Paleoproteomic profiling of organic residues on prehistoric pottery from Malta

Davide Tanasi[1] · Annamaria Cucina[2] · Vincenzo Cunsolo[2] · Rosaria Saletti[2] · Antonella Di Francesco[2] · Enrico Greco[3] · Salvatore Foti[2]

Abstract

Mass spectrometry-based approaches have been successfully applied for identifying ancient proteins in bones and other tissues. On the contrary, there are relatively few examples of the successful recovery and identification of archeological protein residues from ceramic artifacts; this is because ceramics contain much lower levels of proteins which are extensively degraded by diagenetic effects. In this paper, we report the results of the characterization of proteins extracted from pottery of the Maltese site of Baħrija, the guide-site for the Baħrija period (half of 9th–second half of eighth century BCE), recently identified as the final part of the Borġ in-Nadur culture. Proteomic data here reported confirm that one of the major issue of these kind of studies is represented by contamination of animal and human agents that may complicate endogenous protein identification and authentication. The samples tested included a small group of ceramic forms, namely three tableware and six coarse ware thought to have been used in food preparation and/or storage. In this context, the limited availability of paleobotanical and archeozoological analyses may be compensated by the outcomes of the first proteomics profiling which, even if obtained on a limited selection of vessels, revealed the centrality of wheat in the diet of the ancient community of Baħrija. The data have been deposited to the ProteomeXchange with identifier < PXD022848 > .

Introduction

Over the past three decades, studies of ancient biomolecules, in particular ancient DNA (aDNA), proteins, and lipids, allowed to shed light on the biological past, improving our understanding of evolutionary history (Cappellini et al., 2018). Particularly, the interest in the study of paleodiets has grown considerably in tandem with the use of technologies and analytical tools that have made it possible to identify the organic residues more precisely. The use of multiple techniques to validate scientific results has been repeatedly used to identify animal products (Mottram et al. 1999; Spangenberg et al. 2006; Spiteri et al. 2020), milk (Dudd et al. 1998; Evershed et al. 2002) and vegetable lipids (Colombini et al. 2005), fish and shellfish (Hansel et al. 2004) or to check the presence of fermented drinks such as grape wine (Cavalieri et al. 2003; Guasch-Jané et al. 2004), beer (Borek et al. 2007), or other beverages (McGovern et al. 2013). Studies on plant microremains and aDNA trapped in ancient human dental calculus have also been carried out, providing new insights about past human diet and health (Weyrich et al. 2015; Leonard et al. 2015; Søe et al. 2018; Gismondi et al. 2020). On this respect, dental calculus, a mineralized form of dental plaque, represents a long-term reservoir of dietary biomolecules and microfossils, offering a unique opportunity to access primary evidence of ancient diets at an individual level (Hendy et al. 2018a; Velsko et al. 2017; Warinner et al. 2014, 2015).

✉ Vincenzo Cunsolo
vcunsolo@unict.it

1   Department of History, University of South Florida, SOC107 4202 E. Fowler Ave, Tampa, FL 33620, USA

2   Laboratory of Organic Mass Spectrometry, Department of Chemical Sciences, University of Catania, Viale A. Doria 6, 95125 Catania, Italy

3   Aix-Marseille Université, Institut de Chimie Radicalaire, Service 512, Avenue Escadrille Normandie Niemen, 13013 Marseille, France

Proteomics has been widely applied to detect and identify proteins on paleontological and archaeological artefacts. The first research on such interdisciplinary approach dates to early 2000 (Ostrom et al. 2000; Tokarski et al. 2003) and, since then, investigations on ancient samples are steadily increasing (Cappellini et al. 2014a; Dallongeville et al. 2016). Such growing popularity is undoubtedly determined by the significant technological development, in terms of performance and versatility of the MS instrumentation which nowadays represents an indispensable tool in paleoproteomics (Cleland et al. 2018). This has facilitated the phylogenetic identification of extant and extinct species (Cappellini et al. 2014b; Welker et al., 2015), studies of diagenetic and in vivo protein post-translational modifications (PTMs) (Cleland et al. 2015; Mikšík et al. 2016), the characterization of past human diseases (Hendy et al. 2016; D'Amato et al. 2018), and the reconstruction of the human diet (Warinner et al. 2014; Yang et al 2014; Shevchenko et al. 2014; Xie et al. 2016; Greco et al. 2018; Charlton et al. 2019). In comparison with the apparent widespread success of mass spectrometry approaches for identifying ancient proteins in bones and other tissues (Dallongeville et al. 2016), there are relatively few examples of the successful recovery and identification of archeological protein residues from ceramic artifacts (Solazzo et al. 2008; Dallongeville et al. 2011; Buckley et al. 2013; Salque et al. 2013; Shevchenko et al. 2018; Hendy et al. 2018b). Mostly because ceramics contain a much lower amount of proteinaceous material. The above mentioned studies also highlight how challenges routinely come from the heterogenous and unusual physical state of ancient samples and from the high level of degradation (Vinciguerra et al. 2016; Barker et al. 2018). The other major issue is represented by contamination of animal and human agents which could have occurred at pre-depositional (the contamination that a vessel sustains during its 'life') depositional (the contamination it sustains when it becomes part of an archeological context) and post-depositional stage (the contamination it sustains from the time it is unearthed to that when its sample reaches a proteomic laboratory) (Skibo et al. 2015). Concerning contamination, the challenge of discerning contaminant proteins from endogenous ancient molecules becomes particularly problematic. Authentication and validation criteria are needed to discriminate ancient endogenous proteins from contaminants. Deamidation process of asparagine and mainly of glutamine residues is the proposed marker of age that has recently been used in many archeological and paleontological studies (Robinson et al. 2001; Robinson 2002; Wilson et al. 2012; Van Doorn et al. 2012). With this respect, taking into account the complexity and the peculiarity of archeological samples and also their inestimable value, during the last decade

conventional proteomics protocols have been optimized and a series of precautions, from sample selection to data interpretation, have been successfully set, especially for bone remains (Hendy et al. 2018c). On the other hand, the studies on protein residues from ceramics highlighted that diagenesis deeply affects the protein sequences. Diagenesis, including a complex network of reactions such as chemical degradation and molecular breakdown driven by environmental factors during burial and storage, results in protein modifications beyond those produced in vivo and introduces a new challenge for protein identification and authentication (Cleland et al. 2020). Some substrates (e.g., bone, dental calculus, and eggshell) may harbor a better potential for preserving endogenous proteins than other ones. On the contrary, on ancient pottery, the microbial attack of the diagenetic action are more extensive because of the poorer screen effect of ceramics compared to the hydroxyapatite cage that protects bone proteins. However, it should be noted that calcified deposits lining the inner surfaces of archeological vessels may facilitate long-term archeological preservation of proteins and peptides (Hendy et al. 2018b). Detection of surviving proteins is also hampered by their hydrophilic properties (compared to lipids, for example), solubility in water, and the matrix effect of humics. Consequently, the identification of proteins in pottery is extremely challenging compared to other archaeological samples. It should be highlighted that another challenge, for which it does not exist any definitive solution, is related to the reuse of ancient pottery. Indeed, pottery was often subject to reuse for different food preparation purposes, leaving behind a complex and stratified combination of protein traces. However, information derived from an archeological context, a typological study of the pottery, analysis of written and iconographic sources, and ethno-archeological research allow us to significantly narrow down the use or number of uses done for certain vessels (Skibo et al. 2012). Altogether, these aspects highlight the challenges of paleoproteomic studies, such as the characterization of ancient proteins absorbed by archeological pottery. Continuous efforts are aimed to improve experimental and bioinformatic strategies for characterizing protein degradation and contamination, as well as for protein validation and authentication.

Within this research line, this paper reports the results of the first-ever characterization of protein extracted from prehistoric pottery of the Maltese Bronze/Iron Age site of Baħrija, via a shotgun MS-based approach.

Materials and methods

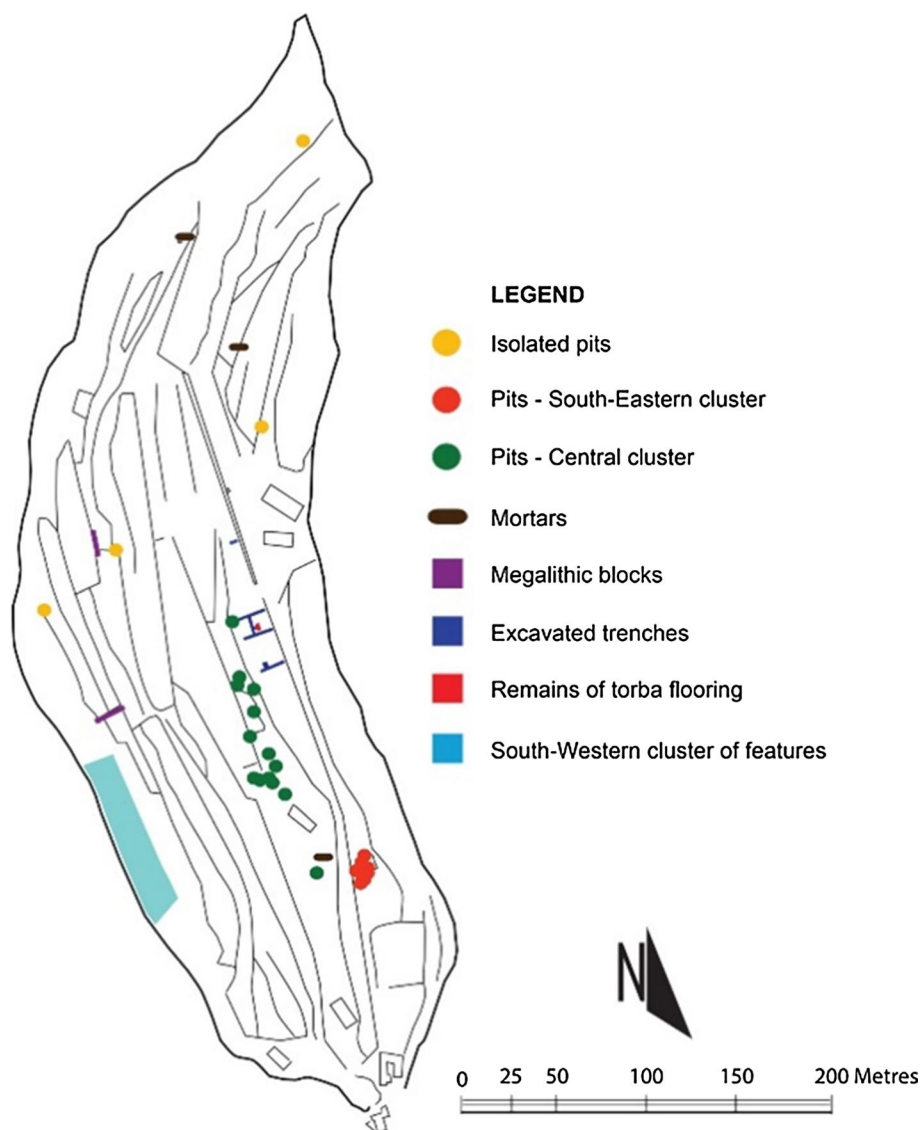### Case study: the iron age settlement at Qlejgħa tal-Baħrija (Malta)

Il-Qlejgħa tal-Baħrija is an Upper Coralline Limestone plateau on the west coast of Malta. Flanked on the eastern side by some of the most fertile lands in the area and offering the strategic advantage and the natural defense of the high ground, the Qlejgħa tal-Baħrija hilltop attracted the attention of human activity in prehistory when a settlement was established on it (Fig. 1). The first archeological exploration of the site was carried out by T.E. Peet in June 1909 (Peet 1910) with a series of long and narrow trenches which intercepted, at about 30–50 cm of depth, a gray layer of soil characterized by abundant pottery and bones (layer 5) and related with at least two structures likely to be interpreted as wattle and daub huts (in trench M and

trench G) (Cardona 2020). Among the materials from layer 5, the vast majority of the vessels were tableware (cups, bowl, jugs) and coarse ware (storage jars) of medium–large size (50–100 cm in height), mostly decorated with incisions and impressions typical of the Baħrija style which emerges in the final stage of the Maltese Bronze Age. Recent studies have characterized such stage as a 'Baħrija period', assignable to mid-9th to mid-eighth century BCE, corresponding to the latest development of the Borġ in-Nadur culture (Tanasi 2020). In 1959, further excavations conducted by D.H. Trump, not far from the area previously investigated, uncovered several other significant features on the plateau and confirmed the evidence of a well-organized settlement (Fig. 2) (Trump 1961). After 1959, no further excavations were ever carried out in the site. However, the publication of an overall reappraisal of the legacy excavation data (Tanasi and Cardona 2020) has now shown evidence of a first Maltese proto-urban society, able to establish new connections with Sicilian and

Fig. 1 Contour Map of il-Qlejgħa tal-Baħrija area, NW Malta (after Cardona and Zammit 2015)
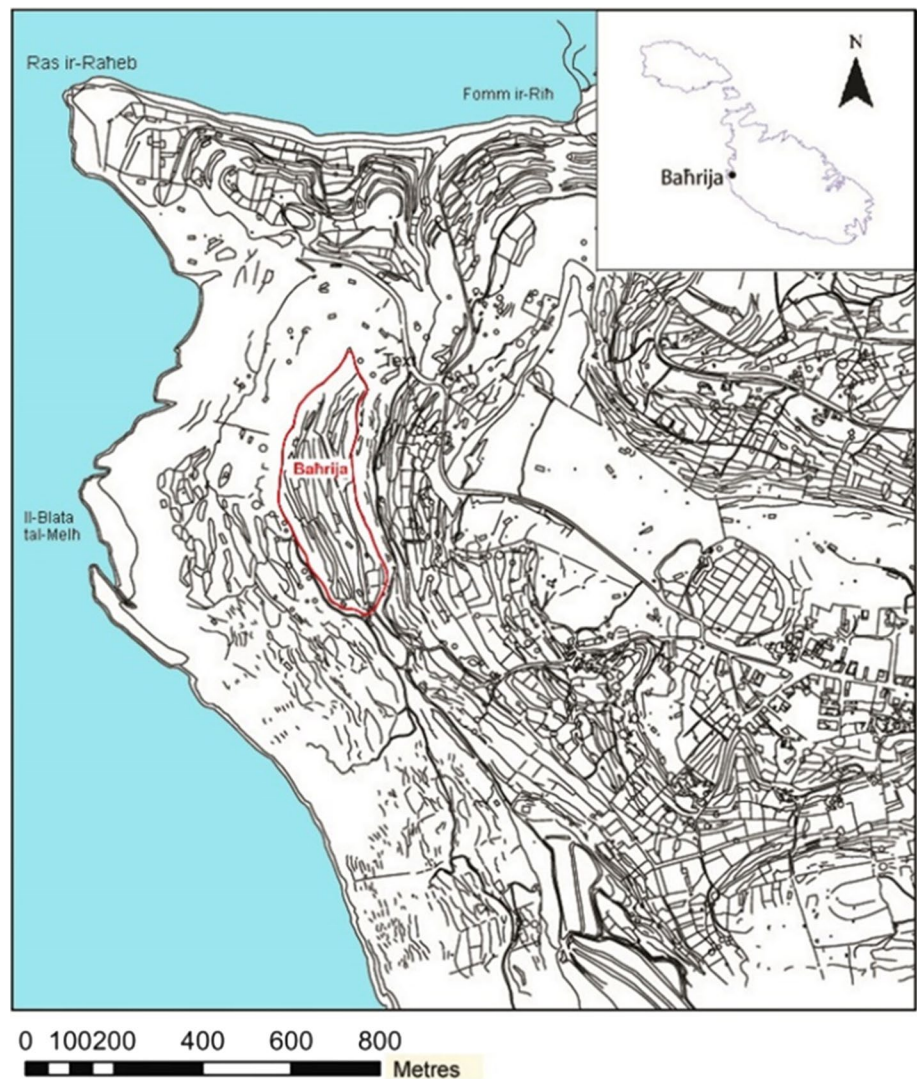
Fig. 2 General site plan of il-Qlejgħa tal-Baħrija showing all the known archeological features (Cardona and Zammit 2015)

LEGEND

🟡 Isolated pits

🔴 Pits - South-Eastern cluster

🟢 Pits - Central cluster

▬ Mortars

🟪 Megalithic blocks

🟦 Excavated trenches

🟥 Remains of torba flooring

🟦 South-Western cluster of features

Aegean communities and exploiting the strategic geography of Malta, at the center of the Mediterranean basin, to influence major historical events before the Greek and Phoenician colonizations BCE (Bonanno 2017).

Chemicals

The chemicals employed during the course of the analysis were of the highest purity commercially available and used without further purification. Tris–HCl was purchased from Carlo Erba (Milan, Italy); formic acid (FA), ammonium bicarbonate, dithiothreitol (DTT), iodoacetamide (IAA) from Aldrich (St. Louis, Missouri, USA), Trifluoroacetic acid (TFA) from Riedel-de Haën; modified porcine trypsin and sodium dodecyl sulfate (SDS) from Promega (Madison, WI, USA); water and acetonitrile (ACN) (OPTIMA® LC/MS grade) for LC/MS analyses from Fisher Scientific

(Milan, Italy). All the chemicals listed above were exclusively employed for the present study.

Sample collection

In consideration of the historical importance of the site of Qlejgħa tal-Baħrija, it becomes crucial to determine whether at the site there was a practice of cereal storage, which in the contemporaneous Sicilian society was an indication of advanced food redistribution policies (Albanese Procelli 2003). To investigate the foodways of the prehistoric community of Qlejgħa tal-Baħrija, in absence of paleobotanical and archeozoological data, and to determine the eventual use for cereal storage of the numerous large storage jars found on site, it was decided to sample specific shapes from the tableware and coarse ware repertoire to perform chemical analyses and explore the possible presence of organic

residues. To this extent, in summer 2017, nine samples (three of tableware and six of coarse ware) were extracted from vessels related to the Late Borġ in-Nadur style and Baħrija style from the Peet excavation (Table 1).

During the Peet excavation in 1909, the materials retrieved were put in large wooden boxes without being previously cleaned or subject to any treatment. After 1958, those same boxes were placed in the store-room of the National Museum of Archaeology of Valletta. No record

Table 1 List of pottery remains analyzed in the present work

Samples	Pottery class	Chronological phase	Hypothesized function	Proteomic Analysis Results by PEAKS (Identified proteins)	
				Contaminants	Food-derived Proteins
100041M	Tableware: cup with inverted rim and carinated body, decorated with impressed linear and geometric motifs	Baħrija (11th -8th century BC)	Used for liquid or semi-liquid food of vegetable or animal origin	Human Keratin and skin-derived proteins (a total of 19 proteins)	Milk-derived proteins (1 protein)
100051A	Coarse ware: strainer vessel with cylindrical body.	Baħrija (11th -8th century BC)	Sieve vessel traditionally interpreted as related to cheese making	-	Milk-derived proteins (2)
100051D	Tableware: strainer spouted jug with a red slipped surface.	Baħrija (11th -8th century BC)	Used for liquid with a residue that needs to be filtered	Human Keratin and skin-derived proteins (30)	Milk-derived proteins (4)
100061B	Tableware: cup with inverted rim and couple of lug handle and dark red/brown slipped lustrous surfaces.	Baħrija (11th -8th century BC)	Used for liquid or semi-liquid food of vegetable or animal origin	Human Keratin and skin-derived proteins (2)	-
100061H	Coarse ware: medium size storage jar decorated with rope bands.	Late Borġ in-Nadur/Baħrija (12th-11th century BCE)	Used for liquid or semi-liquid food of vegetable or animal origin	Human Keratin and skin-derived proteins (31)	-
100087C	Coarse ware: large size storage jar decorated with rope bands.	Late Borġ in-Nadur/Baħrija (12th-11th century BCE)	Used for the storage of oil, wine or cereals	-	Wheat-derived proteins (5)
100087D	Coarse ware large size storage jar with dark red slipped surfaces.	Late Borġ in-Nadur/Baħrija (12th-11th century BCE)	Used for the storage of oil, wine or cereals	Human skin-derived protein (1)	Wheat-derived proteins (4)
100088B	Coarse ware large size storage jar with red slipped surfaces	Late Borġ in-Nadur (12th century BCE)	Used for the storage of oil, wine or cereals	Human Keratin and skin-derived proteins (10)	Cereals-derived proteins (32)
100088D	Coarse ware: large size storage jar decorated with rope bands	Late Borġ in-Nadur/Baħrija (12th-11th century BCE)	Used for the storage of oil, wine or cereals	Human Keratin and skin-derived proteins (4)	Wheat-derived proteins (16)

The list of food-derived proteins identified by the MS-based proteomic analysis is reported in the Table 2. The complete set of proteins, including potential contaminants, is reported in the Supplementary Table S1

Table 2 List of the food-derived proteins identified by PEAKS search in the archeological samples here investigated

Sample	Proteins	UniProt ID	Taxa	InChorus Score (%)	Coverage (%)	Peptides	Unique	
100041 M **100051A**	Alpha-S1-casein	P02662	*Bos taurus*	97.97	10	3	3	
	Beta-lactoglobulin	P02754	*Bos taurus*	83.68	8	2	2	Multispecies identification: Bubalus bubalis
	Beta-casein	P02666	*Bos taurus*	76.82	12	2	2	Multispecies identification: Bubalus bubalis
100051D	Kappa-casein	P02668	*Bos taurus*	98.24	20	3	3	
	Alpha-S1-casein	P02662	*Bos taurus*	97.35	21	6	6	
	Alpha-S2-casein	P02663	*Bos taurus*	83.61	15	2	2	Multispecies identification: Ovis aries
	Beta-casein	P02666	*Bos taurus*	79.64	12	3	3	Multispecies identification: Bubalus bubalis
100061B	No Id							
100061H	No Id							
100087C	Alpha-amylase/trypsin inhibitor CM3	P17314	*Triticum aestivum*	98.35	25	2	2	
	Purothionin A-1	P01543	*Triticum aestivum*	98.29	26	4	1	
	Non-specific lipid-transfer protein (Fragment)	P24296	*Triticum aestivum*	98.25	45	3	3	
	Alpha-1-purothionin (Fragment)	P01544	*Triticum aestivum*	98.24	27	3	2	
	Puroindoline-B	Q10464	*Triticum aestivum*	75.95	18	2	2	
100087D	Alpha-amylase/trypsin inhibitor CM3	P17314	*Triticum aestivum*	98.41	40	3	3	
	Alpha-amylase/trypsin inhibitor CM16	P16159	*Triticum aestivum*	98.07	13	2	2	
	Alpha-amylase inhibitor 0.19	P01085	*Triticum aestivum*	82.99	33	2	2	
	Chymotrypsin inhibitor WCI	P83207	*Triticum aestivum*	58.51	25	2	2	
100088B	Alpha-amylase inhibitor 0.28	P01083	*Triticum aestivum*	99.18	69	10	10	
	Glutenin high molecular weight subunit DY10	P10387	*Triticum aestivum*	99.16	23	7	7	
	Glutenin high molecular weight subunit DX5	P10388	*Triticum aestivum*	99.13	12	7	7	
	Serpin-Z2B	P93692	*Triticum aestivum*	99.13	39	8	7	
	Alpha-amylase/trypsin inhibitor CM3	P17314	*Triticum aestivum*	99.09	49	8	7	
	Beta-amylase	P93594	*Triticum aestivum*	99.05	23	7	7	Multispecies identification: Hordeum vulgare
	Alpha-amylase inhibitor 0.19	P01085	*Triticum aestivum*	99.04	78	10	10	
	Alpha-amylase/trypsin inhibitor CM16	P16159	*Triticum aestivum*	98.92	55	9	9	
	Alpha-1-purothionin	P01544	*Triticum aestivum*	98.84	38	4	2	

Table 2 (continued)

Sample	Proteins	UniProt ID	Taxa	InChorus Score (%)	Coverage (%)	Peptides	Unique	
	Alpha/beta-gliadin clone PW1215	P04726	*Triticum aestivum*	98.78	30	5	2	
	Alpha-amylase/trypsin inhibitor CM1	P16850	*Triticum aestivum*	98.75	39	5	2	
	Puroindoline-A	P33432	*Triticum aestivum*	98.73	45	7	7	
	Alpha-amylase/trypsin inhibitor CM2	P16851	*Triticum aestivum*	98.38	39	4	2	
	Purothionin A-1	P01543	*Triticum aestivum*	98.38	24	6	3	
	Aspartate aminotransferase, cytoplasmic	M7YLN2	*Triticum urartu*	98.35	13	3	3	Multispecies identification: Oryza sativa
	Gamma-gliadin B	P06659	*Triticum aestivum*	98.34	22	5	5	
	Avenin-like b3	P0CZ06	*Triticum aestivum*	98.30	16	4	2	
	Avenin-like b1	Q2A783	*Triticum aestivum*	98.30	16	4	2	
	Type-5 thionin	Q05806	*Triticum aestivum*	98.11	37	3	3	
	Chitin-binding type-1 domain-containing protein	A0A446YZN0	*Triticum durum*	97.96	25	3	2	Multispecies identification: Hordeum vulgare; Secale cereale
	Protein disulfide-isomerase	P52589	*Triticum aestivum*	95.93	8	3	3	Multispecies identification: Hordeum vulgare
	Puroindoline-B	Q10464	*Triticum aestivum*	94.84	32	6	6	
	Chymotrypsin inhibitor WCI	P83207	*Triticum aestivum*	92.04	25	2	2	
	Non-specific lipid-transfer protein	P24296	*Triticum aestivum*	88.01	41	2	2	
	Non-specific lipid-transfer protein 2G	P82900	*Triticum aestivum*	84.47	51	5	2	
	Trypsin/alpha-amylase inhibitor CMX1/CMX3	Q43723	*Triticum aestivum*	84.37	22	3	3	
	Triosephosphate isomerase cytosolic	P34937	*Hordeum vulgare*	84.01	10	2	2	Multispecies identification: Secale cereale
	Alpha/beta-gliadin A-V	P04725	*Triticum aestivum*	83.87	14	4	2	
	Wheatwin-1	O64392	*Triticum aestivum*	61.68	11	2	2	
	16.9 kDa class I heat shock protein 1	P12810	*Triticum aestivum*	61.68	13	2	2	
	Protein H2A.7	Q43312	*Triticum aestivum*	61.66	25	2	2	Multispecies identification: Oryza sativa; Mytilus edulis; etc.
	Carbonic anhydrase chloroplastic	P40880	*Hordeum vulgare*	60.88	9	2	2	
100088D	Alpha-amylase inhibitor 0.28	P01083	*Triticum aestivum*	98.21	27	3	3	
	Alpha-amylase inhibitor 0.19	P01085	*Triticum aestivum*	79.95	37	3	3	
	Purothionin A-1	P01543	*Triticum aestivum*	98.88	29	9	5	
	Alpha-1-purothionin	P01544	*Triticum aestivum*	98.84	40	7	4	
	Glutenin high molecular weight subunit DY10	P10387	*Triticum aestivum*	98.74	7	4	4	

Table 2 (continued)

Sample	Proteins	UniProt ID	Taxa	InChorus Score (%)	Coverage (%)	Peptides	Unique	
	Alpha-amylase/trypsin inhibitor CM16	P16159	*Triticum aestivum*	61.46	28	4	4	
	Alpha-amylase/trypsin inhibitor CM2	P16851	*Triticum aestivum*	83.47	52	6	6	
	Alpha-amylase/trypsin inhibitor CM3	P17314	*Triticum aestivum*	99.05	48	7	6	
	Non-specific lipid-transfer protein	P24296	*Triticum aestivum*	97.25	80	8	8	
	Alpha-2-purothionin	P32032	*Triticum aestivum*	98.86	40	9	5	
	Puroindoline-A	P33432	*Triticum aestivum*	99.00	55	9	9	
	Protein disulfide-isomerase	P52589	*Triticum aestivum*	97.82	8	4	4	
	Non-specific lipid-transfer protein 2G	P82900	*Triticum aestivum*	98.39	65	7	5	
	Beta-amylase	P93594	*Triticum aestivum*	98.10	8	4	4	Multispecies identification: Horedum vulgare
	Chymotrypsin inhibitor WCI	P83207	*Triticum aestivum*	91.58	39	4	4	
	Puroindoline-B	Q10464	*Triticum aestivum*	61.23	9	2	2	

Legend of each column:Sample: Number classification of the archeological pottery; Proteins: The protein's header information; UniProt Id: The UniProt Accession Number as reported in the database; Taxa: Organismal origin. Proteins from different species that are supported by a common set of peptides and could not be differentiated based on MS/MS analysis alone, are also reported as multispecies identification; Protein InChorus Score: The protein confidence score calculated as percentage score when the protein was identified by both PEAKS and Mascot engines (InChorus tool). The InChorus protein percentage score is calculated as the weighted sum of the protein's supporting peptides. The PeptideProphetTM method (Keller et al. Anal. Chem. 2002, 74:5383–92) is then applied to the weighted sum scores of all proteins to convert to a probability score. Coverage: The percentage of the protein sequence that is covered by supporting peptides. #Peptides: The number of high-confidence supporting peptides. #Unique: The number of high-confidence supporting peptides that are mapped to only one protein. Unique peptides with the same sequence but different modifications are only counted once in this number

of any study or treatment of the pottery between 1909 and 1958, and 1958 and 2017 exists in the archive of the Museum. Before sampling, dust was removed from the ceramic fragments using a brush with soft bristle. Encrustations present on the surface of all the samples here investigated were not removed. Nitrile gloves were used during this process and labeled Ziploc bags were employed to store the samples.

Sample preparation for proteomics analyses

The improvement in resolution and accuracy of analytical techniques has required a parallel enhancement in sample preparation procedures to avoid contamination with extraneous species or other molecules that could alter the correct interpretation of the data. Therefore, to minimize contamination, protein extraction and sample handling were performed in a laboratory "clean room" dedicated to ancient protein analysis and using dedicated chemicals, lab glassware, and equipment. Surfaces and equipment were washed with 50% 2-propanol before the use. Non-latex gloves were used.

The nine samples were initially weighed individually and where possible, from 0.1 to 0.5 g were taken. Some samples, which already had sufficient weight and were statistically significant (each piece has an entire section from the inner surface to the outer surface of the ceramic sample), were not further sampled. Others were cut with a micrometric saw with a nano-diamond surface coating (Struers Minitom) at minimum speed to avoid increasing the sample temperature and altering any residues that may be present. After being weighed, each sample was inserted in a glove box with nitrogen gas flow and was ground using an agate mortar and pestle until an impalpable powder was obtained. The samples were weighed one last time to ensure that the amount of powder was sufficient for the analyses and inserted in 1.5 mL Eppendorf. All the equipment was rigorously cleaned with 2-propanol and MilliQ® water between the preparation of one sample and the next. Protein extraction was performed using modified protocols previously reported (Dallongeville et al. 2011; Greco et al. 2018). In particular, to maximize

the protein extraction, the powder of each sample (between 30 and 200 mg) was divided into two aliquots and processed as follows:

- the first aliquot was treated in acidic conditions, pH = 2 (0.1% of trifluoroacetic acid, TFA, and 4% Sodium-dode-cyl-sulfate solution, SDS) for 30 min at 95 °C;
- the second one was used for the extraction of protein content in alkaline reducing conditions, pH = 8.5 (4% SDS, Tris/HCl 100 mM, and 0.1 M dithiothreitol, DTT) for 30 min at 95 °C.

The volume of the extraction solution used was 5 times (in μL) the amount of sample (in mg). Protein extracts were desalted and purified from non-protein contaminants using the PlusOne 2-D Clean-Up kit (GE Healthcare Life Sciences) according to the recommendations of the manufacturer (https://www.gelifesciences.com/en/ca). The obtained pellets were dissolved in 100 μL of 50 mM ammonium bicarbonate (pH 8.3) and DTT 0.1 M to reduce the proteins (3 h, room temperature). The protein concentration for each extract was determined by a fluorimetric assay using the Qubit Protein Assay kit with the Qubit 1.0 Fluorometer (ThermoFisher Scientific, Milan, Italy) (Saletti et al. 2018). Finally, samples were alkylated with 0,2 M IAA (1 h, in the dark at 20 °C) and digested by porcine trypsin (Sequencing Grade Modified Trypsin, Porcine, lyophilized, Promega) at an enzyme–substrate ratio of 1:50 (overnight, 37 °C). The resulting peptide mixture solutions were dried under vacuum (Concentrator Plus, Eppendorf), re-dissolved in 50 μL of 5% aqueous formic acid (FA), filtered by ultracentrifugation (750 μL, 0.2 μm Nonsterile Micro-Centrifugal Filters, Sepachrom, Rho, Milan), and analyzed by UHPLC/high-resolution nanoESI–MS/MS.

Two modern control samples (bovine milk and wheat flour) were also prepared (see Supporting Material). Modern control samples were processed and analyzed by proteomics on the same way of archeological samples.

Mass spectrometry analysis

Mass spectrometry data were acquired via a Thermo Fisher Scientific Orbitrap Fusion Tribrid® (Q-OT-qIT) mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). Liquid chromatography was carried out using a Thermo Scientific Dionex UltiMate 3000 RSLCnano system (Sunnyvale, CA). One microliter of peptide mixture was loaded onto an Acclaim ®Nano Trap C18 Column (100 μm i.d. × 2 cm, 5 μm particle size, 100 Å). After washing the trapping column with solvent A (H$_2$O + 0.1% FA) for 3 min at a flow rate of 7 μL/min, the peptides were eluted from the trapping column onto a PepMap® RSLC

C18 EASY-Spray column (75 μm i.d. × 50 cm, 2 μm particle size, 100 Å) and separated by elution at a flow rate of 0.25 μL/min at 40 °C by a linear gradient of solvent B (ACN + 0.1% FA) in A, 5% for 3 min, followed by 5% to 65% in 85 min, 65% to 95% in 5 min, finishing by holding 95% B 5 min, 95% to 5% in 10 min and re-equilibrating at 5% B for 15 min. The Acclaim ®Nano Trap C18 and Pep-Map® RSLC C18 EASY-Spray columns were dedicated to this set of samples and not previously used. The eluting peptide cations were converted to gas-phase ions by electrospray ionization using a source voltage of 1.75 kV and introduced into the mass spectrometer through a heated ion transfer tube (275 °C). Survey scans of peptide precursors from 200 to 1600 m/*z* were performed at 120 K resolution (@ 200 m/*z*). Tandem MS was performed by isolation at 1.6 Th with the quadrupole, HCD fragmentation with a normalized collision energy of 35, and rapid scan MS analysis in the ion trap (low-resolution MS/MS analysis). Only those precursors with charge state 2 ÷ 4 and intensity above the threshold of $1 \cdot 10^3$ were sampled for MS^2. The dynamic exclusion duration was set to 60 s with a 10 ppm tolerance around the selected precursor and its isotopes. Monoisotopic precursor selection was turned on. The instrument was run in full speed mode with 3 s cycles, meaning it would continuously perform MS^2 events until the list of non-excluded precursors diminished to zero or 3 s, whichever is shorter. MS/MS spectral quality was enhanced by enabling the parallelizable time option (i.e., using all parallelizable time during full scan detection for MS/MS precursor injection and detection). Mass spectrometer calibration was performed by using the Pierce® LTQ Velos ESI Positive Ion Calibration Solution (Thermo Fisher Scientific). MS data acquisition was carried out by utilizing the *Xcalibur* v. 3.0.63 software (Thermo Fisher Scientific). Moreover, a blank laboratory extraction was carried out and analyzed under the same experimental conditions used for ancient samples. Only a meager number of background proteins was identified by database search in blank laboratory extract (see Supplementary Material).

To avoid cross-contamination with other biological samples, all solvents were prepared freshly and ancient samples were not processed or analyzed in one batch with modern references. To avoid carryover during nLC–MS/MS runs, from three to five blank runs were performed before each analysis using the same gradient program. Spectra acquired in the last blank run were searched by PEAKS software against the SwissProt database without species origin restrictions and using the same parameters of the archeological samples (see next paragraph).

Database search and protein identification by PEAKS software

nLC–MS/MS data were analyzed and searched against the comprehensive (all species) SwissProt protein sequences database (January 2019 release, containing 558,475 entries) using integrated PEAKS De Novo sequencing software (v. 10.0, Bioinformatics Solutions Inc., Waterloo, ON Canada) and the Mascot algorithm (Matrix Science, London, UK, version 2.5.1). The first step of database search was carried out using the following parameters: (1) semi-tryptic peptides with a maximum of 3 missed cleavage sites; (2) cysteine carbamidomethylation as a fixed modification; (3) oxidation of methionine, the transformation of N-terminal glutamine and N-terminal glutamic acid residue to pyroglutamic acid form, and the deamidation of asparagine and glutamine as variable modifications. Particularly, the semi-tryptic option was chosen because it is well known that during the hydrolysis of proteins with trypsin, not only specific peptides (i.e., peptides resulting from hydrolysis at the carboxylic side of lysine and arginine) are formed, but also aspecific peptides (i.e., resulting from aspecific cleavages) have commonly been annotated. For instance, during hydrolysis by trypsin the aspecific cleavages after Tyr, Trp and Phe are the most observed (Burkhart et al. 2012; Perutka et al. 2018). On the other hand, the occurrence of spontaneous non-enzymatic peptide backbone cleavage is expected as the degradation effect in ancient samples. Then, based on previous observation of ancient proteome degradation (Cappellini et al. 2012), the unassigned "de novo only" spectra were examined by an additional search function of PEAKS, the Multi-Round Search, which allowed to examine the possible presence of post-translational modifications (unchecked in the first round of search) such as hydroxylation of proline, kynurenine formation, lysine, and arginine carbonylation.

The precursor mass tolerance threshold was set to 15 ppm and the maximum fragment mass error was set to 0.6 Da. Peptide spectral matches (PSM) were validated using a Target Decoy PSM Validator node based on q-values at a 1% False Discovery Rate (FDR). The Mascot score and PEAKS score thresholds for Peptide spectral matches (PSMs) were set to obtain for each database search FDR values, for PSMs, Peptide sequences, and Proteins identified, below the 1.0% value. Finally, all the protein hits obtained by these two approaches were processed by using the *inChorus* function of PEAKS. This tool combines the database search results of PEAKS software with those obtained by the Mascot search engine with the aim not only to increase the coverage but also the confidence since the engines use independent algorithms. Each peptide identified is reported with a percentage confidence score (*InChorus* score), reflecting the probability that the peptide–spectrum match is correct. The percentage score is calculated by the empirical calculation used in PeptideProphet™ (Keller et al. 2002). The protein score is also reported as a percentage score, calculated by adding the peptides score to identify a given protein up to a weighted sum. The PeptideProphet™ method is then applied to the weighted sum scores of all proteins to convert to a probability score. Only protein hits with a minimum of *InChorus* protein score of 50%, sequence coverage above 5% and at least two unique peptides, both having an individual *InChorus* score above the threshold of 50%, were considered valid. Proteins belonging to the same species identified by the same peptides and that could not be differentiated based on MS/MS analysis alone were grouped to satisfy parsimony principles (groups of parsimony). Proteins from different species identified by the same peptides were also grouped to satisfy parsimony principles and reported as multispecies identification (see Supplementary Tables).

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository (Vizcaino et al. 2013) with the dataset identifier <PXD022848>.

Calculation of deamidation level by MaxQuant software

MS raw data were also processed by MaxQuant (MQ) software 1.6.3.4 (https://www.maxquant.org/) to further investigate about deamidation level of the proteins identified (Cox et al. 2008).

In this search, a more specific database from Uniprot (Swiss-Prot and TrEMBL), which included the proteins identified in the PEAKS search, was used. A database including all the milk proteins from the *Bovidae* family and a database including all the proteins from *Triticum* was used for the potentially milk-containing samples (100041 M, 100051A, and 100051D) and wheat-containing samples, respectively (100087C, 100087D, 100088B, and 100088D). Moreover, the common Repository of Adventitious Proteins (c-RAP) contaminant database was enabled in the database search.

Database search was carried out using the following parameters: i) full tryptic peptides with a maximum of 3 missed cleavage sites; ii) cysteine carbamidomethylation as a fixed modification; iii) oxidation of methionine, the transformation of N-terminal glutamine and N-terminal glutamic acid residue to pyroglutamic acid form as variable modifications, hydroxyproline oxidation, and the deamidation of asparagine and glutamine as variable modifications. Match type was "match from and to". The decoy mode was "revert". PSM, Protein and Site decoy fraction FDR were set at 0.01 as threshold for peptide and protein identifications. Minimum score for modified and unmodified peptides was set at 40. All the other parameters were set as default. The principles of parsimony were still applied.

An estimation of the percentage of deamidation for each sample was calculated with the aid of a freely available command-line script for Python 2.x (https://github.com/dblyon/deamidation), which use the MaxQuant "evidence.txt" file. The calculations were done separately for potentially original peptides and potential contaminants peptides as previously reported in Mackie et al. 2018. Briefly, for each peptide containing N and Q residues, the ratio between the number of the residues in the modified form and the total number of the residues was calculated and this value was then multiplied for the intensity of the peptide. The values obtained were summed. The result was then divided by the total sum of all intensity values of the peptides in the modified and unmodified form. 1000 bootstrap iterations were applied to calculate mean, standard deviation, and 95% lower and upper confidence intervals and estimate the calculation error (see the Supplementary Material for details).

Results

Protein extracts from nine powdered samples were analyzed by a MS-based proteomic approach. A preliminary hypothesis about the use of these ancient pottery and the food contained inside was formulated based on archeological studies, and comparisons offered by the contemporaneous Sicilian and South-Italian repertoire of shapes (Recchia et al. 2000). The cups with inverted rims (samples 100041 M and 100061B), usually pedestalled, are considered as sort of serving dishes for semi-solid food more than drinking vessels, while for the strainer vessels there are different hypotheses based on their typological variations. The cylindrical strainer (100051A) has been traditionally interpreted, without any solid evidence though, as connected with cheese making due only to its resemblance to typical wicker cheese baskets of the rural culture. The strainer example (100051D), has instead been identified as a portion of a strainer spouted jug, a typical Sicilian, Aegean, and Levantine Late Bronze and Early Iron Age vessels, considered as a container of a beer-like fermented beverage (Leighton et al. 1981). The five examples of large storage jars (100061H, 100087C, 100087D, 100088B, and 100088D) are commonly interpreted as multi-purpose large size containers of liquids (such as olive oil and wine) or cereals (Tanasi 2020). At times, contents of intact examples can be suggested by the presence of a spout set by the base, used to empty and clean those filled with liquids, as it happens in culturally and chronologically compatible contexts in Sicily. In those specific cases, storage jars appear not to have been subject to reuse or change in the type of content (Albanese Procelli et al. 2003). In the case of these five jars, no clear clues about their function could have been obtained due to their fragmentary

conditions. Proteomic data of these nine samples identified hundreds of peptides. Most of them were from keratins and human skin-derived proteins, but many food proteins were also recognized. Taking into account that archeological materials have intrinsic contamination problems that can be originated at nearly any stage of burial, excavation, storage, and analysis, authentication and validation criteria are needed in order to discriminate ancient endogenous proteins from contaminants. First of all, it is important to highlight that, unlike their modern counterparts, the chemical structure of ancient proteins is altered by a series of complex diagenetic reactions (Cleland et al. 2020). Even if this is still an open question, the deamidation level of glutamine and asparagine is one of the proposed biomolecular indicators of deterioration and natural aging of proteins in artistic and archeological materials (Schroeter et al. 2016). It is important to note that asparagine and glutamine deamidate at vastly different rates, with glutamine being considerably slower (Robinson et al. 2001a; b). Consequently, glutamine deamidation has become more widely used as a means to assess time-dependent damage in ancient sample age. In particular, recent studies on paleontological bone have observed that known contaminants present in ancient protein extracts show very little to no deamidation, while endogenous bone collagens and non-collagenous proteins generally display medium or extensive deamidation (Welker et al. 2016, 2017). However, many environmental conditions, including pH, temperature, protein sequence, protein structure, ionic strength, etc., can affect the rate and total deamidation levels observed in a sample. Therefore, the only use of this modification as a definitive marker to estimate the absolute age of proteins derived from archeological materials may lead to the incorrect classification of exceptionally well-preserved, incompletely deamidated ancient sequences as contamination (Schroeter et al. 2016). Indeed, ancient proteins may also undergo other age-dependent oxidative damages (e.g., kynurenine formation, methionine or tryptophan dioxidation, carbonylation of lysine and arginine etc.) as observed in artificially aged wool and mammoth bone from permafrost (Cappellini et al. 2012) or experience spontaneous peptide backbone cleavages, which should be also examined. Concerning the last damage, protein fragmentation frequently occurs due to spontaneous non-enzymatic peptide backbone cleavage at the carboxyl side of asparagine, aspartic acid, glutamine, and glutamic acid (Joshi et al. 2004). On the other hand, it is also important to highlight that temperature, burial environment, and the fossil chemistry all influence how quickly a protein may decay. Moreover, it is not clear what mechanisms slow down decay, so that some protein sequences might show low damages after thousands of years. As a general result, it remains difficult to know where to look for these ancient sequences.

Proteomic results: identification of proteins and calculation of deamidation level

The complete collection of proteomic data (i.e., list of the identified proteins, taxonomy, percentage of sequence coverage, *InChorus* score, sequence of the corresponding peptides, etc.) obtained by PEAKS software, relative to the nine samples here investigated, is reported in the Supplementary Table S1.

A total of 758 peptides were characterized. 470 were from human proteins; whereas, 288 were related to milk or cereal proteins. Results reveal that, in many samples, a very high level of human skin-derived peptides, which probably represent contaminants, was detected. In particular, these peptides allowed the identification of 97 human skin-derived proteins, 57 cereal proteins, and 7 *Bovinae* milk-derived proteins. In detail, in the samples 100061B and 100061H, only human skin proteins were identified. Samples 100041 M and 100051D present a very high level of human skin protein and only one (α_{s1}-casein) and four milk-derived proteins (α_{s1}-, α_{s2}-, β-, and κ-casein), respectively. Instead, sample 100051A does not present human contaminants, but shows only two milk proteins, a β-casein and β-lactoglobulin, from *Bovinae* species (*Bos taurus or Bubalus bubalis*). With respect to this sample, it is interesting to note that the two β-lactoglobulin-derived peptides (TPEVD<u>D</u>EALEKFDK and TPEVD<u>D</u>EALEK) contain an important variant site that distinguishes *Bovinae* species from sheep. Indeed, an aspartic acid (D) in *Bovinae* (underlined in the previous peptides), is an asparagine (N) in sheep. Considering that the deamidation of asparagine results in its conversion to aspartic acid, protein identification software is unable to distinguish an unmodified *Bovinae* residue (D) from a deamidated sheep residue (DeN) at this position (Hendy et al. 2018a; Warriner et al. 2014). As a consequence, species discrimination is not possible at this stage. Finally, in the samples 100087C, 100087D, 100088B, and 100088D, extracted from four large storage jars (see Table 1) which were probably used for the storage of oil, wine or cereals, many cereal proteins were identified. On the contrary, only few human-derived proteins were detected. In detail, a total of nine wheat proteins (five in the sample 100087C and four in the sample 100087D) were identified; whereas, only one human protein was found (sample 100087D). Proteomic analysis of the samples 100088B and 100088D revealed the presence of 32 and 16 cereal proteins (from *Triticum* and *Hordeum* species), respectively, and a total of 14 human proteins (10 in the 100088B sample and 4 in the 100088D).

To discriminate food-derived proteins, as original components of our samples, from human proteins which probably represent contaminant components related to the post-excavation history of the samples, two aspects were evaluated: peptides deriving from non-enzymatic cleavages, and the deamidation level of asparagine (Asn) and glutamine (Gln) residues. As mentioned above, peptides generated by non-enzymatic cleavages at the carboxyl side of asparagine, glutamine, aspartic acid, and glutamic acid may be related to the degradation effect in ancient samples (Joshi et al. 2004). Moreover, Asn and Gln residues naturally deamidate over time (Robinson et al. 2001a), and these modifications may be used as biomolecular indicators of deterioration and natural aging of proteins in archeological materials. Concerning the evaluation of non-enzymatic cleavages, our results (Supplementary Table S1) show that about 10% (28 out of 288) of food-derived peptides arise by non-enzymatic cleavages, whereas only 2% (9 out of 470) of the human peptides are related to non-enzymatic cleavages. Then, by MQ software, we calculated the deamidation level of asparagine and glutamine residues in food-derived peptides identified in the archeological samples and compared it with that of human-derived peptides (classified as potential contaminants) (see Supporting Information and Supplementary Table S2). With the exclusion of the sample 100041 M, in which the score of all the deamidated and unmodified peptides was below the minimum score set, in the other six samples the deamidation level has been determined. Figure 3 shows that, in general, food-derived peptides present a deamidation level which is five to seven fold as high as that of the potential contaminants (e.g., human-derived peptides). Particularly, in the samples 100051A, 100051D, 100087C, 100087D, and 100088D, the level of deamidation ranges from 23 to 33% for asparagine residues and from 3 to 23% for glutamine residues in food-derived peptides. On the contrary, for the human peptides the level of deamidation ranges from 0 to 4% asparagine residues, and from 0 to 3% for glutamine residues (Fig. 3 and Supplementary Material for details). Instead, the level of deamidation of food-derived peptides in sample 100088B is lower than that in the other samples (6.7% Asn and 2.6% Gln). The lower level of deamidation observed for this sample might be probably due to well-preserved endogenous ancient components. On the other hand, unless other complementary chemical analyses, it cannot be excluded a priori that also food-derived peptides detected in this sample are instead related to the post-excavation history of the samples.

Taking into account that deamidation may also occur in vitro during proteomics sample handling (Xing et al. 2008), we also determined the deamidation level in two modern control samples: a commercial fresh milk sample and a wheat flour sample (see Supplementary Material for details). On average, 8.0% Asn and 9.1% Gln residues in milk, and 7.4% Asn and 4.8% Gln residues in wheat control sample resulted deamidated (Table S3 and Figure S1).
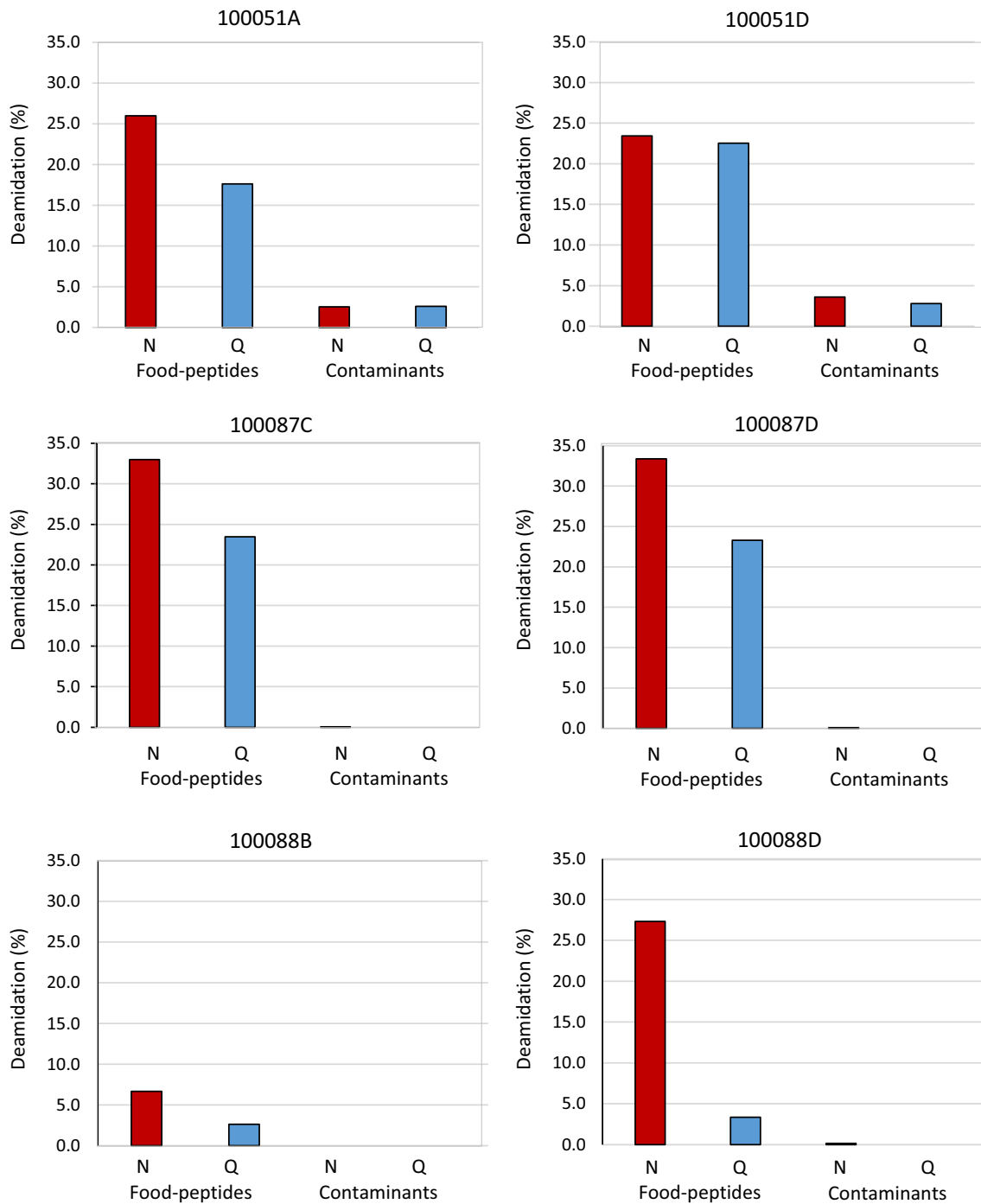
Fig. 3 Deamidation level (reported as percentage) of asparagine (N) and glutamine (Q) amino acids in the archeological samples analyzed

Discussion

Paleoproteomics is currently a growing research area that, by the use of liquid chromatography–tandem mass spectrometry (LC–MS/MS)-based approaches (known as "shotgun" methods), may offer an alternative strategy for identifying foodstuffs prepared in ceramics (Craig et al.

2000). It should be highlighted that, even if "shotgun" proteomic approaches permit the identification of a wide range of proteins in a sample, their application to archeological artifacts has so far been limited to exceptionally well-preserved examples from waterlogged, cold or arid contexts (Buckley et al. 2013; Solazzo et al. 2008).

In this work, we applied a shotgun proteomic approach to proteins extracted from ceramics of the pottery of the Maltese site of Il-Qlejgħa tal-Baħrija, the guide-site for the local Iron Age (11th–eighth century BCE), intending to confirm the preliminary archeological hypothesis about the presence of cereal storage practices, which could offer support to the hypothesis of an emergent proto-urban society based on economic accumulation redistribution of the surplus, and eventually to shed light on indigenous foodways. Our data also highlighted the impact of one of the many difficulties in identifying ancient proteins: the cross-contamination from modern proteins. Protein contamination of ancient and archeological materials are mainly, but not only, due to keratins and human skin-related proteins, that can occur at nearly any stage of the interaction of human agents with the artifact (use, burial, excavation, storage, and sample extraction) and may potentially provide false insights into protein composition, phylogeny, and protein modification, and may also affect the detection and identification of the ancient proteins. Consequently, in compliance with protection guidance for archeological samples, some precautions were adopted to minimize the effects of contamination from modern proteins (see material and methods section).

Proteomics data obtained for many samples here investigated revealed a high level or the exclusive presence of human skin protein contaminants and low amounts of food-derived proteins. To facilitate the interpretation, we subdivided the identified proteins into two groups: human-derived proteins, considered as a common background, were grouped together and classified as potential contaminants; the remaining protein hits derived from foods constituted the second group and were considered as potential authentic ancient proteins. Then, to distinguish between potential contaminants and authentic ancient proteins, we examined two aspects that may be related to diagenetic effects in ancient samples: the number of peptides generated by non-enzymatic cleavages and the deamidation level of Asn and Gln residues. As a general trend, we observed that the number of food-peptides arising by non-enzymatic cleavages is about five fold as high as that of potential contaminants. Similarly, food-derived peptides show a much higher deamidation level regarding that of human contaminants, which is the same as the modern samples used as control. Altogether these data seem to indicate that food-derived proteins identified in our samples can be considered as authentic components. On this respect, interesting archeological implications have the identification of milk in samples 100051A and 100051D. While such discovery for sample 100051A might confirm the common interpretation of such shape as related to cheese making, it is very remarkable that it is instead the case of sample 100051D. This tableware, traditionally interpreted as a beer-like fermented beverage container of foreign origin, eventually in Malta could have been used in a completely original and different way. In light of these archeological implications, it is important to note that the identification of milk proteins as authentic ancient components, and not contaminants, should be confirmed by complementary chemical analyses.

The presence of cereals-derived proteins in samples 100087C, 100087D, 100088B, and 100088D may inform us of the function of the four large storage jars in question as containers of cereals. On the other hand, complementary tests (e.g., lipid biomarker analysis) (Hammann et al., 2018) should be carried out to dispel or not hypothesize that these four jars could have also been contained olive oil or wine. The lower level of deamidation (the same of that calculated for modern control samples) observed for the food-derived peptides identified in the sample 100088B poses the question of whether these peptides may be considered well-preserved endogenous ancient components or are related to modern contaminants. Therefore, to validate the authenticity of these proteomics identification complementary chemical analyses should be performed. At this stage, it can only be hypothesized that encrustations, a layer rich of mineral-organic binder, retained on the surface of all the pottery here investigated, may have facilitated a long-term archeological preservation of ancient proteins and peptides, protecting from degradation and damage driven by external factors.

Finally, the lack of archeozoological and paleobotanical data for the settlement at Il-Qlejgħa tal-Baħrija makes the proteomic analysis outcomes even more relevant in terms of contributing to the reconstruction of foodways and dietary habits of that ancient community. In this respect, the assignment of many cereal proteins to *Triticum* species appears interesting. Indeed, the paleobotanical data, available for other sites in the Maltese archipelago, offer an interesting trend in terms of growth and consumption of cereals (Fiorentino et al. 2013). For most of the Neolithic and Copper Age (5000–2000 BCE ca.), *Hordeum* appears to be predominant with few instances identified of *Triticum*. Subsequently, in the Early Bronze Age (2000–1500 BCE ca.) barley decreases gradually in favor of naked kinds of wheat which reached their peak at this end of this period. But in the second half of the 2nd millennium BCE (*facies* of Borġ in-Nadur—1500–750 BCE ca.) barley increases significantly, up to 55% of the sample assemblage. At least for the specific case of the settlement at Il-Qlejgħa tal-Baħrija, our data, even if obtained by a small selection of vessels, instead seem to indicate that wheat was still consumed.

Conclusions

Food habits are constructed by a broad range of cultural, ideological, and interpersonal factors such as status, religion, gender, age, wealth, and more. In this perspective, food is

not just biologically necessary, but also it becomes a cognitively prominent "material culture" that plays an active role in constructing and negotiating social distinctions. While such very peculiar material in the case of classical antiquity is mainly addressed through the study of written and iconographic sources, concerning prehistoric civilizations, chemical techniques of investigation are currently the sole approach available. In particular, archeological ceramics occupy a place of absolute importance among the artifacts that have come down to us since prehistoric times, taking with them hidden biomolecular information about their function.

Traditional hypotheses on the contents of ancient vessels are based on comparisons with later examples, for which there is some historical certainty, or simply on the observation of shape and decorative style peculiarities. In this perspective, paleoproteomic study of ancient residual food-related compounds absorbed by pottery represents a powerful alternative approach that may inform us about dietary habits, culinary practices, and foodways. Moreover, it can also provide the opportunity for improved tissue and taxonomic resolution, as the identification of proteotypic peptides, i.e., peptides that are unique to the protein sequence specific for an individual organism and not in common to other species, permits to discriminate the organism of origin of the identified proteins. Concerning the set of samples here investigated, even if limited to a small selection of vessels, proteomics results revealed that wheat was central in the diet of the ancient community of Il-Qlejgħa tal-Baħrija. However, it is also important to note that in paleoproteomic studies protein identification is obtained investigating database which contain the nearly complete proteomes of many model organisms (e.g., *Mus musculus*, *Homo sapiens,* etc.), but comprise only partial proteomes of other taxa (e.g., *Capra hircus*, *Ovis aries,* etc.), included the *Triticum* species and other monocots.

In conclusion, it is important to note that as a young discipline, paleoproteomics, especially when applied to ancient ceramics, still needs to be developed in terms of efficiency of protocols and appropriate databases applied. Further studies about the PTMs characterization, amino acid composition, chemical, and physical properties of the identified peptides will give us interesting information about what determines the characteristics and stability of proteins during the time. Also, major understanding of how proteins are transferred from food to the ceramic and how they survive into the pots is needed. The mechanisms behind the preservation of ancient protein are not well understood, but are thought to involve a range of circumstances, including the composition of specific proteins (e.g., hydrophobicity properties), dry conditions, cold temperatures, encapsulation within mineral matrices, and chemical modifications. It is well known that protein residues are subject to transformations resulting from exposure to the environment. Although these modifications introduce a challenge for protein identification and authentication, they are also thought to promote preservation by rendering protein residues less susceptible to microbial attack (Barker et al. 2012, 2018).

Our results also confirm that archeological deposits are usually composed of degraded, mixed and contaminated compounds (mainly human-derived proteins), and their analysis requires a complex analytical approach, including methods to discriminate authentic from contaminant proteins. The lower deamidation level observed in a sample (e.g., 100088B) for cereal-derived peptides might be probably related, but not limited, to the presence of encrustation layer which may have well-preserved protein molecules present into the ceramic. In this respect, although the chemical and mechanical removal of encrustations in ceramic vessels is still considered good practices in archeology (Casaletto et al. 2008), these results confirm the importance of these in situ deposits for preserving protein molecules as already reported (Demarchi et al. 2016; Hendy et al. 2018b). However, it is essential to highlight that different environmental factors and the inherent properties of proteins, including their primary sequences and three-dimensional structures, might have also affected deamidation and other post-translational modifications. On the other hand, this study confirms the challenges related to the investigation of ancient proteins, especially if extracted from ceramic substrates. In this specific case, proteomic data appears attractive in terms of contributing to the reconstruction of dietary habits of the ancient community of the Maltese site of Baħrija, but it would be desirable for them to be cross-validated, in the future, by complementary analyses, such as investigation of lipid biomarkers, aDNA analysis, and carbohydrates characterization. Indeed, it is important to highlight that combining the results of ancient biomolecules (proteins, lipids, carbohydrates, and aDNA) can provide a more coherent approach to tackling complex biological questions than the use of any of these biomolecular proxies alone.

archeological materials. All authors provided intellectual input to the study, reviewed, edited and approved the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding Open Access funding provided by Università degli Studi di Catania. This research was partly supported by "Piano della Ricerca di Ateneo 2016–2018" of the University of Catania, Italy.

Compliance with ethical standards

Conflicts of interest The authors declare no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Albanese Procelli RM (2003) Sicani, Siculi, Elimi: forme di identità, modi di contatto e processi di, trasformazione. Longanesi, Milano

Bonanno A (2017) From Prehistoric to Historic Malta: a Discussion of the Source. In: Mizzi D, Vella NC, Zammit MR, Frendo AJ (eds) "What mean these stones?" (Joshua 4:6, 21): essays on texts, philology, and archaeology. In honour of Anthony J. Frendo, Leuven: Peeters, 243–254

Barker A, Venables B, Stevens SM Jr, Seeley KW, Wang P, Wolverton S (2012) An optimized approach for protein residue extraction and identification from ceramics after cooking. J Archaeol Method Theory 19:407–439. https://doi.org/10.1007/s10816-011-9120-5

Barker A, Dombrosky J, Venables B, Wolverton S (2018) Taphonomy and negative results: An integrated approach to ceramic-bound protein residue analysis. J Archaeol Sci 94:32–43. https://doi.org/10.1016/j.jas.2018.03.004

Borek T, Mowry C, Dean G (2007) Analysis of modern and ancient artifacts for the presence of corn beer; dynamic headspace testing of pottery sherds from mexico and new mexico. MRS Proc. https://doi.org/10.1557/PROC-1047-Y01-05.

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. https://doi.org/10.1002/rcm.6481

Burkhart M, Schumbrutzki C, Wortelkamp S, Sickman A, Zahedi RP (2012) Systematic and quantitative comparison of digest efficiency and specificity reveals the impact of trypsin quality on MS-based proteomics. J Proteom 75:1454–1462. https://doi.org/10.1016/j.jprot.2011.11.016

Cappellini E, Jensen LJ, Szklarczyk D, Ginolhac A, Da Fonseca RA, Stafford TW, Holen SR, Collins MJ, Orlando L, Willerslev E, Gilbert MT, Olsen JV (2012) Proteomic analysis of a pleistocene mammoth femur reveals more than one hundred ancient bone proteins. J Proteome Res 11:917–926. https://doi.org/10.1021/pr200721u

Cappellini E, Collins MJ, Gilbert MT (2014a) Biochemistry. Unlocking ancient protein palimpsests. Science 343:1320–1322. https://doi.org/10.1126/science.1249274

Cappellini E, Gentry A, Palkopoulou E, Ishida Y, Cram D, Roos AM, Watson M, Johansson US, Fernholm B, Agnelli P, Barbagli F, Littlewood DTJ, Kelstrup CD, Olsen JV, Lister AM, Roca AL, Dalén L, Gilbert MTP (2014b) Resolution of the type material of the Asian elephant, *Elephas maximus* Linnaeus, 1758 (Proboscidea, Elephantidae). Zool J Linn Soc 170:222–232. https://doi.org/10.1111/zoj.12084

Cappellini E, Prohaska A, Racimo F, Welker F, Pedersen MW, Allentoft ME, Damgaard PB, Gutenbrunner P, Dunne J, Hammann S, Roffet-Salque M, Ilardo M, Moreno-Mayar JV, Wang Y, Sikora M, Vinner L, Cox J, Evershed RP, Willerslev E (2018) Ancient biomolecules and evolutionary inference. Annu Rev Biochem 87:1029–1060. https://doi.org/10.1146/annurev-biochem-062917-012002

Cardona D (2020), Il-Qlejgej tal-Baal-Ba: the story, D. Tanasi, D. Cardona Eds.; The Maltese Archipelago at the dawn of history. Reassessment of the 1909 and 1959 excavations at Qlejgħa tal-Baħrija and other essays, Oxford 3–18

Cardona D, Zammit M (2015) The Bronze Age settlement at il-Qlejgħa tal-Baħrija, Malta: notes on the rock-cut features. In: Tanasi D, Vella NC (eds.) The Late Prehistory of Malta, Oxford: Archaeopress 161–171

Casaletto MP, Ingo GM, Riccucci C, De Caro T, Bultrini G, Fragalà I, Leoni M (2008) Chemical cleaning of encrustations on archaeological ceramic artefacts found in different Italian sites. Appl Phys A 92:35–42. https://doi.org/10.1007/s00339-008-4519-x

Cavalieri D, McGovern PE, Hartl DL, Mortimer R, Polsinell M (2003) Evidence for *S. cerevisiae* fermentation in ancient wine. J Mol Evol 57:S226–S232. https://doi.org/10.1007/s00239-003-0031-2

Charlton S, Ramsøe A, Collins M, Craig OE, Fischer R, Alexander M, Speller CF (2019) New insights into Neolithic milk consumption through proteomic analysis of dental calculus. Archaeol Anthropol Sci 11:6183–6196. https://doi.org/10.1007/s12520-019-00911-7

Cleland TP, Schroeter ER, Schweitzer MH (2015) Biologically and diagenetically derived peptide modifications in moa collagens. Proc R Soc B 282:20150015. https://doi.org/10.1098/rspb.2015.0015

Cleland TP, Schroeter ER (2018) A comparison of common mass spectrometry approaches for paleoproteomics. J Proteome Res 17:936–945. https://doi.org/10.1021/acs.jproteome.7b00703

Cleland TP, Schroeter ER, Colleary C (2020) Diagenetiforms: a new term to explain protein changes as a result of diagenesis in paleoproteomics. J Proteomics 26:103992. https://doi.org/10.1016/j.jprot.2020.103992

Colombini MP, Giachi G, Modugno F, Ribechini E (2005) Characterisation of organic residues in pottery vessels of the Roman age from Antinoe (Egypt). Microchem J 79:83–90. https://doi.org/10.1016/j.microc.2004.05.004

Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26:1367–1372. https://doi.org/10.1038/nbt.1511

Craig OE, Mulville J, Pearson MP, Sokol R, Gelsthorpe K, Stacey R, Collins M (2000) Detecting milk proteins in ancient pots. Nature 408:312. https://doi.org/10.1038/35042684

D'Amato A, Zilberstein G, Zilberstein S, Compagnoni L, Righetti PG (2018) Of mice and men: traces of life in the death registries of the 1630 plague in Milano. J Proteomics 180:128–137. https://doi.org/10.1016/j.jprot.2017.11.028

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. https://doi.org/10.1007/s00216-010-4218-2

Dallongeville S, Garnier N, Rolando C, Tokarski C (2016) Proteins in art, archaeology, and paleontology: from detection to identification. Chem Rev 116:2–79. https://doi.org/10.1021/acs.chemrev.5b00037

Demarchi B, Hall S, Roncal-Herrero T, Freeman CL, Woolley J, Crisp MK, Wilson J, Fotakis A, Fischer R, Kessler BM, Rakownikow JCR, Olsen JV, Haile J, Thomas J, Marean CW, Parkington J, Presslee S, Lee-Thorp J, Ditchfield P, Hamilton JF, Ward MW, Wang CM, Shaw MD, Harrison T, Domínguez-Rodrigo M, MacPhee R, Kwekason A, Ecker M, Kolska Horwitz L, Chazan M, Kröger R, Thomas-Oates J, Harding JH, Cappellini E, Penkman K, Collins MJ (2016) Protein sequences bound to mineral surfaces persist into deep time. Elife 5:17092. https://doi.org/10.7554/eLife.17092

Dudd SN, Evershed RP (1998) Direct demonstration of milk as an element of archaeological economies. Science 282:1478–1481. https://doi.org/10.1126/science.282.5393.1478

Evershed RP, Dudd SN, Copley MS, Mutherjee A (2002) Identification of animal fats via compound specific δ13C values of individual fatty acids: assessments of results for reference fats and lipid extracts of archaeological pottery vessels. Documenta Praehistorica 29:73–96. https://doi.org/10.4312/dp.29.7

Fiorentino G, D'Oronzo C, Colaianni G (2013) Human-environmental interaction in Malta from Neolithic to the Roman period: Archaeobotanical analyses at Tas-Silġ. Scienze dell'Antichità 18:169–184

Gismondi A, Baldoni M, Gnes M, Scorrano G, D'Agostino A, Di Marco G, Calabria G, Petrucci M, Muldner G, Von Tersch M, Nardi A, Enei F, Canini A, Rickards O, Alexander M, Martínez-Labarga C (2020) A multidisciplinary approach for investigating dietary and medicinal habits of the Medieval population of Santa Severa (7th-15th centuries, Rome, Italy). PLoS ONE 15(1):e0227433. https://doi.org/10.1371/journal.pone.0227433

Greco E, El-Aguizy O, Ali MF, Foti S, Cunsolo V, Saletti R, Ciliberto E (2018) Proteomic analyses on an ancient Egyptian cheese and biomolecular evidence of brucellosis. Anal Chem 90:9673–9676. https://doi.org/10.1021/acs.analchem.8b02535

Guasch-Jané MR, Gómez MI, Lacueva CA, Jáuregui O, Lamuela-Raventós RM (2004) Liquid chromatography with mass spectrometry in tandem mode applied for the identification of wine markers in residues from ancient Egyptian vessels. Anal Chem 76:1672–1677. https://doi.org/10.1021/ac035082z

Hammann S, Cramp LJE (2018) Towards the detection of dietary cereal processing through absorbed lipid biomarkers in archaeological pottery. J Archeol Sci 93:74–81. https://doi.org/10.1016/j.jas.2018.02.017

Hansel FA, Copley MS, Madureira LAS, Evershed RP (2004) Thermally produced ω-(o-alkylphenyl)alkanoic acids provide evidence for the processing of marine products in archaeological pottery vessels. Tetrahedron Lett 45:2999–3002. https://doi.org/10.1016/j.tetlet.2004.01.111

Hendy J, Collins M, Teoh KY, Ashford DA, Oateshi JT, Donoghue ED, Pap I, Minnikin DE, Spigelman M, Buckley M (2016) The challenge of identifying tuberculosis proteins in archaeological tissues. J Archaeol Sci 66:146–153. https://doi.org/10.1016/j.jas.2016.01.003

Hendy J, Warinner C, Bouwman A, Collins MJ, Fiddyment S, Fischer R, Hagan R, Hofman CA, Holst A, Chaves E, Klaus L, Larson G, Mackie M, McGrath K, Mundorff AZ, Radini A, Rao H, Trachsel C, Velsko IM, Speller CF (2018a) Proteomic evidence of dietary sources in ancient dental calculus. Proc R Soc B 285:20180977. https://doi.org/10.1098/rspb.2018.0977

Hendy J, Colonese AC, Franz I, Fernandes R, Fischer R, Orton D, Lucquin A, Spindler L, Anvari J, Stroud E, Biehl PF, Speller C, Boivin N, Mackie M, Jersie-Christensen RR, Olsen JV, Collins MJ, Craig OE, Rosenstock E (2018b) Ancient proteins from ceramic vessels at Çatalhöyük West reveal the hidden cuisine of early farmers. Nat Commun 9:4064–4073. https://doi.org/10.1038/s41467-018-06335-6

Hendy J, Welker F, Speller C, Warinner C, Collins MJ (2018c) A guide to ancient protein studies. Nat Ecol Evol 2:791–799. https://doi.org/10.1038/s41559-018-0510-x

Joshi AB, Kirsch LE (2004) The estimation of glutaminyl deamidation and aspartyl cleavage rates in glucagon. Int J Pharm 273:213–219. https://doi.org/10.1016/j.ijpharm.2004.01.006

Keller A, Nesvizhskii AI, Kolker E, Aebersold R (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem 74:5383–5392. https://doi.org/10.1021/ac025747h

Leighton R (1981) Strainer-spouted jugs and the problem of the earliest Phoenician influence in Sicily. J Mediterranean Anthropol Archaeol 1:280–291

Leonard C, Vashro L, O'Connell JF, Henry AG (2015) Plant microremains in dental calculus as a record of plant consumption: a test with Twe forager-horticulturalists. J Archaeol Sci Rep. https://doi.org/10.1016/j.jasrep.2015.03.009

Mackie M, Rüther P, Samodova D, Di Gianvincenzo F, Granzotto C, Lyon D, Peggie DA, Howard H, Harrison L, Jensen LJ, Olsen JV, Cappellini E (2018) Palaeoproteomic profiling of conservation layers on a 14th century Italian wall painting. Angew Chem Int Ed 57(25):7369–7374. https://doi.org/10.1002/anie.201713020

McGovern PE, Hall GR, Mirzoian A (2013) A biomolecular archaeological approach to 'Nordic grog', Danish. J Arch 2:112–131. https://doi.org/10.1080/21662282.2013.867101

Mikšík I, Sedláková P, Pataridis S, Bortolotti F, Gottardo R (2016) Proteins and their modificationsin a medieval mummy. Protein Sci. https://doi.org/10.1002/pro.3024

Mottram HR, Dudd SN, Lawrence GJ, Stott AW, Evershed RP (1999) New chromatographic, mass spectrometric and stable isotope approaches to the classification of degraded animal fats preserved in archaeological pottery. J Chrom A 833:209–221. https://doi.org/10.1016/S0021-9673(98)01041-3

Ostrom PH, Schall M, Gandhi H, Shen TL, Hauschka PV, Strahler JR (2000) New strategies for characterizing ancient proteins using matrix-assisted laser desorption ionization mass spectrometry. Geochim Cosmochim Acta 64:1043–1050. https://doi.org/10.1016/S0016-7037(99)00381-6

Peet TE (1910) Contributions to the study of the prehistoric period in Malta. Pap Br School Rome 5:141–163. https://doi.org/10.1017/S0068246200005250

Perutka Z, Šebela M (2018) Pseudotrypsin: A Little-Known Trypsin Proteoform. Molecules 23:2637–2651. https://doi.org/10.3390/molecules23102637

Recchia G (2000) La funzione dei contenitori ceramici dell'età del Bronzo nell'Italia Medionale: una prospettiva etnoarcheologica. Archeologia Postmedievale 4:111–122

Robinson NE, Robinson AB (2001a) Molecular clocks. Proc Natl Acad Sci 98:944–949. https://doi.org/10.1073/pnas.98.3.944

Robinson NE, Robinson AB (2001b) Prediction of protein deamidation rates from primary and three-dimensional structure. Proc Natl Acad Sci USA 98:4367–4372. https://doi.org/10.1073/pnas.071066498

Robinson NE (2002) Protein deamidation. Proc Natl Acad Sci 99:5283–5288. https://doi.org/10.1073/pnas.082102799

Saletti R, Reina S, Pittalà MGG, Magrì A, Cunsolo V, Foti S, De Pinto V (2018) Post-translational modifications of VDAC1 and VDAC2 cysteines from rat live mitochondria. Biochim Biophys Acta Bioenergetics 1859:806–816. https://doi.org/10.1016/j.bbabio.2018.06.007

Salque M, Bogucki P, Pyzel J, Sobkowiak-Tabaka I, Grygiel R, Szmyt M, Evershed RP (2013) Earliest evidence for cheese making in the sixth millennium BC in northern Europe. Nature 493:522–525. https://doi.org/10.1038/nature11698

Schroeter ER, Cleland TP (2016) Glutamine deamidation: an indicator of antiquity, or preservational quality? Rapid Commun Mass Spectrom 30:251–255. https://doi.org/10.1002/rcm.7445

Shevchenko A, Yang Y, Knaust A, Thomas H, Jiang H, Lu E, Wang C, Shevchenko A (2014) Proteomics identifies the composition and manufacturing recipe of the 2500-year old sourdough bread from Subeixi cemetery in China. J Proteomics 105:363–371. https://doi.org/10.1016/j.jprot.2013.11.016

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(11):e0206483. https://doi.org/10.1371/journal.pone.0206483

Skibo JM (2012) Understanding Pottery Function. In: Understanding Pottery Function. Manuals in Archaeological Method, Theory and Technique Book 1–25. Springer Ed

Skibo JM (2015) Pottery Use-Alteration Analysis. In: Use-Wear Methodology on the Analysis of Osseous Industries. Manuals in Archaeological Method, Theory and Technique 189–198. Ed. Springer

Solazzo C, Fitzhugh WW, Rolando C, Tokarski C (2008) Identification of protein remains in archaeological potsherds by proteomics. Anal Chem 80:4590–4597. https://doi.org/10.1021/ac800515v

Søe MJ, Nejsum P, Seersholm FV, Fredensborg BL, Habraken R, Haase K, Hald MM, Simonsen R, Højlund F, Blanke L, Inga M, Willerslev E, Kapel CMO (2018) Ancient DNA from latrines in Northern Europe and the Middle East (500 BC–1700 AD) reveals past parasites and diet. PLoS ONE 13:e0195481. https://doi.org/10.1371/journal.pone.0195481

Spangenberg JE, Jacomet S, Schibler J (2006) Chemical analyses of organic residues in archaeological pottery from Arbon Bleiche 3, Switzerland – evidence for dairying in the late Neolithic. J Arch Sci 33(1–13):3. https://doi.org/10.1016/j.jas.2005.05.013

Spiteri C, Belser M, Crispino A (2020) Preliminary results on content analysis of Early Bronze Age vessels from the site of Castelluccio, Noto, Sicily. J Archaeol Sci Rep 31:102355. https://doi.org/10.1016/j.jasrep.2020.102355

Stevens SM Jr, Wolverton 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. https://doi.org/10.1007/s00216-009-3341-4

Tanasi D, (2020) Bronze and Iron Age pottery from the 1909 and 1959 excavations at Qlejgħa tal-Baħrija, D. Tanasi, D. Cardona Eds.; The Maltese Archipelago at the dawn of history. Reassessment of the 1909 and 1959 excavations at Qlejgħa tal-Baħrija and other essays, Oxford 35–92

Tanasi D, Cardona D Eds. (2020) The Maltese Archipelago at the dawn of history. Reassessment of the 1909 and 1959 excavations at Qlejgħa tal-Baħrija and other essays, Oxford

Tokarski C, Martin E, Cren-Olivé C, Rolando C (2003) Protein Studies in Cultural Heritage. Molecular Biology and Cultural Heritage; C. Saiz-Jimenez, Ed.; Sevilla 119−130

Trump DH (1961) The later prehistory of Malta. Proc Prehist Soc 27:253–262

Van Doorn NL, Wilson J, Hollund H, Soressi M, Collins MJ (2012) Site-specific deamidation of glutamine: a new marker of bone collagen deterioration. Rapid Commun Mass Spectrom 26:2319–2327. https://doi.org/10.1002/rcm.6351

Velsko IM, Overmyer KA, Speller C, Klaus L, Collins MJ, Loe L, Frantz LAF, Sankaranarayanan K, Lewis CM Jr, Martinez JBR, Chaves E, Coon JJ, Larson G, Warinner C (2017) The dental calculus metabolome in modern and historic samples. bioRxiv. https://doi.org/10.1101/136176

Vizcaino JA, Cote RG, Csordas A, Dianes JA, Fabregat A, Foster JM, Griss J, Alpi E, Birim M, Contell J, O'Kelly G, Schoenegger A, Ovelleiro D, Perez-Riverol Y, Reisinger F, Rios D, Wang R, Hermjakob H (2013) The Proteomics Identifications (PRIDE) database and associated tools: status in 2013. Nucleic Acids Res 41:D1063-1069. https://doi.org/10.1093/nar/gks1262

Vinciguerra R, De Chiaro A, Pucci P, Marino G, Birolo L (2016) Proteomics strategies for cultural heritage: from bones to paintings. Microchem J 126:341–348. https://doi.org/10.1016/j.microc.2015.12.024

Warinner C, Hendy J, Speller C, Cappellini E, Fischer R, Trachsel C, Arneborg J, Lynnerup N, Craig OE, Swallow DM, Fotakis A, Christensen RJ, Olsen JV, Liebert A, Montalva N, Fiddyment S, Charlton S, Mackie M, Canci A, Bouwman A, Rühli F, Gilbert MTP, Collins MJ (2014) Direct evidence of milk consumption from ancient human dental calculus. Sci Rep 4:7104. https://doi.org/10.1038/srep07104

Warinner C, Speller C, Collins MJ (2015) A new era in palaeomicrobiology: prospects for ancient dental calculus as a long-term record of the human oral microbiome. Phil Trans R Soc B 370:20130376. https://doi.org/10.1098/rstb.2013.0376

Welker F, Collins MJ, Thomas JA, Wadsley M, Brace S, Cappellini E, Turvey ST, Reguero M, Gelfo JN, Kramarz A, Burger J, Thomas-Oates J, Ashford DA, Ashton PD, Rowsell K, Porter DM, Kessler B, Fischer R, Baessmann C, Kaspar S, Olsen JV, Kiley P, Elliott JA, Kelstrup CD, Mullin V, Hofreiter M, Willerslev E, Hublin JJ, Orlando L, Barnes I, MacPhee RDE (2015) Ancient proteins resolve the evolutionary history of Darwin/'s South American ungulates. Nature 522:81–84. https://doi.org/10.1038/nature14249

Welker F, Hajdinjak M, Talamo S, Jaouen K, Dannemann M, David F, Julien M, Meyer M, Kelso J, Barnes I, Brace S, Kamminga P, Fischer R, Kessler BM, Stewart JR, Pääbo S, Collins MJ, Hublin JJ (2016) Palaeoproteomic, evidence identifies archaic hominins associated with the Châtelperronian at the Grotte du Renne. PNAS 113:11162–11167. https://doi.org/10.1073/pnas.1605834113

Welker F, Smith GM, Hutson JM, Kindler L, Garcia-Moreno A, Villaluenga A, Turner E, Gaudzinski-Windheuser S (2017) Middle Pleistocene protein sequences from the rhinoceros' genus Stephanorhinus and the phylogeny of extant and extinct Middle/Late Pleistocene Rhinocerotidae. Peer J 5:e3033–e3050. https://doi.org/10.7717/peerj.3033

Weyrich LS, Dobney K, Cooper A (2015) Ancient DNA analysis of dental calculus. J Hum Evol 79:119–124. https://doi.org/10.1016/j.jhevol.2014.06.018

Wilson J, Van Doorn NL, Collins MJ (2012) Assessing the extent of bone degradation using glutamine deamidation in collagen. Anal Chem 84:9041–9048. https://doi.org/10.1021/ac301333t

Yang Y, Shevchenko A, Knaust A, Abuduresuled I, Li W, Hu X, Wanga C, Shevchenko A (2014) Proteomics evidence for kefir dairy in Early Bronze Age China. J Archaeol Sci 45:178–186. https://doi.org/10.1016/j.jas.2014.02.005

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 (3800 BP, northwestern China). Quat Int 426:158–165. https://doi.org/10.1016/j.quaint.2016.04.015

Xing G, Zhang J, Chen Y, Zhao YJ (2008) Identification of four novel types of in vitro protein modifications. Proteome Res 7:4603–4608. https://doi.org/10.1021/pr800456q