MS, Dickson L, Hansel FA. 2008 Experimental evidence for the processing of marine animal products and other commodities containing polyunsaturated fatty acids in pottery vessels. *Archaeometry* **50**, 101–113. (doi:10.1111/j.1475-4754.2007.00368.x)
51. Charters S, Evershed RP, Quye A, Blinkhorn PW, Reeves V. 1997 Simulation experiments for determining the use of ancient pottery vessels: the behaviour of epicuticular leaf wax during boiling of a leafy vegetable. *J. Archaeol. Sci.* **24**, 1–7. (doi:10.1006/jasc.1995.0091)
52. Bondetti M, Scott E, Courel B, Lucquin A, Shoda S, Lundy J, Labra-Odde C, Drieu L, Craig OE. 2021 Investigating the formation and diagnostic value of ω -(*o*-alkylphenyl)alkanoic acids in ancient pottery. *Archaeometry* **63**, 594–608. (doi:10.1111/arcm.12631)

53. Liu X, Ren M, Fu Y, Hu Y, Wang S, Yang Y. 2023 New insights into the use of neolithic pottery in Guangxi of South China: organic residue analysis of experimental and archaeological pottery. *Herit. Sci.* **11**, 1–15. (doi:10.1186/s40494-023-01045-9)
54. Hammann S, Cramp LJE. 2018 Towards the detection of dietary cereal processing through absorbed lipid biomarkers in archaeological pottery. *J. Archaeol. Sci.* **93**, 74–81. (doi:10.1016/j.jas.2018.02.017)
55. Drieu L, Regert M, Mazuy A, Vieugué J, Bocoum H, Mayor A. 2022 Relationships between lipid profiles and use of ethnographic pottery: an exploratory study. *J. Archaeol. Method Theory* **29**, 1294–1322. (doi:10.1007/s10816-021-09547-1)
56. Barker A, Venables B, Stevens SM, Seeley KW, Wang P, Wolverson S. 2012 An optimized approach for protein residue extraction and identification from ceramics after cooking. *J. Archaeol. Method Theory* **19**, 407–439. (doi:10.1007/s10816-011-9120-5)
57. Evershed RP, Tuross N. 1996 Proteinaceous material from potsherds and associated soils. *J. Archaeol. Sci.* **23**, 429–436. (doi:10.1006/jasc.1996.0038)
58. Hughes CS, Foehr S, Garfield DA, Furlong EE, Steinmetz LM, Krijgsveld J. 2014 Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol. Syst. Biol.* **10**, 757. (doi:10.15252/msb.20145625)
59. Hughes CS, Moggridge S, Müller T, Sorensen PH, Morin GB, Krijgsveld J. 2019 Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat. Protoc.* **14**, 68–85. (doi:10.1038/s41596-018-0082-x)
60. Cleland TP. 2018 Human bone paleoproteomics utilizing the single-pot, solid-phase-enhanced sample preparation method to maximize detected proteins and reduce humics. *J. Proteome Res.* **17**, 3976–3983. (doi:10.1021/acs.jproteome.8b00637)
61. Bleasdale M *et al.* 2021 Ancient proteins provide evidence of dairy consumption in eastern Africa. *Nat. Commun.* **12**, 632. (doi:10.1038/s41467-020-20682-3)
62. Ventresca Miller AR *et al.* 2022 The spread of herds and horses into the Altai: how livestock and dairying drove social complexity in Mongolia. *PLoS One* **17**, e0265775. (doi:10.1371/journal.pone.0265775)
63. Sakalauskaite J, Marin F, Pergolizzi B, Demarchi B. 2020 Shell palaeoproteomics: first application of peptide mass fingerprinting for the rapid identification of mollusc shells in archaeology. *J. Proteomics* **227**, 103920. (doi:10.1016/j.jprot.2020.103920)
64. Lim MY, Paulo JA, Gygi SP. 2019 Evaluating false transfer rates from the match-between-runs algorithm with a two-proteome model. *J. Proteome Res.* **18**, 4020–4026. (doi:10.1021/acs.jproteome.9b00492)
65. Conway JR, Lex A, Gehlenborg N. 2017 UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33**, 2938–2940. (doi:10.1093/bioinformatics/btx364)
66. Lex A, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H. 2014 UpSet: visualization of intersecting sets. *IEEE Trans. Vis. Comput. Graph.* **20**, 1983–1992. (doi:10.1109/TVCG.2014.2346248)
67. Sharpe JR, Sammons RL, Marquis PM. 1997 Effect of pH on protein adsorption to hydroxyapatite and tricalcium phosphate ceramics. *Biomaterials* **18**, 471–476. (doi:10.1016/S0142-9612(96)00157-3)
68. Stevens SM, Wolverson S, Venables B, Barker A, Seeley KW, Adhikari P. 2010 Evaluation of microwave-assisted enzymatic digestion and tandem mass spectrometry for the identification of protein residues from an inorganic solid matrix: implications in archaeological research. *Anal. Bioanal. Chem.* **396**, 1491–1499. (doi:10.1007/s00216-009-3341-4)
69. Collins MJ, Bishop AN, Farrimond P. 1995 Sorption by mineral surfaces: rebirth of the classical condensation pathway for kerogen formation? *Geochim. Cosmochim. Acta* **59**, 2387–2391. (doi:10.1016/0016-7037(95)00114-F)
70. Jung F, Frey K, Zimmer D, Mühlhaus T. 2023 DeepSTABp: a deep learning approach for the prediction of thermal protein stability. *Int. J. Mol. Sci.* **24**, 7444. (doi:10.3390/ijms24087444)
71. Charoenkwan P, Ahmed S, Nantasenamat C, Quinn JMW, Moni MA, Lio' P, Shoombuatong W. 2022 AMYPred-FRL is a novel approach for accurate prediction of amyloid proteins by using feature representation learning. *Sci. Rep.* **12**, 7697. (doi:10.1038/s41598-022-11897-z)
72. Erdős G, Dosztányi Z. 2020 Analyzing protein disorder with iupred2a. *Curr. Protoc. Bioinformatics* **70**, e99. (doi:10.1002/cpbi.99)
73. Kozłowski LP. 2016 IPC: isoelectric point calculator. *Biol. Direct* **11**, 55. (doi:10.1186/s13062-016-0159-9)
74. Wang S, Li W, Liu S, Xu J. 2016 RaptorX-property: a web server for protein structure property prediction. *Nucleic Acids Res.* **44**, W430–W435. (doi:10.1093/nar/gkw306)
75. Wang S, Peng J, Ma J, Xu J. 2016 Protein secondary structure prediction using deep convolutional neural fields. *Sci. Rep.* **6**, 18962. (doi:10.1038/srep18962)
76. Buric F, Viknander S, Fu X, Lemke O, Zrimec J, Szyrwiel L, Mueleder M, Ralsler M, Zelezniak A. 2023 The amino acid sequence determines protein abundance through its conformational stability and reduced synthesis cost. *bioRxiv*. (doi:10.1101/2023.10.02.560091)
77. Pollastri G, Przybylski D, Rost B, Baldi P. 2002 Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles. *Proteins* **47**, 228–235. (doi:10.1002/prot.10082)
78. Demarchi B *et al.* 2016 Protein sequences bound to mineral surfaces persist into deep time. *eLife* **5**, e17092. (doi:10.7554/eLife.17092)
79. Scott A. 2022 Investigation of ancient proteins in archaeological material. PhD thesis, Friedrich Schiller University, Jena, Germany.
80. Zhang Y, Che H, Li C, Jin T. 2023 Food allergens of plant origin. *Foods* **12**, 2232. (doi:10.3390/foods12112232)
81. Scott A *et al.* 2021 Exotic foods reveal contact between South Asia and the Near East during the second millennium BCE. *Proc. Natl Acad. Sci. USA* **118**, e2014956117. (doi:10.1073/pnas.2014956117)
82. Hilscher A, Heister K, Siewert K, Knicker H. 2009 Mineralisation and structural changes during the initial phase of microbial degradation of pyrogenic plant residues in soil. *Org. Geochem.* **40**, 332–342. (doi:10.1016/j.orggeochem.2008.12.004)
83. Knicker H, Río JC del, Hatcher PG, Minard RD. 2001 Identification of protein remnants in insoluble geopolymers using TMAH thermochemolysis/GC–MS. *Org. Geochem.* **32**, 397–409. (doi:10.1016/S0146-6380(00)00186-8)

84. Knicker H. 2004 Stabilization of N-compounds in soil and organic-matter-rich sediments—what is the difference? *Mar. Chem.* **92**, 167–195. (doi:10.1016/j.marchem.2004.06.025)
85. Knicker H, Hatcher PG. 1997 Survival of protein in an organic-rich sediment: possible protection by encapsulation in organic matter. *Naturwissenschaften* **84**, 231–234. (doi:10.1007/s001140050384)
86. Panettieri M, Knicker H, Murillo JM, Madejón E, Hatcher PG. 2014 Soil organic matter degradation in an agricultural chronosequence under different tillage regimes evaluated by organic matter pools, enzymatic activities and CPMAS ¹³C NMR. *Soil Biol. Biochem.* **78**, 170–181. (doi:10.1016/j.soilbio.2014.07.021)
87. del Rio JC, Olivella MA, Knicker H, de las Heras FXC. 2004 Preservation of peptide moieties in three Spanish sulfur-rich tertiary kerogens. *Org. Geochem.* **35**, 993–999. (doi:10.1016/j.orggeochem.2004.06.002)
88. Taurozzi AJ et al. 2024 Deep-time phylogenetic inference by paleoproteomic analysis of dental enamel. *Nat. Protoc.* **19**, 2085–2116. (doi:10.1038/s41596-024-00975-3)
89. Ramsøe A, Heekeren V, Ponce P, Fischer R, Barnes I, Speller C, Collins MJ. 2020 DeamiDATE 1.0: site-specific deamidation as a tool to assess authenticity of members of ancient proteomes. *J. Archaeol. Sci.* **115**, 105080. (doi:10.1016/j.jas.2020.105080)
90. Tanriver G, Monard G, Catak S. 2022 Impact of deamidation on the structure and function of antiapoptotic Bcl-xL. *J. Chem. Inf. Model.* **62**, 102–115. (doi:10.1021/acs.jcim.1c00808)
91. Griffin RC, Penkman KEH, Moody H, Collins MJ. 2010 The impact of random natural variability on aspartic acid racemization ratios in enamel from different types of human teeth. *Forensic Sci. Int.* **200**, 148–152. (doi:10.1016/j.forsciint.2010.04.005)
92. Cao B, Porollo A, Adamczak R, Jarrell M, Meller J. 2006 Enhanced recognition of protein transmembrane domains with prediction-based structural profiles. *Bioinformatics* **22**, 303–309. (doi:10.1093/bioinformatics/bti784)
93. Siano F, Picariello G, Caruso T, Esposito S, Rescigno C, Addeo F, Vasca E. 2022 Proteomics and integrated techniques to characterize organic residues in funerary findings from Italic populations of the first millennium BC. *J. Proteome Res.* **21**, 1330–1339. (doi:10.1021/acs.jproteome.2c00093)
94. Yorkshire Water. 2023 *York East 2019 (9809519) water supply zone*. See [https://www.ywsonline.co.uk/web/WQZ.nsf/0/F0638E59C295EEA280258A36003EDA17/\\$file/York%20East%202019%20WSZ.pdf](https://www.ywsonline.co.uk/web/WQZ.nsf/0/F0638E59C295EEA280258A36003EDA17/$file/York%20East%202019%20WSZ.pdf) (accessed 12 July 2024).
95. Barnard H et al. 2007 Mixed results of seven methods for organic residue analysis applied to one vessel with the residue of a known foodstuff. *J. Archaeol. Sci.* **34**, 28–37. (doi:10.1016/j.jas.2006.03.010)
96. Heaton K, Solazzo C, Collins MJ, Thomas-Oates J, Bergström ET. 2009 Towards the application of desorption electrospray ionisation mass spectrometry (DESI-MS) to the analysis of ancient proteins from artefacts. *J. Archaeol. Sci.* **36**, 2145–2154. (doi:10.1016/j.jas.2009.05.016)
97. Perez-Riverol Y et al. 2022 The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552. (doi:10.1093/nar/gkab1038)
98. Evans M, Hagan R, Boyd O, Bondetti M, Craig OE, Collins M. 2024 Data from: The impact of cooking and burial on proteins: a characterisation of experimental foodcrusts and ceramics. Figshare. (doi:10.6084/m9.figshare.c.7431942)