Styring, A., Maier, U., Stephan, E., Schlichtherle, H., Bogaard, A., 2016. Cultivation of choice: new insights into farming practices at Neolithic lakeshore sites. Antiquity 90, 95-110. https://doi.org/10.15184/aqy.2015.192

Styring, A.K., Charles, M., Fantone, F., Hald, M.M., McMahon, A., Meadow, R.H., Nicholls, G.K., Patel, A.K., Pitre, M.C., Smith, A., Sołtysiak, A., Stein, G., Weber, J.A., Weiss, H., Bogaard, A., 2017. Isotope evidence for agricultural extensification reveals how the world's first cities were fed. Nature Plants 3, 17076. https://doi.org/10.1038/nplants.2017.76

Szpak, P., Chiou, K.L., 2019. A comparison of nitrogen isotope compositions of charred and desiccated botanical remains from northern Peru. Vegetation History and Archaeobotany 29, 527-538. https://doi.org/10.1007/s00334-019-00761-2

Szpak, P., Metcalfe, J.Z., Macdonald, R.A., 2017. Best practices for calibrating and reporting stable isotope measurements in archaeology. Journal of Archaeological Science: Reports 13, 609-616. https://doi.org/10.1016/j.jasrep.2017.05.007

Vaiglova, P., Bogaard, A., Collins, M., Cavanagh, W., Mee, C., Renard, J., Lamb, A., Gardeisen, A., Fraser, R., 2014. An integrated stable isotope study of plants and animals from Kouphovouno, southern Greece: a new look at Neolithic farming. Journal of Archaeological Science 42, 201-215. https://doi.org/10.1016/j.jas.2013.10.023

Vaiglova, P., Halstead, P., Pappa, M., Triantaphyllou, S., Valamoti, S.M., Evans, J., Fraser, R., Karkanas, P., Kay, A., Lee-Thorp, J., Bogaard, A., 2018. Of cattle and feasts: Multi-isotope investigation of animal husbandry and communal feasting at Neolithic Makriyalos, northern Greece. PLoS ONE 13, e0194474. https://doi.org/10.1371/journal.pone.0194474

Ventresca Miller, A.R., Bragina, T.M., Abil, Y.A., Rulyova, M.M., Makarewicz, C.A., 2019. Pasture usage by ancient pastoralists in the northern Kazakh steppe informed by carbon and nitrogen isoscapes of contemporary floral biomes. Archaeological and Anthropological Sciences 11, 2151-2166. https://doi.org/10.1007/s12520-018-0660-4

Wallace, M., Jones, G., Charles, M., Fraser, R., Halstead, P., Heaton, T., Bogaard, A., 2013. Stable carbon isotope analysis as a direct means of inferring crop water status and water management practices. World Archaeology 45, 388-409.

CHAPTER 9

Sampling for ancient biomolecular data: DNA, proteins and lipids from specific archaeological sample types

Macleod Ruairidh *[1,2,3], Miranda Evans *[2], Ainara Sistiaga *[1,4], Theis T. Jensen *[1], Trish Biers *[2], Ian Barnes *[5] and Matthew J. Collins *[1,2]

1. Globe Institute, University of Copenhagen, Denmark
2. Department of Archaeology, University of Cambridge, UK
3. Research Department of Genetics, Evolution and Environment, University College London, UK
4. Earth, Atmospheric and Planetary Sciences Department, Massachusetts Institute of Technology, USA
5. Ancient DNA Laboratory, Natural History Museum, London, UK

*contact:
r@palaeome.org (RM)
miranda@palaeome.org (ME)
asistiaga@palaeome.org (AS)
theistrollejensen@palaeome.org (TJ)
tmb40@cam.ac.uk (TB)
i.barnes@nhm.ac.uk (IB)
matthew@palaeome.org (MC)

Introduction

Ancient biomolecular analyses have rapidly risen to prominence in archaeological research, generating a considerable breadth of novel evidence for behaviour, population history and subsistence in the past.

Carbohydrates rarely feature in studies at present; although important as foodstuffs (e.g. sugars, starch), glues, textiles (e.g. flax, cotton), writing media (papyrus, paper) and in construction (papier-mâché), they survive poorly in the archaeological record. RNA also survives poorly, presumably mainly due to enzymolysis, but has been recovered over millennia. Lignin is highly resistant to biodegradation under anoxic conditions, but contains limited biomolecular information. Metabolites other than lipids have only recently been systematically analysed in archaeological remains and are typically highly susceptible to biodegradation. These classes are not considered further in this chapter.

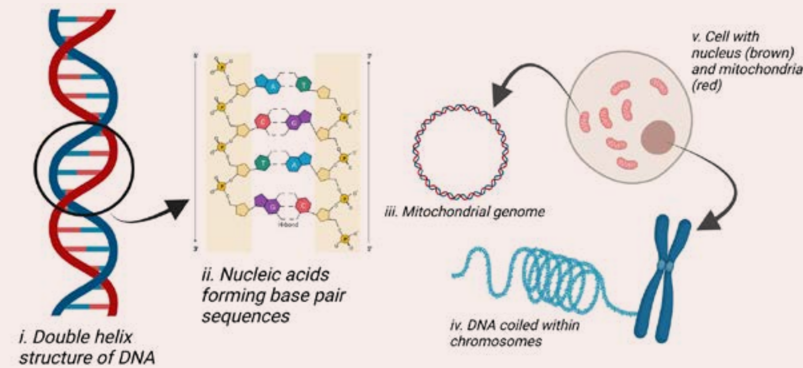
Three main classes of biomolecule are the focus of these guidelines: DNA (the fragile complex molecule that encodes heritable information), proteins (large biomolecules that perform many functional roles within cells, and are often preserved as polypeptide chains) and lipids (molecules involved in storing energy and in cell structures, such as fats, that are insoluble in water). Descriptions of sampling approaches for these biomolecule classes are grouped here by types of preserved substrate (e.g. hard tissue, soft tissue) given the distinct considerations required of each.

The principal limiting factor to ancient biomolecular analysis is preservation, and considering each substrate type individually is intended to take into account their differing preservational characteristics. The key determinant of biomolecule preservation is the environmental history of the site. Hot temperatures, humidity and unstable burial environments (e.g. unsaturated zone), will accelerate degradational processes while cool, stable environments will decelerate these. In terms of biodegradation, DNA is typically the most fragile, disintegrating into short fragments upon cell death. Proteins (degraded as peptides) may persist considerably longer when associated within minerals such as enamel. Lipids are the most stable class, with their carbon backbones surviving deep into the geological record.

However, proteins and lipids are presently especially vulnerable to contamination in analyses, as post-hoc contamination detection methods do not exist for these in the same way as for DNA. Guidelines on optimal sample storage are also included in each section (together with other key sampling details), bearing in mind that access to ideal facilities may often be lacking in the field. This overview is intentionally simplified to describe typical field considerations, and there will be exceptions to the general rules we outline below.

Figure 1: DNA, proteins and lipids explanatory figure. Figure created in Biorender.com

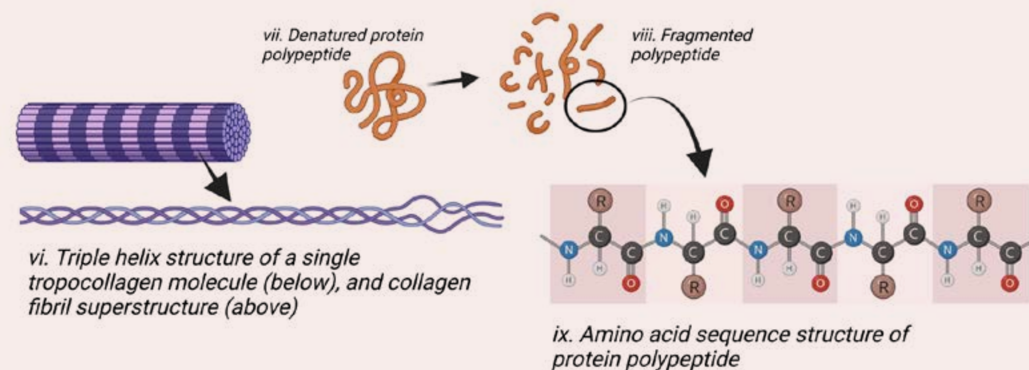
DNA



DNA (deoxyribonucleic acid) is a long double-stranded helical molecule (*i*) composed of four repeating nucleobases (*ii*) adenine (A), thymine (T), guanine (G) and cytosine (C). The sequence of the base pair units encodes genetic information relating to the organism. DNA is located within various locations in the cell, although the principal sources are within the chromosomes inside the cell nucleus (the nuclear genome) and the many mitochondria within the cell cytoplasm (mitochondrial genomes). In humans, the nuclear genome comprises 3.2 billion base pairs, and the mitochondrial genome 16,569 base pairs.

During sexual reproduction, chromosomal DNA recombines to produce new combinations of gene variants. Mitochondrial DNA and Y-chromosomal DNA do not recombine however, and are inherited directly from mother to offspring and father to offspring, respectively.

PROTEINS

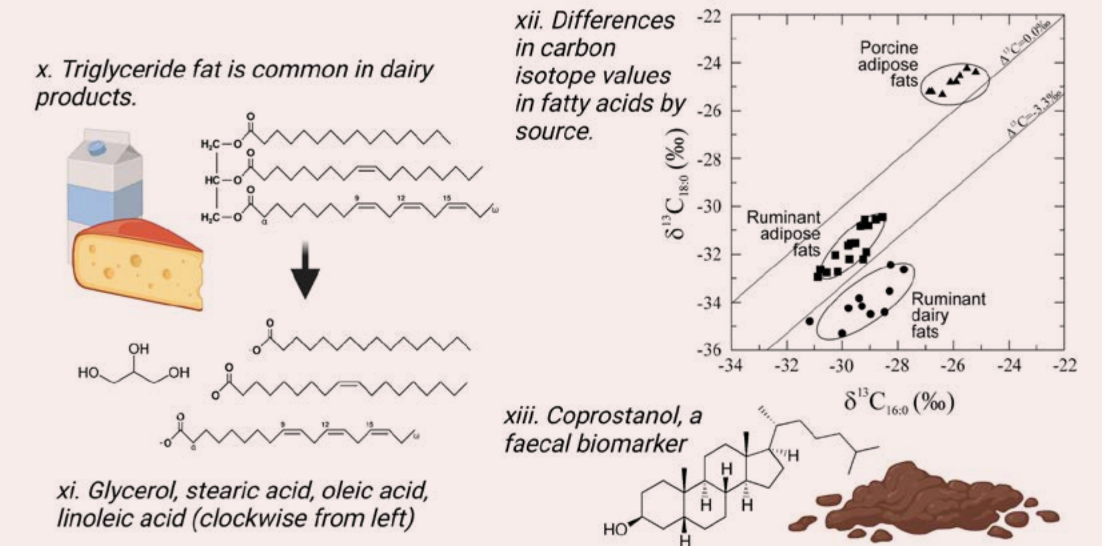


Proteins are large complex molecules that perform a considerable variety of functions within organisms. One of the most ubiquitous in all animals is collagen, an abundant structural protein composed of triple helices of polypeptide chains (tropocollagen, *vi*), bundled as collagen fibrils that in turn are grouped as collagen fibres. All proteins have a primary structure as a polypeptide chain (*vii*) prior to folding and other molecular modifications. This consists of peptide-bonded amino acids (*ix*). In archaeological contexts, these polypeptides may be highly fragmented (*viii*), with shorter sequences of little use for analyses.

The sequence of amino acids within polypeptides is determined by genetic information from DNA, via messenger RNA sequences.

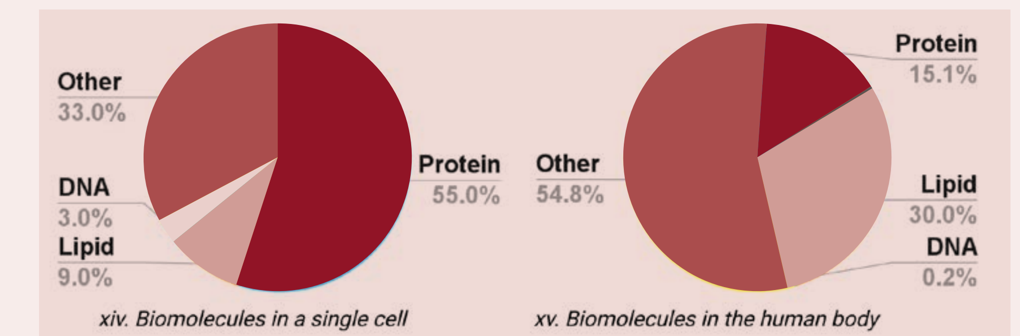
LIPIDS

Structural chemical formula from Wikimedia Commons. Plot *xii* from Copley et al. 2003.



Lipids are a broad group of hydrophobic molecules involved in a range of biological functions, including energy storage, cell structure and signalling. These have a relatively simple structure in comparison to proteins and DNA, and are typically hydrocarbons formed of ketoacyl and isoprene groups. These fall within the broader category of metabolites as biomolecules that are mediated by proteins in many molecular processes.

Larger molecules such as triglyceride (the main constituent of body fat and common in many foodstuffs, *x*) are readily hydrolysed to release component fatty acids (*xi*). While these alone cannot reveal the source of archaeological lipids, compound-specific carbon isotope ratios ($\delta^{13}C$) can be used to discriminate ruminant dairy lipids from adipose lipids. The potential long-term stability of lipids also provides ideal suitability as a range of biomarkers, for example coprostanol is frequently used as an indicator of faeces (xiii).



Relative abundance of DNA, Proteins and Lipids in life. Percentage biomolecule composition for a single cell (*E. coli*, *xiv*) and for a healthy human (70kg body mass, *xv*). Based on data from the BioNumbers database and calculated by the authors.

Ancient biomolecular analyses are inherently (if minimally) destructive, an important consideration in designing sampling strategies. Unlike conventional biomolecular research, where undegraded samples are abundant, archaeological samples are often irreplaceable. Optimising findings obtained from any destructive sampling has led us to produce the chapter in the current format, centred around different preservational substrate types. Table 1, while not comprehensive, provides an overview of the applications of biomolecular data within archaeological research, informing the choice of sampling approach for desired research questions. When sampling, be aware that left-over residues or supernatants from one approach can potentially be stored and utilised in others (see below).

Archaeological Applications of Ancient Biomolecules

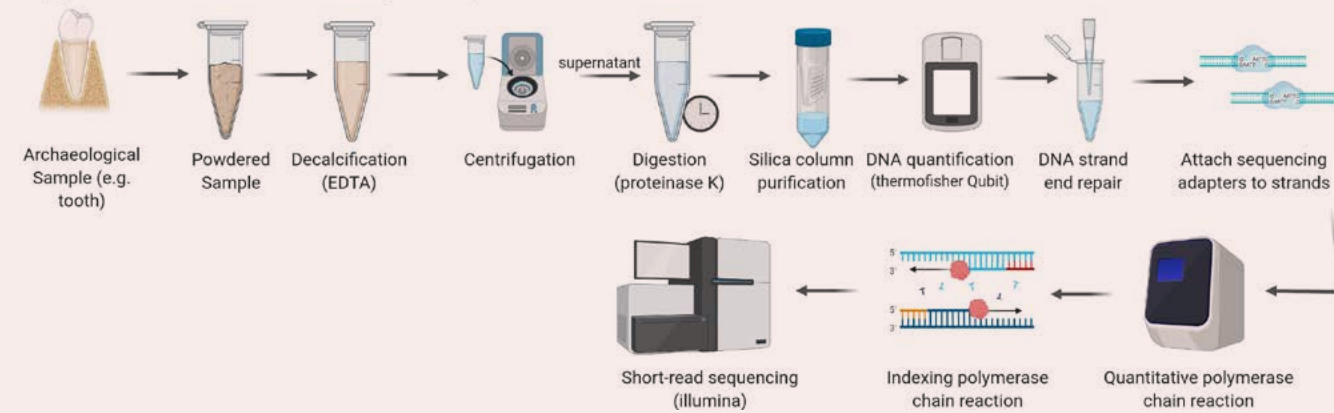
Desired Evidence of:	Potentially Suitable Data:	Uses, Pros and Cons:
Species Identification (e.g. faunal assemblages or worked animal products)	Peptide Mass Fingerprinting (ZooMS)	Rapid and inexpensive solution for species ID, potentially minimally destructive. Limited taxonomic specificity for some genera; only applicable to animal remains.
	DNA barcoding sequence analysis or mitochondrial DNA sequences	High level of accuracy and taxonomic precision. Significantly more costly and time-consuming than ZooMS.
Biological Identity and Kinship	Enamel proteome sexing	X and Y amelogenin identification can be used to sex individuals. Useful for highly fragmented remains. Male sex attributions (XY) are more secure than female (XX) due to the possibility of poor preservation removing evidence of Y variant. Does not work for all mammals, optimal for primates.
	Uniparental marker DNA (Y-chromosome and mitochondrion)	Less costly sequencing option, and suitable for poorer DNA preservation for mitochondrial sequences. Only provides patrilineal or matrilineal information.
	Whole genome sequence data	Provides greatest available amount of evidence for an individuals' identity by descent relative to others; costly and dependent on sufficient endogenous DNA preservation.
Diet*	Protein sequences in residues (e.g. on ceramics) and well preserved food remains	Facilitates both tissue and taxonomic specificity, and authentication as ancient material. Proteins survive well, especially when bound in a mineral matrix, and allows identification of mixed foods. Extraction potentially difficult from ceramics, mechanisms of protein preservation in this context still poorly understood, expensive.
	Lipid residues and paleo faeces	Can discriminate plant or animal and in some cases be species-specific. More information can be obtained from single compound isotope analyses (e.g. ruminant dairy). Faecal biomarkers can provide information on the proportions of meat and plant intake in the diet. Inexpensive sample screening.
	Metabarcoding or Metagenomics	Species identification with high specificity, from mixed or morphologically degraded material. Resource costly, and dependent upon suitable preservation.
Animal and Plant Domestication	Genetic data (whole or partial genome)	Data suitable for domesticate-associated selection, bottlenecks, and recent phylogenetic profiling. This requires good reference genomic data for taxa, and nuclear genome methods required are costly.
	Proteomics or peptide fingerprint data	Peptide fingerprinting can discriminate early domesticates from morphologically similar wild-types. Proteomic data is also suitable for phylogenetic reconstruction (though less accurate than nuclear genome), and can reveal early use of domesticate products. Considerably better preserved than DNA, especially in hot environments in the Neolithic. Protein databases lack many plant sequences.
	Lipid residues	Identification of early domesticate products including dairy and beeswax.
Environmental Context	Environmental aDNA	Metagenomics provides ecological profiling of palaeoenvironments, for animal taxa and microbial communities. Unable to provide reliable quantitative data for taxa within ecosystems, and rarer taxa may be missed. Risk of contamination is also much greater (especially between contexts). Environmental aDNA can provide full genomes as well (e.g. Neanderthals), though with very low coverage.
	Lipid biomarkers	Plant lipids and microbial biomarkers may provide specific information on the environmental context. Biomarkers can provide information on precipitation, pH, salinity and temperature.
Health and Disease	Genomes	Genetic diseases, and outcomes of deleterious inbreeding, as well as inference of phenotype through Polygenic Risk Scores (PRS). PRS is still subject to considerable debate for reliability.
	Metagenomics	Accurate pathogen identification and evolution; resource costly, though generated alongside genomic data in shotgun sequencing.
	Proteomes	Pathogen virulence factors and host immunity factors can provide detailed evidence of infection interactions, though requires very good preservation (e.g. within well-preserved dental calculus).
	Lipid biomarkers	Microbial biomarkers can identify certain infectious diseases and pathogens.
Population Dynamics and Movement	Uniparental DNA haplogroups (Y-chromosome and mitochondrion)	Rapid and less costly option for genetic data; facilitates inference of uniparental haplogroups. More specific inferences made on population history are less reliable than nuclear genomic data.
	Nuclear genomes	Many options available for inference of population structure and admixture, though requiring good reference datasets. Population size estimation through Bayesian skyline plots (only reliable with larger sample numbers). Resource costly, and analyses can be time-consuming.
Phylogenetic History	Genomes	Mitochondrial genomes provide suitable data for looking at taxonomic relationships deeper in time, though full genomes provide greater accuracy. Ancient genomes limited by preservation (e.g. subarctic Pleistocene fauna well studied).
	Proteomes	Address evolutionary relationships and in vivo modification: provides accurate taxonomy-level phylogenetic modelling, and is often well-preserved even under adverse conditions, though analytically costly for research. Proteomes will provide significant utility in contexts lacking DNA preservation (e.g. Pleistocene and Pliocene remains).

* see also Chapter 8, this volume

Figure 2: Figural labwork protocols. Figure created in Biorender.com

ANCIENT DNA *[i]*

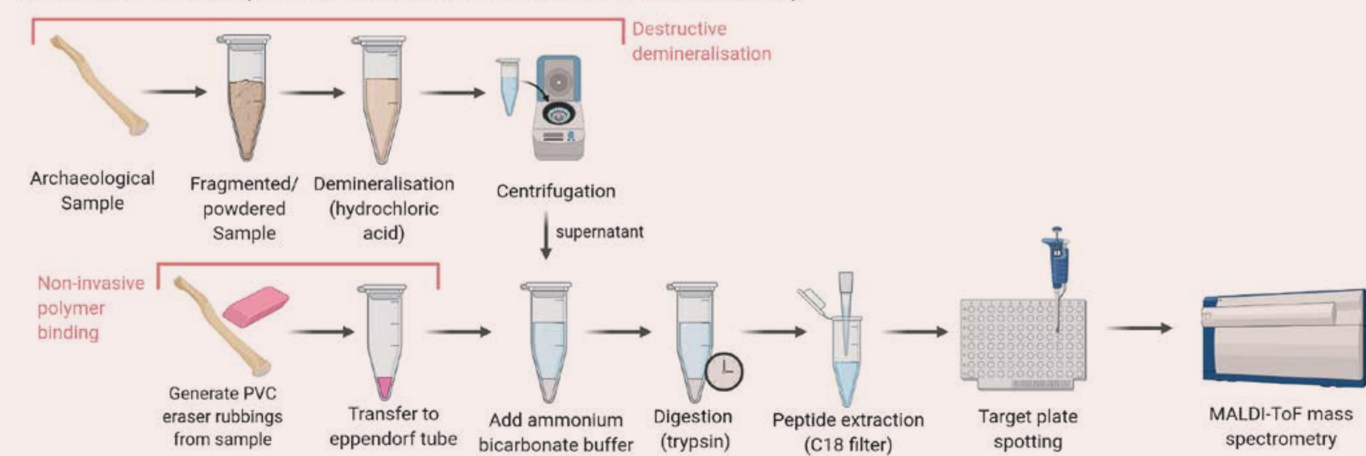
i. Typical aDNA Extraction and Sequencing Protocol



The process of ancient DNA extraction and sequencing is relatively complex, with many stages that risk introducing contamination. The requirements of costly reagents and a dedicated clean lab facility presently make research in this area prohibitively expensive for routine purposes. The dominant cost is for the sequencing required for good coverage of degraded, low-abundance ancient DNA. Currently Illumina, Inc. is the principal sequencing platform used by labs, and allows significant parallelisation of sample sequencing through multiplexing. Many laboratories will therefore wait until sufficient samples are prepared to optimise this cost before making a sequencing run.

ZooMS *[ii]*

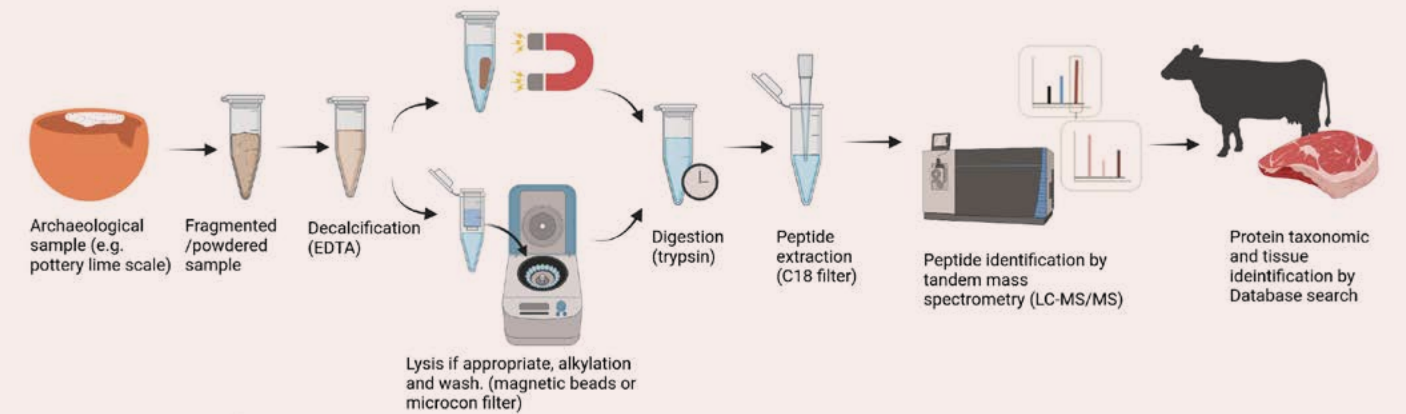
ii. ZooMS Protocol (Non-invasive and Destructive Demineralisation)



ZooMS (Zooarchaeology by Mass Spectrometry) uses MALDI-ToF-MS to obtain masses of an abundant protein, commonly type 1 collagen. Differences in the amino acid sequence appear within spectra of peptide masses, allowing species or genus identification. Analysis turnaround is very rapid (labwork can potentially be completed in a single day if necessary) and has moderate capacity for scalability. A protein quantification step (using spectrophotometry) is not included here, but can be used for calculating the volume of trypsin or determining whether mass spectrometry is worthwhile.

PALAEOPROTEOMICS *[iii]*

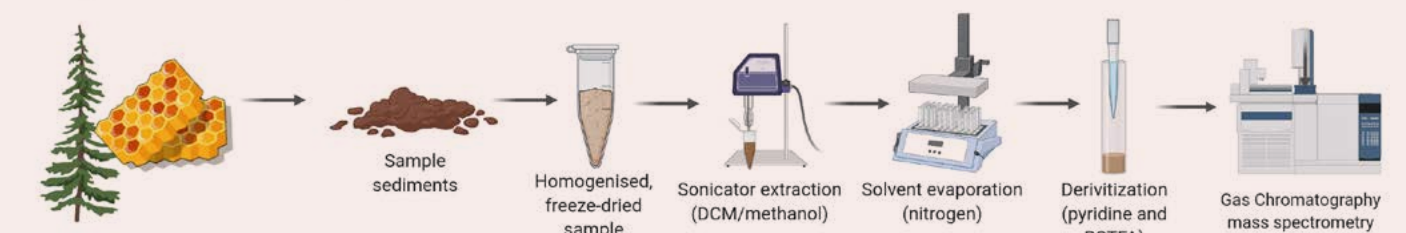
iii. Typical Proteomics protocol



Comprehensive proteomic coverage through Liquid Chromatography Tandem Mass Spectrometry (LC-MS/MS) provides data for all proteins present within a sample (rather than one abundant protein for taxonomic identification, as in ZooMS). This can reveal biological origins of samples, including specific tissues, and mixed samples such as food remains. This method is significantly more costly than ZooMS however, and requires dedicated palaeoproteomic facilities as the risk of contamination is much greater.

ANCIENT LIPIDS *[iv]*

iv. Lipid Extraction and Analysis Protocol



Lipid analysis utilises gas chromatography mass spectrometry (GC-MS) to determine the molecular structure of these organic compounds, allowing identification of the biological source of a sample. Compound Specific Isotope Analysis (CSIA) is often subsequently utilised to further elucidate the origin of lipids, especially around diet. Variations on the protocol described above may be utilised, dependent on research questions and sample characteristics.

SIZE COMPARISON *[v]*

for sampling

Pea (petit pois): ~100mg

Rice grain: ~25mg

Sesame seed: 5mg

Scale numbered in centimetres.

Mineralised Hard Tissue (Bone and Teeth)

Skeletal remains constitute the overwhelming majority of excavated organic material retained on sites, and are a principal focus for biomolecular analysis, especially for DNA and proteins.

Endogenous aDNA is significantly more likely to be gained from non vascularized mineralized tissues, as evidenced by the success of DNA recovery from fish bone. In mammalian skeletons, dental cementum or the petrous part of the temporal bone (found to be seven and eighteen times greater respectively than parietal bone, Hansen et al. 2017). Ear ossicles have also been found to show high levels of endogenous aDNA (Sirak et al., 2020). Recently, this has seen significant efforts to sample many important petrous, ossicles, and tooth specimens. Given the value of these to other analyses, destructively sampling should be carefully weighed against other research priorities within project designs.

ZooMS (Buckley et al., 2009) provides rapid taxonomic identification through peptide mass fingerprinting, and though limited for taxon specificity, can be performed with minimal sample masses, or even simply on synthetic polymers that have been in contact with the sample (McGrath et al., 2019). More recently, SPIN may provide a more comprehensive sequence-based means of rapidly identifying taxa (Rüther et al., 2021). For proteomics, dental enamel has been demonstrated to be an excellent substrate for preservation, allowing phylogenetic resolution from the early Pleistocene (Cappellini et al., 2019) and biological sexing, which can be achieved using minimally invasive methods (Stewart et al. 2017).

Key Sampling Details

(ZooMS here and subsequently refers to Peptide Mass Fingerprinting, and proteomics refers to LC-MS/MS)

	aDNA	ZooMS	Proteomics	Lipids
Vulnerability to contamination during excavation	Medium – utilise face masks and gloves	Low	Medium – consider using gloves, avoid processing finds near food, avoid protein-containing clothes (wool, silk), avoid protein-containing conservation materials (latex gloves, fish glue).	Medium – consider using gloves and sterile or clean sampling tools. Avoid sampling near food or plastics.
Typical indicators of suitable preservation	Samples retain original bone-like characteristics and are not soft or crumbly, or diagenetically altered.	Samples are not diagenetically altered or highly desiccated or crumbly. Especially suitable for diaphyseal bone.	As for aDNA and ZooMS	Presence of other organic compounds. Well preserved tissue. 'Greasy' feel.
Typical mass of sample required	*Petrous or cementum*: 50mg or less (2 grains of rice) *Other bone*: 100mg or less (1 pea or less)	*Noninvasive*: ~0mg (requires visible bone surface and very good preservation) *Demineralisation protocol*: 50mg or less (2 grains of rice) *Empty tube:* re-extract from tubes used to prepare collagen for other analyses	Dependant on analyses: ~10-100mg	Dependant on analyses: ~10-200mg
Optimal sample storage practices	Keep samples within two or more unused sealed plastic sample bags, or an eppendorf tube for smaller samples. Monitor humidity within containers and store in cold conditions if possible, avoiding sunlight exposure. Avoid storing for extended periods in conditions where condensation forms inside sample containers.	No special storage is required. Place the sample in a plastic bag or tube. The sampling container may be extracted for collagen.	Store in clean or sterile sealed containers at room temperature.	Store samples in clean (ideally combusted) aluminum foil, preferably in cold conditions. Samples can be then stored in cloth bags, sterile containers or unused plastic bags. If stored in plastic, monitor humidity and integrity of the aluminum foil.

Guidance for Hard Tissue sampling for proteins and DNA	
Drilling for bone powder:	A clean, small diameter drill bit (e.g. 2-4 mm) can be used to drill into the compact bone. Drilling should be done in short intervals at low revolutional frequency, as introducing excessive heat will cause collagen to gelatinise. One point of caution when producing bone powder is that it can become electrostatic. If it becomes static, it can be difficult to handle. When drilling, it is advisable to have tin foil placed underneath to capture the extracted bone powder. When enough is obtained (20-50 mg), the tin foil can be wrapped like an envelope and placed in a ziplock bag.
Sawing for bone chips:	A rotary saw like a Dremel can be used to obtain a solid bone chip (20-50 mg). Here it is of less importance that some heat is produced, since heat is produced in a relatively small area. However, less heat is advisable. The obtained bone chip can then be placed in a plastic tube or in tin foil in a ziplock bag until further analyses.

Skins, claws, hooves and horns

That many keratinous tissues are non-vascularised suggests a source of resistance to microbial degradation via this route, and articles such as wool and hair can be sampled with minimal visual damage. However, untreated keratinous tissues are scarce in the archaeological record, and often especially culturally or morphologically valuable. The underlying biogenesis of keratinous tissues is fundamentally similar; keratinocytes dividing within the cuticle or follicle undergo cell death to form the completely keratinised structure above. This results in optimal DNA sources at the root of keratinised structures (where active cells were situated), though the robust keratin structure itself may also provide effective preservation (e.g. hair shafts). The structure of feathers is considerably more complex, with a fractal-like geometry of rachis, ramus and barbules. Biomolecular preservation in these is still not well-characterised, though the calamus (root) is potentially most suitable.

Protein fingerprinting (Hollemeyer et al., 2002) and sequencing can differentiate morphologically similar hairs (such as guanaco and vicuña, and domesticates llama and alpaca (Azémard et al., 2021) or even distinguish individuals (Macri et al., 2020). Proteomics readily provides species identification for animal derived textiles: wool, down, silk, etc. (Solazzo, 2019), though is especially vulnerable to contamination from modern fabrics. Larger keratin structures (claws, horns, scales and scutes) can also provide excellent biomolecule preservation due to their dense non-vascularised composition, providing high taxonomic specificity through proteomics (or genomics where also suitable).

Parchment and leather are collagen-based skin products. Leather is formed by application of tanning agents to animal skin, utilising either metals-based or plants-based chemicals to increase physical durability. Leather often survives well in cold, waterlogged archaeological contexts as a result, and is often suitable for species characterisation through protein analysis (including non-invasive ZooMS if samples are not considerably degraded). DNA preservation is much more variable; if DNA survives tanning, the increased material robusticity may serve to preserve DNA well, though it cannot be easily predicted if this is the case. Lipid analysis and GC-MS may also reveal the tanning process. Parchment is highly amenable to non-invasive ZooMS (where peptides can be extracted simply by making contact with a synthetic polymer like a PVC eraser (Fiddyment et al., 2015), due to its high collagen content. Larger quantities of eraser rubbings ("erdu") may also be suitable for DNA extraction as well.

Key Sampling Details

	aDNA	ZooMS	Proteomics	Lipids
Vulnerability to contamination during excavation	Medium – utilise face masks and gloves, and hairnets or hoods if sampling hair or textiles.	Low – take multiple sets of eraser rubbings from the same site for noninvasive sampling, as first sets will be most contaminated.	Medium – Avoid wearing protein-containing clothing (wool, silk), avoid protein containing conservation materials (wool or silk thread, latex gloves)	High vulnerability to extraction of plasticizers (if plastic materials used for collection and storage)
Typical indicators of suitable preservation	Lack of diagenetic alteration or total disintegration.	Avoid highly conserved parchment. Still flexible leather is still suitable for eraser sampling, if spalling, use loose fragments.	If tissues are preserved and not diagenetically altered, keratin / collagen will be present.	Avoid sites with excessive water flow (e.g. unsaturated zone of soil), extreme antiquity and high temperatures. Sunlight + oxygen will promote free radical attack.
Typical mass of sample required	*Hair*: 100mg or less (1 pea or less) *Claw, horn, scales etc.*: 50mg or less (2 grains of rice) *Leather*: 50mg or less (2 grains of rice) *Well-preserved parchment*: non-invasive extraction from eraser rubbings (150–250 µl)	*Leather*: 25mg or less (1 grain of rice) *Parchment*: 50µl eraser rubbings (approx. volume)	*Hair*: very little as it almost all protein (potentially < 1 µg), the issue is sample handling of such tiny pieces.	Lipids associated with sheen (e.g. lanolin, fatty acids) are more abundant than sterols/ hormones. K_{ow} will predict selective leaching in aqueous environments.
Optimal sample storage practices	Store sampled material in eppendorf tubes within sealed plastic bags, preferably in cold or freezing conditions (avoid humidity).	*Leather*: As for DNA, but can be stored in plastic bags. *Rubber erasings*: store within eppendorfs at room temperature	Store in sealed sterile containers, ideally at cold temperatures.	Store samples in clean (ideally combusted) aluminum foil, preferably in cold conditions.

Other Mineralised Substrates (Calculus, Shell, Eggshell)

Mineralised substrates preserve far better than soft-tissue, and rarer mineralised materials may provide significant biomolecular insight. In particular, dense mineral matrices can provide an ideal environment for protein preservation, for up to millions of years (Demarchi et al., 2016). Recently, dental calculus has especially been a key focus – its formation entraps biofilms within hardened calcium phosphate, providing an ideal preservational environment for the oral microbiome and meta-proteome. Dental calculus should be sampled directly from the tooth within a laboratory environment, if possible; it is not necessary for teeth to be disarticulated for this. During excavation, minimal or no cleaning should be undertaken of samples. Lipids in dental calculus are also useful in the characterization of the oral health and metabolome in historic samples (Velsko et al., 2017). Combined biomolecular analyses from calculus may be especially informative for studying archaeological pathogens, providing both metagenomic profiling of microbial diversity and direct evidence of pathogenic activity through bacterial and viral proteins. Proteomic and genetic analyses may also provide evidence of diet (together with identifiable plant micro-remains). Internal lithiasis pathologies (e.g. gallstones) are not well-studied, though considerations for sampling are likely similar to dental calculi.

Molluscan shells are often robustly preserved in archaeology, featuring notably in midden deposits. DNA preservation within shells is variable dependent on the specific biomineral structure and impacts of erosion, though potentially highly effective under stable conditions (Sarkissian et al., 2020). Metagenomic analyses may also identify pathogenic infections within shells, though generally DNA extractions are not applied unless for specific questions around species' ecology, due to the associated expense. Proteomics provides a reliable means of taxonomic identification, especially for worked artefacts that are otherwise not identifiable (Sakalauskaite et al., 2019). The spire and apex are especially morphologically informative in helical shells, and sampling these should be avoided. Hard eggshell (from birds and reptiles), if preserved in intact fragments, may provide phylogenetic data through proteomics over considerable time-periods (through protein binding sites within the calcium carbonate matrix), well into the Pliocene and contemporary to hominin evolution (Demarchi et al., 2016). DNA is also well-preserved within eggshell (for example 125 times less microbial DNA was obtained in shells than in bird bones from the species (Oskam et al., 2010) due to the protective crystalline structure, though considerably less so than protein. Sample cleaning within the field should similarly be avoided, particularly using liquids.

Key Sampling Details

	aDNA	ZooMS (eggshell only)	Proteomics	Lipids
Vulnerability to contamination during excavation	Medium – utilise face masks and gloves for sampling, and avoid cleaning samples as much as possible.	Low (unless extremely old, in which case, Medium: use nitrile gloves and avoid wearing protein-containing clothes)	Medium; consider utilising gloves and face mask for sampling. Avoid processing in places where food has been present. Avoid protein containing conservation materials (lacquer, fish glue, latex gloves). Avoid protein containing clothing (wool, silk).	Medium – consider using gloves and sterile or clean sampling tools. Avoid sampling near food or plastics
Typical indicators of suitable preservation	*Dental calculus*: present and not diagenetically altered. *Molluscan shell*: original lustrous characteristics, not crumbly. *Eggshell*: not diagenetically altered or excessively fragmented.	Present, and diagenetically unaltered. Eggshell within archaeological contexts can be extremely fragmented and requires sieving for retrieval, but this is still suitable for ZooMS.	Generally good preservation. Not substantially diagenetically altered.	Generally good preservation. Not substantially diagenetically altered.
Typical mass of sample required	*Dental calculus*: 25mg or less (1 grain of rice) *Molluscan shell*: 100mg or less (1 pea) *Eggshell*: 50mg or less (2 grains of rice)	1-5mg (1 sesame seed or less)	2-20mg (½ to 4 sesame seeds)	Dependant on analyses: 20-100mg (smaller quantities required for Py-GCMS)
Optimal sample storage practices	Store whole un-subsampled samples uncleaned in sealed containers, preferably in cold conditions. Monitor humidity within containers.	Store in clean or sterile sealed containers at room temperature.	Store in clean or sterile sealed containers at room temperature.	Store samples in clean (ideally combusted) aluminum foil, preferably in cold conditions.

Plant Remains

Plants are biologically very different to animals, containing significant quantities of polysaccharides such as cellulose and lignin, and lacking mineralised components that might be resistant to decomposition. Nonetheless, waterlogged, desiccated and carbonised remains have been utilised, though typically with poorer biomolecular preservation than animal remains. Preservation by freezing or desiccation generally provides the best likelihood of biomolecule extraction.

For DNA extraction, seeds have been the main target of analysis, though given their small size, this generally requires the destruction of the entire seed. Destruction of the entire sample has led to concerns over replicability of analyses, although with current contamination detection methods, sequencing results can be reliably verified. Extractions from wood have been more limited, though well-preserved water-logged material in northern latitudes from the Neolithic and subsequently have provided chloroplast haplotype data (although the utility of this for archaeological questions is in most cases tenuous, as morphological preservation is sufficient for taxonomic identification). Though data is limited, sapwood provides considerably better endogenous preservation than heartwood (Wagner et al., 2018). Protein analysis (amino acid composition and racemisation rate) has also been applied upon archaeological seeds including grains, providing data on molecular taphonomy and taxonomy (Cappellini et al., 2010). Protein taxonomy can corroborate archaeobotanical evidence, or identify plant species if necessary.

Most plants have a protective waxy coating mainly composed of a complex mixture of long straight-chain hydrocarbons, alkanols and alkanoic acids, which are often well preserved in plant material, ceramics or sediments. Although these provide less taxonomic specificity, leaf waxes and terpenoids can be very useful to identify the vegetation when fragments are not morphologically identifiable or DNA and proteins are not well preserved.

Key Sampling Details

	aDNA	Proteomics	Lipids
Vulnerability to contamination during excavation	Medium – utilise gloves and facemask while sampling	Medium – consider utilising gloves. Avoid processing in places where food has been present. Avoid protein containing conservation materials (lacquer, fish glue, latex gloves). Avoid protein containing clothing (wool, silk).	Medium – consider using gloves and sterile or clean sampling tools. Avoid reusing tools between samples and processing samples near plants, food or plastics.
Typical indicators of suitable preservation	*Seeds*: retain original characteristics, or are not fully carbonised or crumbly. *Wood*: retains original wood-like characteristics; not diagenetically altered.	*Seeds*: retain original characteristics, or are not fully carbonised.	*Seeds and wood*: not fully carbonised; not diagenetically altered
Typical mass of sample required	*Seeds*: ~25mg (1 rice grain) or whole seed. *Wood*: ~0.5cm^3 solid fragment (volume)	*Seeds*: <5mg (<1 sesame seed)	*Seeds*: whole seed. *Wood*: ~1cm^3 solid fragment (volume)
Optimal sample storage practices	Store remains in sterile containers and keep at a cool temperature. Store waterlogged wood in original waterlogging water if possible, similarly at cold temperature.	Store remains in sterile containers and keep at a cool room temperature.	Store samples in clean (ideally combusted) aluminum foil, preferably in cold conditions. Samples can be then stored in cloth bags, sterile containers or unused plastic bags. If stored in plastic, monitor humidity and integrity of the aluminum foil.

Other Organic Substrates

When working in the field, be creative in considering the possibility of trapped biomolecular records. Natural products can provide a wealth of biomolecular data, especially as many are composed of stable hydrophobic substances (such as beeswax or resins) that provide for the entrapment of biomolecules. Beeswax is widely used and commonly recovered in the archaeological record as a sealant, although may be far more abundant in settlements than appreciated. The Kenward method[1] applied to strange soft lumps in sediments may yield a rich treasure trove of organic materials. DNA extracted from modern (Modi et al., 2021) and Mesolithic resins has – in the latter case (Jensen et al., 2019) provided evidence of the chewer and the chewed; attempts to extract DNA and proteins from fossil plant resins have been less successful. Metagenomics of organic substrates may characterise microbial communities and species diversity within samples, particularly where these originate or are utilised by animals, and can reveal evidence of diet or internal microbiome.

1 Beeswax can be identified in the field when "gently touched with a heated needle… the characteristic and delightful aroma of beeswax can quite distinctly be smelled" (Kenward, 1991, p. 3).

Adipocere and other preserved soft tissues may provide an excellent source of biomolecular information if well-preserved, particularly for microbial decomposition (the necrobiome). Highly desiccated remains (such as mummified Ancient Egyptian samples) often provide poor endogenous DNA preservation, however, due to thermal exposure, are usually rich in lipids. Tissue-specific proteomic data can also be applied for evidence of health and inference of post-mortem intervals.

Key Sampling Details

	aDNA	Proteomics	Lipids
Vulnerability to contamination during excavation	High potential for surface contamination; avoid cleaning samples using any liquids, and use sterile gloves. Avoid drilling.	Probably low, insufficient data to assess. Nonetheless, utilise precautions such as sterile gloves, minimal cleaning or exposure to liquids.	Medium – consider using gloves and sterile or clean sampling tools. Avoid reusing tools between samples and processing samples near food or using plastics.
Typical indicators of suitable preservation	Hard to determine from physical sample alone – sequencing needs to be carried out in order to check for preservation.	Insufficient data to assess	Dependent on material. High preservation potential in waxy tissues such as beeswax and adipocere. Lower preservation potential in desiccated remains.
Typical mass of sample required	20-30mg (~1 grain of rice)	Insufficient data to assess; highly variable yield dependent upon preservation.	Dependent on analyses. 30-250mg (3-5 grains of rice).
Optimal sample storage practices	In sterile sealed containers	Sealed container, at room temperature.	Store samples in clean (ideally combusted) aluminum foil, preferably in cold conditions.

Residues on Artefacts

Both lipid analysis and proteomics can be applied to residues on artefacts to identify their composition, most frequently for probable food residues and well preserved organic remains in vessels. Preserved lipids have provided extraordinary insight regarding ancient diets and subsistence strategies. Organic residues can be extracted from *in situ* preserved contents, visible residues that survived in the inner or outer surface of the vessel or as absorbed residues in the ceramic matrix (Evershed 2008). The combination of organic residue analysis with compound specific isotopic analysis is capable of identifying a wide range of commodities, including dairy, ruminant carcass fats, non-ruminant carcass fats, marine fats, and waxes or resins derived from plants. However, contamination in the field or the lab with fats from the hands, food or plastics can lead to the misidentification of certain biomarkers and spurious results, and food mixing causes identification issues.

Since the relatively-recent inception of palaeoproteomics, several metaproteomic matrices have been investigated. Particularly promising materials for proteomic analysis

are well-preserved organic substrates preserved by arid, cold or waterlogged conditions. In the absence of such remains, while proteins have been identified in the ceramic matrix itself in a few studies (Hendy et al., 2018; Solazzo et al., 2008), a more accessible material is limescale residue, where proteins can be readily extracted and survive well (Hendy et al., 2018). Proteomics can often detect both the taxonomy (eg. sheep) and tissue (eg. milk) present in residues on artefacts such as ceramics, as well mixtures of different proteinaceous ingredients. In certain cases proteomics has also revealed food preparation practices such as sourdough bread and kefir-making (Shevchenko et al., 2014; Yang et al., 2014). Contamination in the field and lab is similarly important for proteomic analysis.The presence of certain degradation patterns can sometimes differentiate evidence for ancient and modern proteins (Ramsøe et al., 2020), although this is not always possible (Ramsøe et al., 2021). As such contamination from modern food, protein-containing clothing (wool, silk) and labelling/conservation materials (lacquer, fish glue, latex gloves) should be avoided, as should cross-contamination between ancient samples. If possible, cleaning and excessive handling of artefacts should be avoided, and sampling in the field may be appropriate depending on the size, fragility and scarcity of the sample in question. Sampling involves carefully scraping the interior surface of the ceramic to remove a sample of adhering residue.

Metagenomic DNA analysis similarly allows identification of food residues with high phylogenetic accuracy, though lacking the tissue-specificity afforded by proteomics. DNA extraction is also generally less suitable for samples from hotter and old environments. Metagenomics provides detailed evidence of microbial communities within samples, which may be especially relevant for fermented foods, as well as molecular taphonomy.

Key Sampling Details

	aDNA	Proteins	Lipids
Vulnerability to contamination during excavation	Medium-High – utilise face masks and gloves; sterilise any sampling equipment.	Medium-High; consider utilising gloves. Avoid processing in places where food has been present. Avoid protein containing conservation materials (lacquer, fish glue, latex gloves) Avoid protein containing clothing (wool, silk)	Medium-high – consider using gloves and sterile or clean sampling tools. Avoid sampling near food or using plastics.
Typical indicators of suitable preservation	Suitable for trace quantities of residue, though only up to a couple millenia, or within stable cold conditions.	Internal lime scale visible in pot.	Visible residue or foodcrust inside the container.
Typical mass of sample required	25-50mg (1-2 grains of rice)	10-100mg	1-2g
Optimal sample storage practices	Store remains in sterile containers (e.g. eppendorf tubes), preferably at cold temperature.	Store remains in sterile containers.	Store samples in clean (ideally combusted) aluminum foil, preferably in cold conditions.

PROTEOMIC DATA FROM FOOD RESIDUES

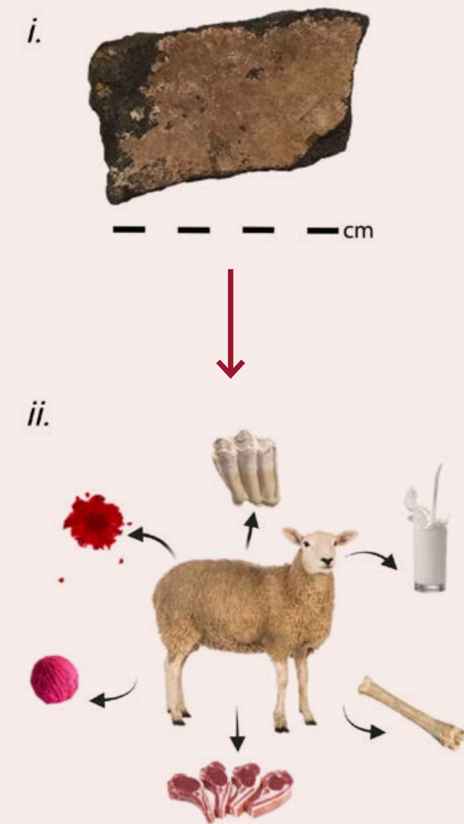


Figure 3: Foodcrusts and tissue specific data from preserved proteins. Figure created in Biorender.com

(*i*) An example of exceptional limescale preservation on a pottery sherd, from a vessel used for food preparation.

(*ii*) Proteins obtained from the residue on this sherd can provide both taxonomic and tissue-specific information.

Environmental Samples (Sediments and Ice cores)

Biomolecules from environmental samples have provided considerable new insight for palaeoecological profiling, allowing detection of organisms solely by their biomolecular traces. This mode of sampling is one of most vulnerable to contamination, however, and sampling should be undertaken with precaution.

Ancient environmental DNA (eDNA) is especially susceptible to contamination, both from modern DNA and from cross contamination between horizons that would be impossible to subsequently detect. Trace substances, like unique plasmid DNA, can be applied to exposed surfaces and equipment to determine whether contaminants may have entered samples at any point during sampling or storage. Laboratory strains of bacteria have also been used to reveal contamination during coring, by distributing these around coring sites, although such precautions should be judged within the specific context of analyses. More recently though, with better contamination detection methods, subsampling of extracted cores within the lab has provided sufficient results. Archaeological sediments should be carefully sampled using full protective equipment (see below), with samples sealed inside collection tubes as quickly as possible once complete. In sampling from lake sediments, cores obtained should be plugged and taped immediately, and sent directly to a laboratory for subsampling. Similarly, ice cores should be kept frozen, and sub-sampling may then be made from internal contents within laboratories. Proteomic sampling of sediments is presently less commonplace, though has been explored in a few studies, with potential to reveal allochthonous proteins from animals within the local environment (Oonk et al., 2012), and soil microbe activity within sites. Molecular imprints of fabrics can potentially also be detected within sediments (Li et al., 2021).

Similar considerations should be made for lipid analysis, although full protective equipment is not often required. Lipids from plants, microbes and other organisms can survive in sediments and rocks for millions of years (Peters et al., 2004). Their analysis can provide unique insights into the environmental context, climate (e.g.: precipitation and temperature), organisms that were present, fecal input from animals and humans and elucidate the function of different areas within the archaeological site (Sistiaga et al., 2020). Material can be subsampled from cores in the lab or directly sampled from archaeological sites or geological stratigraphic profiles, using clean tools (non-plastic) and avoiding mixing sediments from different layers.

Key Sampling Details

	aDNA	Proteomics	Lipids
Vulnerability to contamination during excavation	High – wear sterile gloves, protective sleeves, shoe covers and face masks, avoid loose-fitting clothes. Full cleansuit gowns should also be worn if possible.	High, as eDNA	High – consider using gloves and sterile or clean sampling tools. Avoid reusing tools between samples and processing samples near plants, food or plastics.
Typical indicators of suitable preservation	Stable cool temperature within the environment.	As eDNA	High preservation potential.
Typical mass of sample required	Variable (typically subsampled within lab). Generally between 3ml and 50ml volume per sample.	Insufficient data to assess.	Dependent on age. *Sediment*: 15gr or less.
Optimal sample storage practices	*Sediment*: excavate sample into sterile conical centrifuge tubes, seal in bags and store at cold temperature, refrigerated or continuously frozen, if possible. *Lake sediments*: Use a sterile PVC pipe or similar for coring, that should be sealed immediately and cold-stored if possible.	As eDNA	Store samples in clean (ideally combusted) aluminum foil, preferably in cold conditions. Samples can be then stored in cloth bags, sterile containers or unused plastic bags. If stored in plastic, monitor humidity and integrity of the aluminum foil.

Digested Remains (Coprolites, Abdominal Sediments, Cesspits)

Vertebrate faeces contain a series of lipids known as faecal biomarkers, which have been used to ascertain the diet and origin of the remains (Bull et al., 2002; Sistiaga et al., 2014). Differences in the proportions of these steroidal biomarkers are a consequence of dietary preferences, endogenous biosynthesis of cholesterol and the capacity to produce these metabolites by gut bacteria (Prost et al., 2017). With continued methodological refinement, lipidomic and broader metabolomic data can provide significant insight into nutritional metabolics and health in the past. Sample collection for the analysis of lipids and metabolites from coprolites, abdominal sediments and cesspits should be conducted following similar considerations as with environmental samples and other organic substrates. Coprolites and palaeo-faeces should not be washed or handled without gloves. If possible, subsampling should be done within laboratories.

Considerations for aDNA sampling within coprolites and abdominal sediments are broadly similar as for environmental DNA above (but without necessarily such strict anti-contamination protocols as tracing allochthonous sequences from animals are not a principal target). DNA can provide mitochondrial or low-coverage nuclear genomes for the originating individual, as well as metagenomic evidence of diet, gut microbial diversity and pathogens. The latter can provide an additional level of insight into paleoparasitology, and has recently been explored more in tropical climates (with greater pathogen rain), though within environmentally stable settings (such as caves). Proteomics has recently been applied to coprolites to investigate the diet and also reveal the host proteome (Runge et al., 2021), though these are permafrost-preserved, and further research is required to understand the utility of proteins under other conditions.

Key Sampling Details

	aDNA	Proteins	Lipids
Vulnerability to contamination during excavation	Medium-High: wear sterile gloves, protective sleeves and face-masks; prevent sampled material coming in contact with any external liquids or solids.	Medium-High; consider utilising gloves. Avoid processing in places where food has been present. Avoid protein containing conservation materials (lacquer, fish glue, latex gloves) Avoid protein containing clothing (wool, silk).	Medium – consider using gloves and sterile or clean sampling tools. Avoid reusing tools between samples and processing samples near food or using plastics.
Typical indicators of suitable preservation	Cool, stable, undisturbed environments especially favourable for good preservation.	Permafrost; otherwise insufficient data to say.	High preservation potential in most cases.
Typical mass of sample required	Variable, dependent on research aims; 50mg - 250mg (3 peas or less).	Insufficient data to assess (150mg used by Runge et al. 2021).	Dependant on age and material. *Paleo faeces*: 150-500mg (less material required if analyzed using Py-GCMS). *Abdominal sediments and cesspits*: 5-15gr
Optimal sample storage practices	Store sample in sterile container (sealed tube or bag) in refrigerator or freezer until subsampling or analysis.	Store as close to the condition in which the sample was found as possible (e.g. frozen for permafrost remains).	Store samples in clean (ideally combusted) aluminum foil, preferably in cold conditions. Samples can be then stored in cloth bags, sterile containers or unused plastic bags.

General Considerations

In this chapter, we have considered the range of bioarchaeological substrates, describing their distinct preservational characteristics, and the sampling considerations required of each. However, variability across and within archaeological contexts, coupled with the cryptic nature of the deposition process, will always be a significant issue for biomolecular recovery, and impossible to take account of in a chapter of this length. Options for sample screening to mitigate this are limited outside the lab, beyond gross preservation. However, portable Near Infrared (NIR) spectroscopy has recently presented one suitable option to detect collagen (Sponheimer et al. 2019). More generally, the inherent variability of archaeological preservation means that decisions on sampling appropriateness are best made by those who are closely familiar with the specific circumstances of the site, working with those conducting analyses, rather than by a set of generic guidelines. The risk of contamination within samples should also be judged in the same way, with a primary requirement being that as much information as possible is openly presented and available, so the wider community can form their own assessments of results rather than trying to rely on 'foolproof' specific protocols or guidelines (Gilbert et al., 2005).

For materials that are inherently rarer or only found in small quantities (e.g. dental calculus), prioritising maximum data retrieval from destructive sampling should always be considered. Prior to destructive analyses, capture of visual data by photography, surface-model scanning or micro-CT scanning is important for preserving morphological data. In the lab, there is considerable potential to combine methods, including with other approaches that make use of preserved biomolecules, such as stable isotope analysis and radiocarbon dating. The latter two require considerably greater volumes of collagen than ZooMS, which can be undertaken simply using traces of extracted collagen leftover in tubes. ZooMS is also applied to identify fragmentary bone from specific taxa (e.g. human) for subsequent aDNA analysis (Douka et al., 2019). Analyses can be methodologically combined, for example in simultaneous extractions for aDNA and proteins (Fagernäs et al. 2020 - also plant microremains). Future developments may make it possible to apply a range of methods for different biomolecules to the mass of material currently required for only a single analysis. However, this prediction should not be seen as an encouragement to stop destructive sampling of material until some unspecified point in the future, as the development of this field, and archaeology more widely, requires the constant injection of new information to develop new hypotheses and approaches. Neither would we advocate requiring multi-substrate analyses after sampling if the archaeological questions can be addressed by only a single biomarker. Combining analyses should be seen as a way to recover more information from archaeological materials, not restrict analyses to only a small set of well-resourced labs.

Finally, ancient biomolecular sampling will always contain an element of unpredictability, due to the stochasticity of preservational environments. Heterogeneity in preservation is also apparent in individual samples. Even if all sampling advice here is scrupulously followed and analyses are undertaken flawlessly in a state-of-the-art lab, there is still a real risk that samples will fail to produce any results due to a lack of biomolecular preservation. This risk should not be treated excessively cautiously; it is similarly possible that apparently poorly preserved remains may occasionally produce extraordinary results. In both cases, providing as much data as possible to the wider research community upon dissemination is the best way of facilitating the process of understanding and predicting how this occurs.

References

Azémard, C., Dufour, E., Zazzo, A., Wheeler, J.C., Goepfert, N., Marie, A., Zirah, S., 2021. Untangling the fibre ball: Proteomic characterization of South American camelid hair fibres by untargeted multivariate analysis and molecular networking. J. Proteomics 231, 104040. https://doi.org/10.1016/j.jprot.2020.104040

Buckley, M., Collins, M., Thomas-Oates, J., Wilson, J.C., 2009. Species identification by analysis of bone collagen using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. Rapid Communications in Mass Spectrometry 23, 3843-3854. https://doi.org/10.1002/rcm.4316

Bull, I.D., Lockheart, M.J., Elhmmali, M.M., Roberts, D.J., Evershed, R.P., 2002. The origin of faeces by means of biomarker detection. Environ. Int. 27, 647-654. https://doi.org/10.1016/s0160-4120(01)00124-6

Cappellini, E., Gilbert, M.T.P., Geuna, F., Fiorentino, G., Hall, A., Thomas-Oates, J., Ashton, P.D., Ashford, D.A., Arthur, P., Campos, P.F., Kool, J., Willerslev, E., Collins, M.J., 2010. A multidisciplinary study of archaeological grape seeds. Naturwissenschaften 97, 205-217. https://doi.org/10.1007/s00114-009-0629-3

Cappellini, E., Welker, F., Pandolfi, L., Ramos-Madrigal, J., Samodova, D., Rüther, P.L., Fotakis, A.K., Lyon, D., Moreno-Mayar, J.V., Bukhsianidze, M., Rakownikow Jersie-Christensen, R., Mackie, M., Ginolhac, A., Ferring, R., Tappen, M., Palkopoulou, E., Dickinson, M.R., Stafford, T.W., Jr, Chan, Y.L., Götherström, A., Nathan, S.K.S.S., Heintzman, P.D., Kapp, J.D., Kirillova, I., Moodley, Y., Agusti, J., Kahlke, R.-D., Kiladze, G., Martínez-Navarro, B., Liu, S., Sandoval Velasco, M., Sinding, M.-H.S., Kelstrup, C.D., Allentoft, M.E., Orlando, L., Penkman, K., Shapiro, B., Rook, L., Dalén, L., Gilbert, M.T.P., Olsen, J.V., Lordkipanidze, D., Willerslev, E., 2019. Early Pleistocene enamel proteome from Dmanisi resolves Stephanorhinus phylogeny. Nature 574, 103-107. https://doi.org/10.1038/s41586-019-1555-y

Copley, M.S., Berstan, R., Dudd, S.N., Docherty, G., Mukherjee, A.J., Straker, V., Payne, S. and Evershed, R.P., 2003. *Direct chemical evidence for widespread dairying in prehistoric Britain.* Proc. Natl. Acad. Sci. U. S. A. 100(4). 1524-1529. https://doi.org/10.1073/pnas.033595510

Demarchi, B., Hall, S., Roncal-Herrero, T., Freeman, C.L., Woolley, J., Crisp, M.K., Wilson, J., Fotakis, A., Fischer, R., Kessler, B.M., Rakownikow Jersie-Christensen, R., Olsen, J.V., Haile, J., Thomas, J., Marean, C.W., Parkington, J., Presslee, S., Lee-Thorp, J., Ditchfield, P., Hamilton, J.F., Ward, M.W., Wang, C.M., Shaw, M.D., Harrison, T., Domínguez-Rodrigo, M., MacPhee, R.D.E., Kwekason, A., Ecker, M., Kolska Horwitz, L., Chazan, M., Kröger, R., Thomas-Oates, J., Harding, J.H., Cappellini, E., Penkman, K., Collins, M.J., 2016. Protein sequences bound to mineral surfaces persist into deep time. Elife 5. https://doi.org/10.7554/eLife.17092

Douka, K., Brown, S., Higham, T., Pääbo, S., Derevianko, A., Shunkov, M., 2019. FINDER project: collagen fingerprinting (ZooMS) for the identification of new human fossils. Antiquity 93. https://doi.org/10.15184/aqy.2019.3

Evershed, R. P. (2008). Experimental approaches to the interpretation of absorbed organic residues in archaeological ceramics. World Archaeology, 40, 26-47. https://www.tandfonline.com/doi/abs/10.1080/00438240801889373

Fiddyment, S., Holsinger, B., Ruzzier, C., Devine, A., Binois, A., Albarella, U., Fischer, R., Nichols, E., Curtis, A., Cheese, E., Others, 2015. Animal origin of 13th-century uterine vellum revealed using noninvasive peptide fingerprinting. Proc. Natl. Acad. Sci. U. S. A. 112, 15066-15071. https://doi.org/10.1073/pnas.151226411

Gilbert, M.T.P., Bandelt, H.-J., Hofreiter, M., Barnes, I., 2005. Assessing ancient DNA studies. Trends Ecol. Evol. 20, 541-544. https://doi.org/10.1016/j.tree.2005.07.005

Hansen, H.B., Damgaard, P.B., Margaryan, A., Stenderup, J., Lynnerup, N., Willerslev, E., Allentoft, M.E., 2017. Comparing Ancient DNA Preservation in Petrous Bone and Tooth Cementum. PLoS One 12, e0170940. https://doi.org/10.1371/journal.pone.0170940

Hendy, J., Colonese, A.C., Franz, I., Fernandes, R., Fischer, R., Orton, D., Lucquin, A., Spindler, L., Anvari, J., Stroud, E., Biehl, P.F., Speller, C., Boivin, N., Mackie, M., Jersie-Christensen, R.R., Olsen, J.V., Collins, M.J., Craig, O.E., Rosenstock, E., 2018. Ancient proteins from ceramic vessels at Çatalhöyük West reveal the hidden cuisine of early farmers. Nat. Commun. 9, 1-10. https://doi.org/10.1038/s41467-018-06335-6

Hollemeyer, K., Altmeyer, W., Heinzle, E., 2002. Identification and quantification of feathers, down, and hair of avian and mammalian origin using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Anal. Chem. 74, 5960-5968. https://doi.org/10.1021/ac020347f

Jensen, T.Z.T., Niemann, J., Iversen, K.H., Fotakis, A.K., Gopalakrishnan, S., Vågene, Å.J., Pedersen, M.W., Sinding, M.-H.S., Ellegaard, M.R., Allentoft, M.E., Lanigan, L.T., Taurozzi, A.J., Nielsen, S.H., Dee, M.W., Mortensen, M.N., Christensen, M.C., Sørensen, S.A., Collins, M.J., Gilbert, M.T.P., Sikora, M., Rasmussen, S., Schroeder, H., 2019. A 5700 year-old human genome and oral microbiome from chewed birch pitch. Nat. Commun. 10, 5520. https://doi.org/10.1038/s41467-019-13549-9

Kenward, H.K., 1991. A sceptical view of the Coppergate" beehive. Interim: Archaeology in York 15, 20-24.

Li, L., Zhu, L., Xie, Y., 2021. Proteomics analysis of the soil textile imprints from tomb M6043 of the Dahekou Cemetery site in Yicheng County, Shanxi Province, China. Archaeol. Anthropol. Sci. 13, 7. https://doi.org/10.1007/s12520-020-01258-0

Macri, A.M., Lam, S., Powers, R.H., Marsico, A.L.M., 2020. Differentiation of Morphologically Similar Human Head Hairs from Two Demographically Similar Individuals Using Amino Acid Ratios. J. Forensic Sci. 65, 1745-1751. https://doi.org/10.1111/1556-4029.14489

McGrath, K., Rowsell, K., Gates St-Pierre, C., Tedder, A., Foody, G., Roberts, C., Speller, C., Collins, M., 2019. Identifying Archaeological Bone via Non-Destructive ZooMS and the Materiality of Symbolic Expression: Examples from Iroquoian Bone Points. Sci. Rep. 9, 11027. https://doi.org/10.1038/s41598-019-47299-x

Modi, A., Vergata, C., Zilli, C., Vischioni, C., Vai, S., Tagliazucchi, G.M., Lari, M., Caramelli, D., Taccioli, C., 2021. Successful extraction of insect DNA from recent copal inclusions: limits and perspectives. Sci. Rep. 11, 6851. https://doi.org/10.1038/s41598-021-86058-9

Oonk, S., Cappellini, E., Collins, M.J., 2012. Soil proteomics: An assessment of its potential for archaeological site interpretation. Org. Geochem. 50, 57-67. https://doi.org/10.1016/j.orggeochem.2012.06.012

Oskam, C.L., Haile, J., McLay, E., Rigby, P., Allentoft, M.E., Olsen, M.E., Bengtsson, C., Miller, G.H., Schwenninger, J.-L., Jacomb, C., Walter, R., Baynes, A., Dortch, J., Parker-Pearson, M., Gilbert, M.T.P., Holdaway, R.N., Willerslev, E., Bunce, M., 2010. Fossil avian eggshell preserves ancient DNA. Proc. Biol. Sci. 277, 1991-2000. https://doi.org/10.1098/rspb.2009.2019

Peters, K.E., Walters, C.C., Moldowan, J.M., 2004. The Biomarker Guide. https://doi.org/10.1017/cbo9781107326040

Prost, K., Birk, J.J., Lehndorff, E., Gerlach, R., Amelung, W., 2017. Steroid biomarkers revisited – improved source identification of faecal remains in archaeological soil material. PLoS One 12, e0164882. https://doi.org/10.1371/journal.pone.0164882

Ramsøe, A., Crispin, M., Mackie, M., McGrath, K., Fischer, R., Demarchi, B., Collins, M.J., Hendy, J., Speller, C., 2021. Assessing the degradation of ancient milk proteins through site-specific deamidation patterns. Sci. Rep. 11, 7795. https://doi.org/10.1038/s41598-021-87125-x

Ramsøe, A., van Heekeren, V., Ponce, P., Fischer, R., Barnes, I., Speller, C., Collins, M.J., 2020. DeamiDATE 1.0: Site-specific deamidation as a tool to assess authenticity of members of ancient proteomes. J. Archaeol. Sci. 115, 105080. https://doi.org/10.1016/j.jas.2020.105080

Runge, A.K.W., Hendy, J., Richter, K.K., Masson-MacLean, E., Britton, K., Mackie, M., McGrath, K., Collins, M., Cappellini, E., Speller, C., 2021. Palaeoproteomic analyses of dog palaeofaeces reveal a preserved dietary and host digestive proteome. Proc. Biol. Sci. 288, 20210020. https://doi.org/10.1098/rspb.2021.0020

Rüther, P.L., Husic, I.M., Bangsgaard, P., Murphy Gregersen, K., Pantmann, P., Carvalho, M., Godinho, R.M., Friedl, L., Cascalheira, J., Jørkov, M.L.S., . Benedetti, M.M., Haws, J., Bicho, N., Welker, F., Cappellini, E., Olsen, J.V., 2021. SPIN - Species by Proteome INvestigation. bioRxiv. https://doi.org/10.1101/2021.02.23.432520

Sakalauskaite, J., Andersen, S.H., Biagi, P., Borrello, M.A., Cocquerez, T., Colonese, A.C., Dal Bello, F., Girod, A., Heumüller, M., Koon, H., Mandili, G., Medana, C., Penkman, K.E., Plasseraud, L., Schlichtherle, H., Taylor, S., Tokarski, C., Thomas, J., Wilson, J., Marin, F., Demarchi, B., 2019. "Palaeoshellomics" reveals the use of freshwater mother-of-pearl in prehistory. Elife 8. https://doi.org/10.7554/eLife.45644

Sarkissian, C.D., Moller, P., Hofman, C.A., Ilsoe, P., Rick, T.C., Schiotte, T., Sorensen, M.V., Dalen, L., Orlando, L., 2020. Unveiling the Ecological Applications of Ancient DNA From Mollusk Shells. Frontiers in Ecology and Evolution 8. https://doi.org/10.3389/fevo.2020.00037

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

Sirak, K., Fernandes, D., Cheronet, O., Harney, E., Mah, M., Mallick, S., Rohland, N., Adamski, N., Broomandkhoshbacht, N., Callan, K., Candilio, F., Lawson, A.M., Mandl, K., Oppenheimer, J., Stewardson, K., Zalzala, F., Anders, A., Bartík, J., Coppa, A., Dashtseveg, T., Évinger, S., Farkaš, Z., Hajdu, T., Bayarsaikhan, J., McIntyre, L., Moiseyev, V., Okumura, M., Pap, I., Pietrusewsky, M., Raczky, P., Šefčáková, A., Soficaru, A., Szeniczey, T., Szőke, B.M., Van Gerven, D., Vasilyev, S., Bell, L., Reich, D., Pinhasi, R., 2020. Human auditory ossicles as an alternative optimal source of ancient DNA. Genome Res. 30, 427-436. https://doi.org/10.1101/gr.260141.119

Sistiaga, A., Husain, F., Uribelarrea, D., Martín-Perea, D.M., Ferland, T., Freeman, K.H., Diez-Martín, F., Baquedano, E., Mabulla, A., Domínguez-Rodrigo, M., Summons, R.E., 2020. Microbial biomarkers reveal a hydrothermally active landscape at Olduvai Gorge at the dawn of the Acheulean, 1.7 Ma. Proc. Natl. Acad. Sci. U. S. A. 117, 24720-24728. https://doi.org/10.1073/pnas.2004532117

Sistiaga, A., Mallol, C., Galván, B., Summons, R.E., 2014. The Neanderthal meal: a new perspective using faecal biomarkers. PLoS One 9, e101045. https://doi.org/10.1371/journal.pone.0101045

Solazzo, C., 2019. Characterizing historical textiles and clothing with proteomics. Conserv. Patrim. 31, 97-114. https://doi.org/10.14568/cp2018031

Solazzo, C., Fitzhugh, W.W., 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

Sponheimer, M., Ryder, C. M., Fewlass, H., Smith, E. K., Pestle, W. J., & Talamo, S. (2019). Saving old bones: a non-destructive method for bone collagen prescreening. Scientific Reports, 9, 1-7. https://www.nature.com/articles/s41598-019-50443-2

Stewart, N.A., Gerlach, R.F., Gowland, R.L., Gron, K.J., Montgomery, J., 2017. Sex determination of human remains from peptides in tooth enamel. Proc. Natl. Acad. Sci. U. S. A. 114, 13649-13654. https://doi.org/10.1073/pnas.1714926115

Velsko, I.M., Overmyer, K.A., Speller, C., Klaus, L., Collins, M.J., Loe, L., Frantz, L.A.F., Sankaranarayanan, K., Lewis, C.M., Jr, Martinez, J.B.R., Chaves, E., Coon, J.J., Larson, G., Warinner, C., 2017. The dental calculus metabolome in modern and historic samples. Metabolomics 13, 134. https://doi.org/10.1007/s11306-017-1270-3

Wagner, S., Lagane, F., Seguin-Orlando, A., Schubert, M., Leroy, T., Guichoux, E., Chancerel, E., Bech-Hebelstrup, I., Bernard, V., Billard, C., Billaud, Y., Bolliger, M., Croutsch, C., Čufar, K., Eynaud, F., Heussner, K.U., Köninger, J., Langenegger, F., Leroy, F., Lima, C., Martinelli, N., Momber, G., Billamboz, A., Nelle, O., Palomo, A., Piqué, R., Ramstein, M., Schweichel, R., Stäuble, H., Tegel, W., Terradas, X., Verdin, F., Plomion, C., Kremer, A., Orlando, L., 2018. High-Throughput DNA sequencing of ancient wood. Mol. Ecol. 27, 1138-1154. https://doi.org/10.1111/mec.14514

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. https://doi.org/10.1016/j.jas.2014.02.005