AutoZooMS: Integrating robotics into high-throughput ZooMS for the species identification of archaeofaunal remains at Grotte Mandrin, France

Ellie-May Oldfield, Mark Dunstan, Manasij Pal Chowdhury, Ludovic Slimak, Michael Buckley

AutoZooMS: Integrating robotics into high-throughput ZooMS for the species identification of archaeofaunal remains at Grotte Mandrin, France

Ellie-May Oldfield
University of Manchester
Mark Dunstan
University of Manchester
Manasij Pal Chowdhury
University of Manchester
Ludovic Slimak
University of Toulouse-Jean Jaurès
Michael Buckley (m.buckley@manchester.ac.uk)
University of Manchester

Research Article

Keywords: palaeoproteomics, zooarchaeology by mass spectrometry, automation, archaeofauna, Grotte Mandrin

Posted Date: April 6th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2762261/v1

Abstract

As the vast majority of excavated archaeofaunal skeletal remains are fragmentary to the extent that they cannot be identified by morphological analysis alone, various molecular methods have been considered to retrieve information from an otherwise underutilised resource. The introduction of collagen fingerprinting, known as Zooarchaeology by Mass Spectrometry (ZooMS), has become one of the most popular approaches to improve taxonomic data yields from fragmentary bone. However, few studies have analysed large numbers of samples. Here we test the incorporation of liquid-handling robots to further develop ZooMS into a more automated technique using samples excavated from Grotte Mandrin, France. By increasing the faunal identifications of the B2 layer (~42-44 Ka) at Grotte Mandrin from 55 to 1037, we identified a wider range of taxa, now including bear and mammoth, as well as further remains of hominin. AutoZooMS has the capacity to investigate larger proportions of archaeofaunal assemblages rapidly and cost effectively whilst requiring little human intervention, aiming to improve our understanding of the human past.

1. Introduction

The archaeological and palaeontological records are largely dominated by skeletal vertebrate remains, the biomineralised tissues of bone and teeth that on average preserve for far longer than soft-tissues (Briggs 2003). However, depending on the nature of the faunal assemblage, relatively large numbers of remains can be fragmentary at excavation before any taxonomic identification can take place. Bone fragmentation results from taphonomic and anthropogenic processes, such as dispersal and accumulation via scavengers and floods or butchery by hominins (Noe-Nygaard 1977; Denys 2002). These fragmentary remains are often unidentifiable by morphological analysis and can make up 60–70% of excavated Pleistocene assemblages (Morin et al. 2017). Thus, fragmentation renders most excavated bone taxonomically redundant and is a major driver of low taxonomic data yields from archaeofaunal assemblages. Increasing these data yields is of great interest not only to reduce archaeofaunal waste and improve the efficiency of zooarchaeology and palaeobiology, but also to minimise any bias this large proportion of unidentifiable material has caused to current understandings of matters such as, the ecology of ancient environments and ancient hominin subsistence behaviours. The growing uptake of ancient biomolecular analysis in zooarchaeology and palaeobiology has begun to reduce this waste, whereby several approaches have been used, including DNA-based (Newman et al. 2002; Kahila Bar-Gal et al. 2003) and proteomic-based methods (Buckley et al. 2009; Brandt et al. 2014; Presslee et al. 2017), the latter of which is gaining in popularity due to its compromise between cost, success rate and taxonomic resolution in relation to DNA and lipids. Species identification by collagen fingerprinting, also known as ZooMS (for Zooarchaeology by Mass Spectrometry) is the most promising of these.

1.1 Zooarchaeology by Mass Spectrometry (ZooMS)

Bone and tooth dentine is dominated by type I collagen, a three-stranded molecule, with a triple helical structure (Brodsky and Ramshaw 1997). Tetrapods typically have two identical α1(I) strands and one

α2(I) strand (Henriksen and Karsdal 2019). The α1 strands contain a repeated glycine-proline-hydroxyproline motif (Chu et al. 1984), this is highly conserved throughout many species so it is not suitable for use in species identification. However, the α2 strand is genetically variable, with more distant relatives exhibiting greater variation from one another (Buckley 2018). This strand has a highly conserved charge distribution; as trypsin typically cleaves at the C-terminal of arginine and lysine (Olsen et al. 2004), the α2 chain exhibits conserved cleavage patterns during tryptic digestion.

This has all allowed for the development of a technique called ZooMS (Buckley et al. 2009),a collagen peptide mass fingerprinting technique which can extract taxonomic data from fragmentary bone most often down to genus level (Buckley et al. 2017, 2019) and to species level in camels, some rodents and bats (Rybczynski et al. 2013; Buckley and Herman 2019), turtles (Harvey et al. 2019), frogs (Buckley and Cheylan 2020) and fish (Rick et al. 2019; Hawkins et al. 2022), including the ability to identify hybrids (Marković et al. 2021). ZooMS was originally developed using the insoluble collagen fragment (Buckley et al. 2009), while newer methods utilise the soluble fragment (van der Sluis et al. 2014). The insoluble route poses less contamination risk while the soluble route can produce greater collagen yields, but both routes produce a similar number of peptide mass fingerprints (Buckley 2018). Generally, contamination is less of an issue in ZooMS than aDNA sequencing as there is no amplification involved and low-level contaminants are relatively easy to detect and thus disregard.

The soluble ZooMS methodology is currently performed via two operating procedures both low-throughput (LT) (Buckley et al. 2009) and high-throughput (HT) (Buckley et al. 2016). The LT method uses singular Eppendorf tubes, whereas the HT method uses 96-well microtiter plates. The LT method is better suited to processing precious specimens or low sample sizes due to risks relating to clogging and contamination, as well as yielding better quality spectra. HT is much more time efficient and can process up to 384 samples at once (sample estimate based on being able to centrifuge four 96 well plates simultaneously, increased, or decreased centrifuge access would alter this estimation). However, since Buckley *et al.* introduced HT ZooMS in 2016 only two other papers have analysed the spectra of ≥ 1000 samples (Brown et al. 2021; Harvey et al. 2022). This is because, funding, time, and staff availability all heavily limit the uptake of ZooMS. Despite, being more time efficient than LT, HT is still a manual process and requires substantial staff time, which is usually the costliest expense in a laboratory.

Processing large sample sizes has further benefit in zooarchaeology and palaeoecology than just increasing the utilisation of excavated archaeofaunal material. ZooMS analyses are a looking glass through which we can view ancient ecosystems and by expanding the size of our analysis we expand our field of view into the past and thus the scope of our investigations. Larger sample sizes can approximate the actual population with much greater precision and power than smaller sample sizes. In larger samples sizes, the potential effect of sampling bias is reduced, and outliers have less influence. This is important as often ZooMS analyses are used to make ecological inferences about ancient ecosystems. This can include proposing phylogenies, inferring anthropogenic effects on ancient biodiversity, and informing on hominin subsistence behaviours and the development of animal husbandry (Harvey et al. 2019, 2021; Janzen et al. 2021). By only sampling a small proportion of excavated archaeofaunal

material, outliers and uncontrollable sampling biases can heavily influence research outputs. Increasing sample sizes in ZooMS analyses will increase likelihood that research inferences are representative of the population being studied.

1.2 Grotte Mandrin, France

Grotte Mandrin is a rock shelter in the Rhône Valley, Drôme, France, overlooking the eastern bank of the middle Rhône River Valley. Mandrin's 3-m deep stratigraphic Pleistocene sequence has been broken down into 12 archaeological layers (J-B1), with radiocarbon dating placing the oldest layer at marine isotope stage 5, with the younger layers being dated to the end of the middle palaeolithic and the advent of the Upper Palaeolithic (Slimak et al. 2022; Metz et al. 2023).

This site has long been of interest in the study of European hominin evolution. Hominin fossils have been excavated across multiple layers of the site, including 9 dental specimens and ~ 60,000 lithics (Slimak et al. 2022). These hominin artifacts are representative of both early anatomically modern human and Neanderthal occupations. Compared to other regions of the world, evidence of early modern human emergence in Europe out of Africa is relatively late, potentially due to well established Neanderthal occupation (Hublin 2015; Liu et al. 2015; Clarkson et al. 2017). At present Mandrin is home to the earliest documented anatomically modern humans in Europe, with the remains being dated to 56,800–51,700 calibrated years before present (cal. B.P.) using accelerator mass spectrometry (AMS) (Slimak et al. 2022). AMS radiocarbon dating, lithic analysis and dental morphology has helped reveal that throughout the last thousand years of the Neanderthal existence, multiple successive replacement hominin occupations occurred at Mandrin of both Neanderthal and anatomically modern human groups (Slimak et al. 2022). Therefore, Mandrin provides a valuable insight into a period in which Neanderthal and early modern human coexisted in Europe.

Grotte Mandrin faces north towards the Mistral wind (Vandevelde et al. 2017). The Mistral is a cold and dry mesoscale wind that is found in the valleys of south-eastern France, it blows from the north through the Rhône Valley before eventually reaching the northwest coast of the Gulf of Lion, Mediterranean Sea (Obermann et al. 2018). Mandrin sits at 225 m elevation, with little protection from the Mistral, many sediments have settled here, riding the tails of the strong wind into the shelter (Slimak et al. 2022). This sustained history of sediment deposit at the site has aided in the preservation of the large collection of archaeological material. Research at Mandrin began in 1990 and the shelter has since undergone multiple excavations. At present ~ 70,000 archaeofaunal remains have been excavated, a large proportion of which are fragmentary and unable to be identified by morphological analysis (this proportion varies with taxonomic level and the type of site, but typically 70–90% of recovered archaeological remains are not identified beyond family level). The extensive archaeofaunal assemblage at Grotte Mandrin is a great exemplary candidate for the need of an automated ZooMS technique. A sizeable investment of time and funding has allowed for Grotte Mandrin to be excavated for the last three decades. In this study we describe an automated ZooMS (AutoZooMS) approach that we use to help ensure maximal taxonomic data yields can be secured from this investment.

All samples in this study are sourced from Grotte Mandrin, layer B2, which dates to ~ 42,000–44,000 cal. B.P. and attributed to the Post-Neronian II culture representing here the last expression of the Mousterian in Mediterranean France (Higham et al. 2014). Neanderthal remains have been excavated from this layer, along with Mousterian technologies that are heavily associated with European Neanderthal occupation (Klein 1992). Therefore, the fauna identified in this project have the potential to provide insight into Neanderthal subsistence behaviours and due to factors previously discussed, the large sample size in the study helps provide more accurate inferences than previously achieved. To be a viable method of identifying whole archaeofaunal assemblages and to increase uptake of HT ZooMS in the scientific community, the efficiency and robustness of HT ZooMS needs to improve. The aims of this study were to explore the inclusion of liquid handling robots into the ZooMS protocol, both for improving accuracy, repeatability, and robustness (e.g., to multichannel pipette clogging error), hereafter referred to as 'AutoZooMS', which is the HT method adapted for automation. We hope that the fast-processing power of this technique will promote increased utilisation of excavated archaeofaunal material to increase taxonomic data yields, reduce archaeofaunal waste and encourage larger sample sizes in future zooarchaeological research. AutoZooMS has been developed using the archaeofaunal assemblage at layer B2, Grotte Mandrin, thus the secondary aim of this study has been to conduct a zooarchaeological analysis of this site.

2. Methods And Materials

2.1 Zooarchaeological Samples

A total of 1707 samples were involved in the AutoZooMS zooarchaeological analysis of Grotte Mandrin, all from layer B2 from multiple excavations between 2018–2020. For comparison of success rate and species richness, 96 samples were run via LT and 469 via HT.

2.2 Zooarchaeology by Mass Spectrometry

LT, HT, and AutoZooMS used identical reagents: Hydrochloric acid (HCL) and sequencing-grade trypsin were sourced from Fisher Chemicals (U.K.) and Promega (U.K.) respectively. Acetonitrile (ACN), ammonium bicarbonate (ABC), trifluoroacetic acid (TFA), α-cyano-4-hydroxycinnamic acid (matrix) and mass spectrometric standards were obtained from Sigma-Aldrich (U.K.).

2.2.1 Low Throughput Protein Extraction

Following an adapted protocol laid out by van der Sluis et al. (2014) bone fragments were decalcified overnight in 1.5 mL Eppendorf tubes using 1 mL 0.6 M HCL. The next day, 0.5 mL of acid was transferred into 10 kDa molecular weight cut off (MWCO) ultrafilters (Vivaspin, U.K.) and centrifuged for 30 minutes at 12,400 rpm. After discarding flow through 0.5 mL 50 mM ABC was added and the ultrafilters were centrifuged. This process was conducted twice. The acid-soluble collagen peptides were resuspended in 0.2 mL 50 mM ABC and transferred to 1.5 mL Eppendorf tubes. Then 0.2 µg of trypsin was added to each sample. The soluble collagen fragments were then left to digest overnight on a heating block at 37°C. 1

µL sample was then diluted in 19 µL 0.1% TFA. For each sample 1 µL of matrix solution (10 mg/mL in 50% ACN + 0.1% TFA) and 1 µL of diluted sample was spotted onto a 384-well stainless steel MALDI target plate and left to air dry.

2.2.2 High Throughput Protein Extraction

HT protein extraction was based on the protocol laid out by Buckley et al. (2016). HT uses the same reagent volumes and concentrations as LT, the procedure is adapted to a plate processing format using multi-channel pipettes. Bones were decalcified in 48-well Corning™ Costar™ cell culture-treated flat-bottom plates. Soluble protein was filtered in 1 mL well 10 kDa MWCO 96-well ultrafiltration plates (Pall Corporation). These ultrafiltration plates were sat on 1.1 mL 96 deep well plates (Axygen) that collected any liquid flow through during centrifugation. A microplate sealing film (Sigma-Aldrich, U.K.) was then placed on top of the ultrafiltration plate and the structure was taped together. The plates were then centrifuged for 45 minutes at 23 °C, 3700 mph. Then two ABC washes were conducted. After the last centrifuge, soluble protein was resuspended in 200 µL ABC and 100 µL of sample was transferred into 2 x 96-well Corning™ Costar™ cell culture-treated flat-bottom plates. Overnight trypsin digest took place on a heating block at 37°C (0.2 ug trypsin per sample). 40 µL of matrix solution (10 mg/ml in 50% ACN + 0.1% TFA) was added to a 1.1 mL 96 deep well plate (Axygen), then 2 µL of sample was mixed into the matrix solution. Finally, 2 µl of this mixture was spotted onto a 384 stainless steel MALDI target plate and left to air dry.

2.2.3 Automated Protein Extraction

AutoZooMS was developed on Hamilton Method Editor for the Hamilton Microlab STAR Liquid Handling System using the multiprogramming software Venus. The method uses 50 µL, 300 µL and 1000 µL CO-RE Hamilton Tips and is split into 5 runnable methods (Fig. 1): 1 - HCl addition, 2 - HCl removal, 3 - filter plate wash and elute, 4 - trypsin addition and 5 - matrix addition and spotting. Reagent volumes differ slightly to HT. The machine used 8 individual pipetting channels to add 0.95 mL 0.6 M HCl to each 48-well plate and 1.9 mL to each 24-well plate. Then of this acid-soluble protein, 0.3 mL was removed and ultrafiltered; 10 kDa molecular weight cut-off (MWCO) ultrafiltration plates were used during method optimisation and 30 kDa MWCO plates thereafter. A 96-channel Multi-Probe-Head (MPH) was used for ABC addition, sample resuspension and transfer to 96-well plates. Manual intervention was required at each centrifugal stage. Trypsin (0.2 µg trypsin per sample in initial testing, 0.1 µg in final method) was added using 8 individual channels. Manual intervention was required at this point to transfer samples to heating block. Matrix addition and mixing with sample was conducted using the 96 MPH head. All spotting was carried out using the 8 individual channels. Troughs of HCL, ABC and matrix provided the machine with the necessary reagent reservoirs. Trypsin was sourced from first column (A-H) of a PCR plate as the cone shaped bottom favors retrieval of small volumes by the machine.

Each method within the programme could be run independently to allow multiple batches of samples to be processed simultaneously to maximize the speed and number of samples being processed, up to 3 x

384 batches could be processed simultaneously. Using parameters laid out in Supplementary Material S1 & S2, it would be feasible to process > 1000 samples within a 24-hour period.

2.2.4 Mass Spectrometry

Collagen fingerprints were generated using a Bruker RapifleX matrix assisted laser desorption/ionisation (MALDI) time of flight (TOF) mass spectrometer. Each spot was calibrated against a 1 in 10 diluted MS standard which contained five known peptides. This standard was made using a ProteoMass™ Protein MALDI-MS Calibration Kit (Sigma-Aldrich, UK) following supplier's protocol. MALDI-TOF mass spectra were acquired by collecting up to 2000 laser shots per sample spot and peaks within the *m/z* range 1100–3200 were compared with those published previously (Buckley and Collins 2011; Welker et al. 2016; Buckley et al. 2017) for anticipated Middle Palaeolithic megafauna common to Europe. In initial method testing, collagen fingerprinting success was scored 0, 1 or 2 (Pothier Bouchard et al. 2019). Spectra scoring 0 failed to produce any collagen biomarkers, spectra scoring 1 exhibit one or more identifiable collagen biomarker, species ID was able to be successfully assessed from spectra scoring 2.

3. Results

3.1 Low Throughout vs High Throughput vs AutoZooMS

Once the AutoZooMS programme had been successfully developed, an initial test of 95 samples was analysed to compare the success and spectra quality of LT, HT, and AutoZooMS. This test was conducted to ensure the programme was not only functional but was as successful as HT ZooMS which it was adapted from. Each of the 95 samples were broken into three pieces and each piece was ran by LT, HT or AutoZooMS respectively. For fair comparison all three runs maintained similar parameters in terms of trypsin concentration and decalcification and digest length. The results of this comparison can be seen in table 1. AutoZooMS exhibited a collagen success rate of 62.1%, which was ~1.2x higher than its HT counterpart, and a species ID success rate of 50.5% which was ~1.5x higher than HT. LT did exhibit higher success rates than both HT and AutoZooMS which is to be expected. Thus, AutoZooMS provides a new alternate high throughput methodology that is more successful at species identification than current HT techniques.

Table 1: Comparison of LT, HT and AutoZooMS Collagen Fingerprinting and Species Identification Success.

	Low Throughput	High Throughput	Automated
Species Identification Success (n)	57	31	48
Species Uncertain Collagen Success (n)	14	18	11
Total Success (n)	71	49	59
Fail (n)	24	46	36
Collagen Success (%)	74.7	51.6	62.1
Identification success (%)	60.0	32.6	50.5

3.2 AutoZooMS Analysis of Grotte Mandrin, France

After the completion of all method optimisation 1612 samples were processed using AutoZooMS and then manual species identification of collagen yielding samples took place (e.g., Fig. 2A-E). AutoZooMS had a 60.7% species identification success rate (Fig. 2F), slightly higher than the 60% LT success rate seen in the AutoZooMS initial testing (Table 1), indicating that alterations to sample to matrix spotting ratio were successful in improving AutoZooMS identification success. In comparing the composition of the archaeofaunal assemblage at layer B2 Grotte Mandrin (Fig. 2G), this analysis revealed further evidence of hominin occupation, with ~5% of the assemblage yielding a hominin signature (likely Neanderthal). The layer was largely dominated by deer (36.4%), the majority of which will be red deer, followed by horse (19.4%) and pig (13.2%); a total of 11 taxonomic groups were identified, including, mammoth, bovine (*Bos/Bison*), reindeer, caprine (*Ovis/Capra*), lagomorph (*Oryctolaglus/Lepus*), pig, woolly rhino, horse, hominin, cervid (deer excluding reindeer) and bear (Fig. 2G).

3.3 Effect of Sample Size on Species Abundance and Richness

To assess the impact ZooMS sample size has on our understanding of the composition of archaeofaunal assemblages, the species abundance and richness of 96 LT samples, 469 HT samples and 1707 AutoZooMS samples were compared (Figure 3; Table 2). Results showed that by facilitating the processing of larger sample sizes AutoZooMS provides higher resolving power to identify less abundant species such as mammoth; LT exhibited the lowest species richness and AutoZooMS the highest. Assemblage composition does differ between the three data sets, notable differences include the percentage of pig (*Sus*), which in HT analyses comprised only 4.7% of the identifiable assemblage whereas in AutoZooMS and LT it was 13.5% and 17.2% respectively. Equine was also notable, composing 33.0% in HT analyses, despite only making up 20.7% and 19.0% of LT and AutoZooMS analysis. Similarly, lagomorph composed 4.7% in HT but 1.7% and 1.3% in LT and AutoZooMS. Woolly rhinoceros was not present at all in LT analysis but were present 0.5% and 1.1% in HT and AutoZooMS respectively. Overall, we observed notable differences caused by sampling bias when subsampling from the Grotte Mandrin archaeofaunal assemblage, particularly the lack of resolving power to identify lower abundance species such as mammoth and woolly rhinoceros.

Table 2: Species Identification Success and Species Richness for LT, HT and AutoZooMS

	Low Throughput	High Throughput	Automated
Total Sampled	96	469	1707
Species ID Success (%)	60	40.7	60.7
Uncertain (%)	14.6	10.0	0.644
Species Count	9	10	11

4. Discussion

4.1 Automated ZooMS

The overriding focus of this study was to develop an automated method of extracting taxonomic data from archaeofaunal assemblages with the aims to increase zooarchaeological data yields and reduce current archaeofaunal waste. Our AutoZooMS approach was shown to be more successful than traditional HT collagen fingerprinting for species identification, yielding the potential to increase acquisition of large-scale ZooMS datasets (e.g., Buckley et al. 2016; Brown et al. 2021). Across a five-day working week, AutoZooMS could process ~ 2000 samples with under 1-hour human intervention each day; all 1612 samples used in this analysis were spotted in < 3 hours with 30 minutes of human intervention. If transferred to a fully automated system (that includes centrifugation and incubation) using parameters laid out in the supplementary material (Supplementary Figs S1 & S2), > 1000 samples could be fully processed in 24 hours. One other factor to consider would relate to costing and laboratory consumable efficiency; as protein preservation varies between sites, when working with a sample set from a new zooarchaeological site it would be beneficial to evaluate protein abundance and establish optimum protease to protein ratios before any ZooMS analysis takes place, in this case 20 µg trypsin/0.4 mL buffer.

Recently, Rüther et al. (2022) presented a shortened LC-MS/MS process, proposing that up to 200 samples could be analysed per day. By comparison, ZooMS can process ~ 2x this, but more taxonomically limited relative to LC-based methods. LC-MS/MS approaches have been used for decades to identify protein (and species) composition within a sample, to greater accuracy than ZooMS, but the taxonomic resolution is more closely linked to amino acid sequences present in the search databases. For example, Mackie et al. (2018) were able to distinguish between chicken and duck egg white from a 14th century Italian wall painting whereas due to the highly conserved nature of bird collagen and the coincidental matching of peptides with different amino acid sequences, such taxonomic separation is difficult to confidently achieve with ZooMS peptide mass fingerprinting (Buckley 2018). However, due to funding constraints in zooarchaeology, AutoZooMS is more suited to the mass processing of whole archaeofaunal assemblages than LC-MS/MS approaches.

4.2 Archaeofaunal Assemblage of Grotte Mandrin, France

The composition of layer B2 deduced in this study is similar to that of previous manual ZooMS analysis whereby Pal Chowdhury et al. (2019) reported that red deer (29%), aurochs/bison (20%) and horse (21%) made up most of the layer B2 assemblage (n = 70) with a small proportion of pig (1%), but no mammoth, hominin, bear, or rabbit recorded in their study. This therefore shows that by processing much larger sample sizes, AutoZooMS can provide a more accurate picture of archaeofaunal assemblages and ancient environments. The ZooMS data from this project may be useful in the study of Neanderthal subsidence behaviours, further supporting previous evidence of hominin occupations in Grotte Mandrin (Vandevelde et al. 2017).

A limitation of ZooMS is that it produces highly inflated number of identified specimen counts (NISP) compared to traditional morphological analysis (Buckley et al. 2017; Sinet-Mathiot et al. 2019). NISP counts are a measure of species abundance and as specimens processed via ZooMS are often highly fragmentary, it lends itself to an overinflated assumption that each fragment represents a unique individual, creating bias in our understanding of these sites. This would otherwise have been mitigated by the more ideal zooarchaeological measure of minimum number of individuals (MNI) relying on using unique skeletal (including sided) elements. As Grotte Mandrin has recorded hominin presence, fragmentary bone could be representative of hominin butchery for subsistence or tool making, i.e., bones being purposefully broken into smaller pieces. Hominin butchery could also be contributing to the failure rate of the ZooMS analysis presented in this study, as heat applied to bone during cooking would be expected to increase collagen degradation. Moreover, this anthropic fragmentation could be distinct between the different species exploited as related to hominin choices and traditions. For example, bone marrow extraction, that induce an anthropic fragmentation of the bone could be limited only to particular species. Finally, the relatively large representation of hominin bones found in this study could have been biased by the chosen sample locations that were focused in the direct vicinity of a Neanderthal body that has been slowly excavated in the area since 2015. As the archaeofaunal sample is specific to a particular spatial area of the site, we cannot infer that 5% of the fragmentary bones from throughout this layer (B2) belong to Neanderthal remains. The deployment of this high-throughput AutoZooMS method in the same layer but across a wider region of the site, once excavated, might then be highly informative on improved understanding of the cultural anthropology of these societies.

4.3 Conclusions

To conclude, presented here is a new AutoZooMS methodology, that is just as, if not more successful than prior HT ZooMS methods. This technique can process thousands of samples in a couple of days at relatively lost cost, thus making it a great technique to scan whole archaeofaunal assemblages. With advances in method development such as proposed here, more ZooMS analyses could be subsequently carried out using much larger sample sizes that are more truly representative of the contemporary ecology, or at least reducing one of the main taphonomic biases to such interpretation of ancient life.

Statements And Declarations

Competing Interests: The authors have no competing interests.

Data availability: mzXML files to be uploaded to Figshare.

Author contribution: E.M.O., M.P.C. and M.D. carried out laboratory analyses. L.S. provided samples. E.M.O. and M.B. wrote the main manuscript text. E.M.O. and M.D. prepared figures 1-5. All authors reviewed the manuscript.

Funding statement: The acknowledge support from the Henry Royce Institute for Advanced Materials (funded through EPSRC grants nos. EP/R00661X/1, EP/S019367/1, EP/P025021/1 and EP/P025498/1) for access to their facilities.

References

1. Brandt LØ, Schmidt AL, Mannering U, et al (2014) Species Identification of Archaeological Skin Objects from Danish Bogs: Comparison between Mass Spectrometry-Based Peptide Sequencing and Microscopy-Based Methods. PLOS ONE 9:e106875. https://doi.org/10.1371/journal.pone.0106875
2. Briggs D (2003) The Role of Decay and Mineralization in the Preservation of Soft-Bodied Fossils. Annu Rev Earth Planet Sci 31:275–301. https://doi.org/10.1146/annurev.earth.31.100901.144746
3. Brodsky B, Ramshaw JA (1997) The collagen triple-helix structure. Matrix Biol J Int Soc Matrix Biol 15:545–554. https://doi.org/10.1016/s0945-053x(97)90030-5
4. Brown S, Wang N, Oertle A, et al (2021) Zooarchaeology through the lens of collagen fingerprinting at Denisova Cave. Sci Rep 11:15457. https://doi.org/10.1038/s41598-021-94731-2
5. Buckley M (2018) Zooarchaeology by Mass Spectrometry (ZooMS) Collagen Fingerprinting for the Species Identification of Archaeological Bone Fragments. In: Giovas CM, LeFebvre MJ (eds) Zooarchaeology in Practice: Case Studies in Methodology and Interpretation in Archaeofaunal Analysis. Springer International Publishing, Cham, pp 227–247
6. Buckley M, Cheylan M (2020) Collagen fingerprinting for the species identification of archaeological amphibian remains. Boreas 49:709–717. https://doi.org/10.1111/bor.12443
7. Buckley M, Collins M, Thomas-Oates J, Wilson JC (2009) Species identification by analysis of bone collagen using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. Rapid Commun Mass Spectrom 23:3843–3854. https://doi.org/10.1002/rcm.4316
8. Buckley M, Collins MJ (2011) Collagen survival and its use for species identification in Holocene-lower Pleistocene bone fragments from British archaeological and paleontological sites. Antiqua 1:e1–e1. https://doi.org/10.4081/antiqua.2011.e1
9. Buckley M, Gu M, Shameer S, et al (2016) High-throughput collagen fingerprinting of intact microfaunal remains; a low-cost method for distinguishing between murine rodent bones. Rapid Commun Mass Spectrom 30:805–812. https://doi.org/10.1002/rcm.7483
10. Buckley M, Harvey VL, Chamberlain AT (2017) Species identification and decay assessment of Late Pleistocene fragmentary vertebrate remains from Pin Hole Cave (Creswell Crags, UK) using collagen fingerprinting. Boreas 46:402–411. https://doi.org/10.1111/bor.12225

11. Buckley M, Herman J (2019) Species identification of Late Pleistocene bat bones using collagen fingerprinting. Int J Osteoarchaeol 29:1051–1059. https://doi.org/10.1002/oa.2818
12. Buckley M, Recabarren OP, Lawless C, et al (2019) A molecular phylogeny of the extinct South American gomphothere through collagen sequence analysis. Quat Sci Rev 224:105882. https://doi.org/10.1016/j.quascirev.2019.105882
13. Chu ML, de Wet W, Bernard M, et al (1984) Human pro alpha 1(I) collagen gene structure reveals evolutionary conservation of a pattern of introns and exons. Nature 310:337–340. https://doi.org/10.1038/310337a0
14. Clarkson C, Jacobs Z, Marwick B, et al (2017) Human occupation of northern Australia by 65,000 years ago. Nature 547:306–310. https://doi.org/10.1038/nature22968
15. Denys C (2002) Taphonomy and experimentation. Archaeometry 44:469–484. https://doi.org/10.1111/1475-4754.00079
16. Harvey VL, Egerton VM, Chamberlain AT, et al (2019) Interpreting the historical terrestrial vertebrate biodiversity of Cayman Brac (Greater Antilles, Caribbean) through collagen fingerprinting. The Holocene 29:531–542. https://doi.org/10.1177/0959683618824793
17. Harvey VL, Keating JN, Buckley M (2021) Phylogenetic analyses of ray-finned fishes (Actinopterygii) using collagen type I protein sequences. R Soc Open Sci 8:201955. https://doi.org/10.1098/rsos.201955
18. Harvey VL, LeFebvre MJ, Sharpe AE, et al (2022) Collagen fingerprinting of Caribbean archaeological fish bones: Methodological implications for historical fisheries baselines and anthropogenic change. J Archaeol Sci 145:105642. https://doi.org/10.1016/j.jas.2022.105642
19. Hawkins AL, Buckley M, Needs-Howarth S, Orchard TJ (2022) Practice makes perfect? Inter-analyst variation in the identification of fish remains from archaeological sites. Int J Osteoarchaeol 32:694–705. https://doi.org/10.1002/oa.3096
20. Henriksen K, Karsdal MA (2019) Type I collagen. In: Biochemistry of Collagens, Laminins and Elastin. Elsevier, pp 1–12
21. Higham T, Douka K, Wood R, et al (2014) The timing and spatiotemporal patterning of Neanderthal disappearance. Nature 512:306–309. https://doi.org/10.1038/nature13621
22. Hublin J-J (2015) The modern human colonization of western Eurasia: when and where? Quat Sci Rev 118:194–210. https://doi.org/10.1016/j.quascirev.2014.08.011
23. Janzen A, Richter KK, Mwebi O, et al (2021) Distinguishing African bovids using Zooarchaeology by Mass Spectrometry (ZooMS): New peptide markers and insights into Iron Age economies in Zambia. PLOS ONE 16:e0251061. https://doi.org/10.1371/journal.pone.0251061
24. Kahila Bar-Gal G, Ducos P, Kolska Horwitz L (2003) The application of ancient DNA analysis to identify neolithic caprinae: a case study from the site of Hatoula, Israel. Int J Osteoarchaeol 13:120–131. https://doi.org/10.1002/oa.666
25. Klein RG (1992) The archeology of modern human origins. Evol Anthropol Issues News Rev 1:5–14. https://doi.org/10.1002/evan.1360010105

26. Liu W, Martinón-Torres M, Cai Y, et al (2015) The earliest unequivocally modern humans in southern China. Nature 526:696–699. https://doi.org/10.1038/nature15696
27. Mackie M, Rüther P, Samodova D, et al (2018) Palaeoproteomic Profiling of Conservation Layers on a 14th Century Italian Wall Painting. Angew Chem Int Ed Engl 57:7369–7374. https://doi.org/10.1002/anie.201713020
28. Marković N, Ivanišević V, Baron H, et al (2021) The Last Caravans in Antiquity: Camel Remains from Caričin Grad (Justiniana Prima). J Archaeol Sci Rep
29. Metz L, Lewis JE, Slimak L (2023) Bow-and-arrow, technology of the first modern humans in Europe 54,000 years ago at Mandrin, France. Sci Adv 9:eadd4675. https://doi.org/10.1126/sciadv.add4675
30. Morin E, Ready E, Boileau A, et al (2017) Problems of Identification and Quantification in Archaeozoological Analysis, Part II: Presentation of an Alternative Counting Method. J Archaeol Method Theory 24:938–973. https://doi.org/10.1007/s10816-016-9301-3
31. Newman ME, Parboosingh JS, Bridge PJ, Ceri H (2002) Identification of Archaeological Animal Bone by PCR/DNA Analysis. J Archaeol Sci 29:77–84. https://doi.org/10.1006/jasc.2001.0688
32. Noe-Nygaard N (1977) Butchering and marrow fracturing as a taphonomic factor in archaeological deposits. Paleobiology 3:218–237. https://doi.org/10.1017/S0094837300005285
33. Obermann A, Bastin S, Belamari S, et al (2018) Mistral and Tramontane wind speed and wind direction patterns in regional climate simulations. Clim Dyn 51:1059–1076. https://doi.org/10.1007/s00382-016-3053-3
34. Olsen JV, Ong S-E, Mann M (2004) Trypsin cleaves exclusively C-terminal to arginine and lysine residues. Mol Cell Proteomics MCP 3:608–614. https://doi.org/10.1074/mcp.T400003-MCP200
35. Pal Chowdhury M, Wogelius R, Manning PL, et al (2019) Collagen deamidation in archaeological bone as an assessment for relative decay rates. Archaeometry 61:1382–1398. https://doi.org/10.1111/arcm.12492
36. Pothier Bouchard G, Mentzer SM, Riel-Salvatore J, et al (2019) Portable FTIR for on-site screening of archaeological bone intended for ZooMS collagen fingerprint analysis. J Archaeol Sci Rep 26:101862. https://doi.org/10.1016/j.jasrep.2019.05.027
37. Presslee S, Wilson J, Woolley J, et al (2017) The identification of archaeological eggshell using peptide markers. STAR Sci Technol Archaeol Res 3:89–99. https://doi.org/10.1080/20548923.2018.1424300
38. Rick T, Harvey VL, Buckley M (2019) Collagen fingerprinting and the Chumash billfish fishery, Santa Barbara Channel, California, USA. Archaeol Anthropol Sci 11:6639–6648. https://doi.org/10.1007/s12520-019-00930-4
39. Rüther PL, Husic IM, Bangsgaard P, et al (2022) SPIN enables high throughput species identification of archaeological bone by proteomics. Nat Commun 13:2458. https://doi.org/10.1038/s41467-022-30097-x
40. Rybczynski N, Gosse JC, Richard Harington C, et al (2013) Mid-Pliocene warm-period deposits in the High Arctic yield insight into camel evolution. Nat Commun 4:1550.

https://doi.org/10.1038/ncomms2516

41. Sinet-Mathiot V, Smith GM, Romandini M, et al (2019) Combining ZooMS and zooarchaeology to study Late Pleistocene hominin behaviour at Fumane (Italy). Sci Rep 9:12350. https://doi.org/10.1038/s41598-019-48706-z
42. Slimak L, Zanolli C, Higham T, et al (2022) Modern human incursion into Neanderthal territories 54,000 years ago at Mandrin, France. Sci Adv 8:eabj9496. https://doi.org/10.1126/sciadv.abj9496
43. van der Sluis L, Hollund H, Buckley M, et al (2014) Combining histology, stable isotope analysis and ZooMS collagen fingerprinting to investigate the taphonomic history and dietary behaviour of extinct giant tortoises from the Mare aux Songes deposit on Mauritius. Palaeogeogr Palaeoclimatol Palaeoecol 416:. https://doi.org/10.1016/j.palaeo.2014.06.003
44. Vandevelde S, Brochier JÉ, Petit C, Slimak L (2017) Establishment of occupation chronicles in Grotte Mandrin using sooted concretions: Rethinking the Middle to Upper Paleolithic transition. J Hum Evol 112:70–78. https://doi.org/10.1016/j.jhevol.2017.07.016
45. Welker F, Hajdinjak M, Talamo S, et al (2016) Palaeoproteomic evidence identifies archaic hominins associated with the Châtelperronian at the Grotte du Renne. Proc Natl Acad Sci U S A 113:11162–11167. https://doi.org/10.1073/pnas.1605834113

Figures

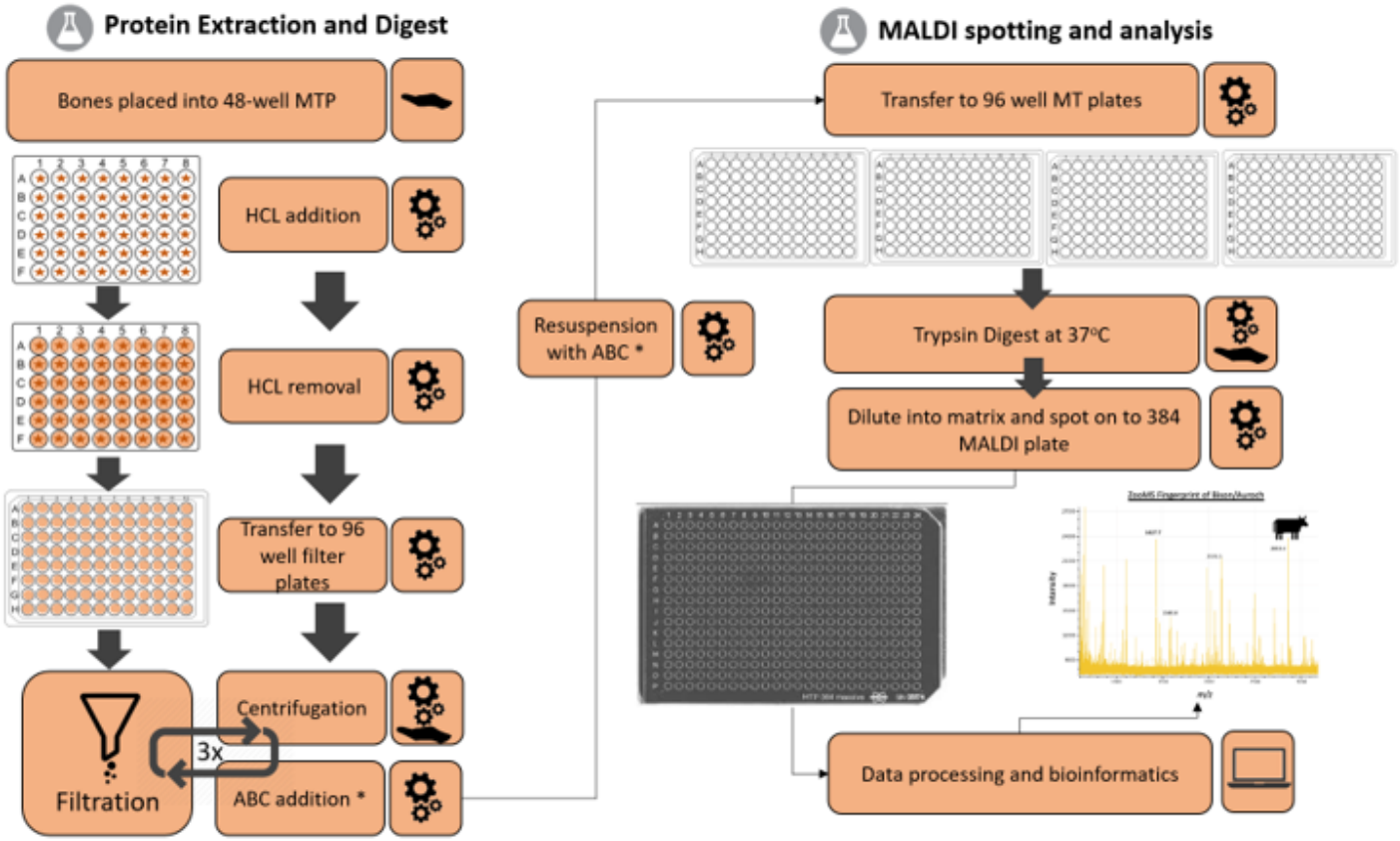

Figure 1

AutoZooMS Workflow.

Hand symbol represents manual intervention, e.g., removing plates from machine and placing in centrifuge or incubator. *Third and final ABC addition is used for resuspension in ABC.

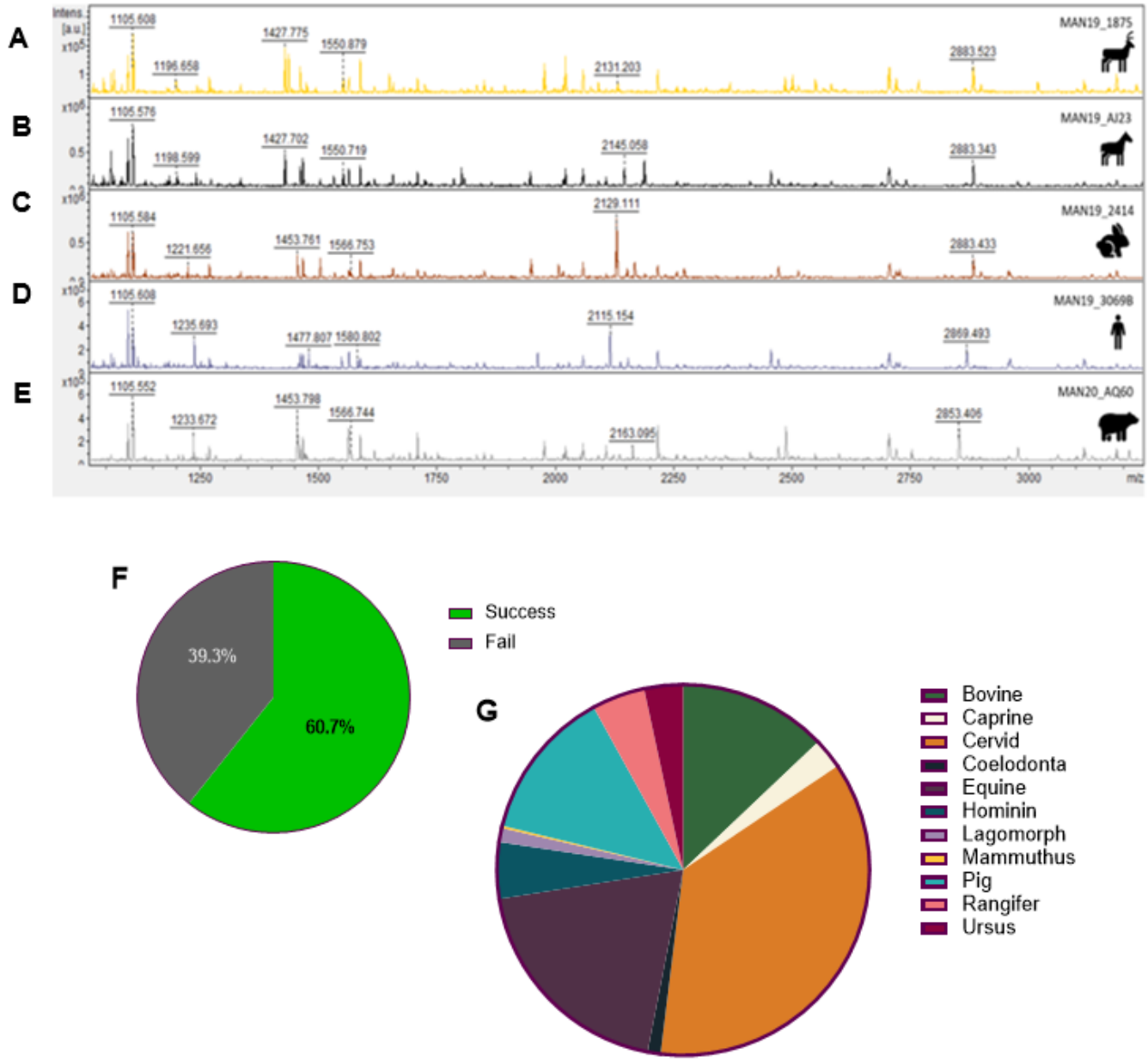

Figure 2

AutoZooMS analysis of Grotte Mandrin, France.

A-E – AutoZooMS collagen fingerprints of deer (A), horse (B), lagomorph (C) hominin (D) and bear (E). F – Species identification success rate of analysis (n=1612). G - The composition of the archaeofaunal assemblage at layer B2, Grotte Mandrin (n=978). Composition: Bovine - 12.9%, Caprine – 2.7%, Cervid –

36.4%, Coelodonta – 1.1%, Equine – 19.4%, Hominin – 4.9%, Lagomorph – 1.2%, Mammuthus – 0.2% Pig – 13.2%, Rangifer – 4.6% and Ursus – 3.4%.

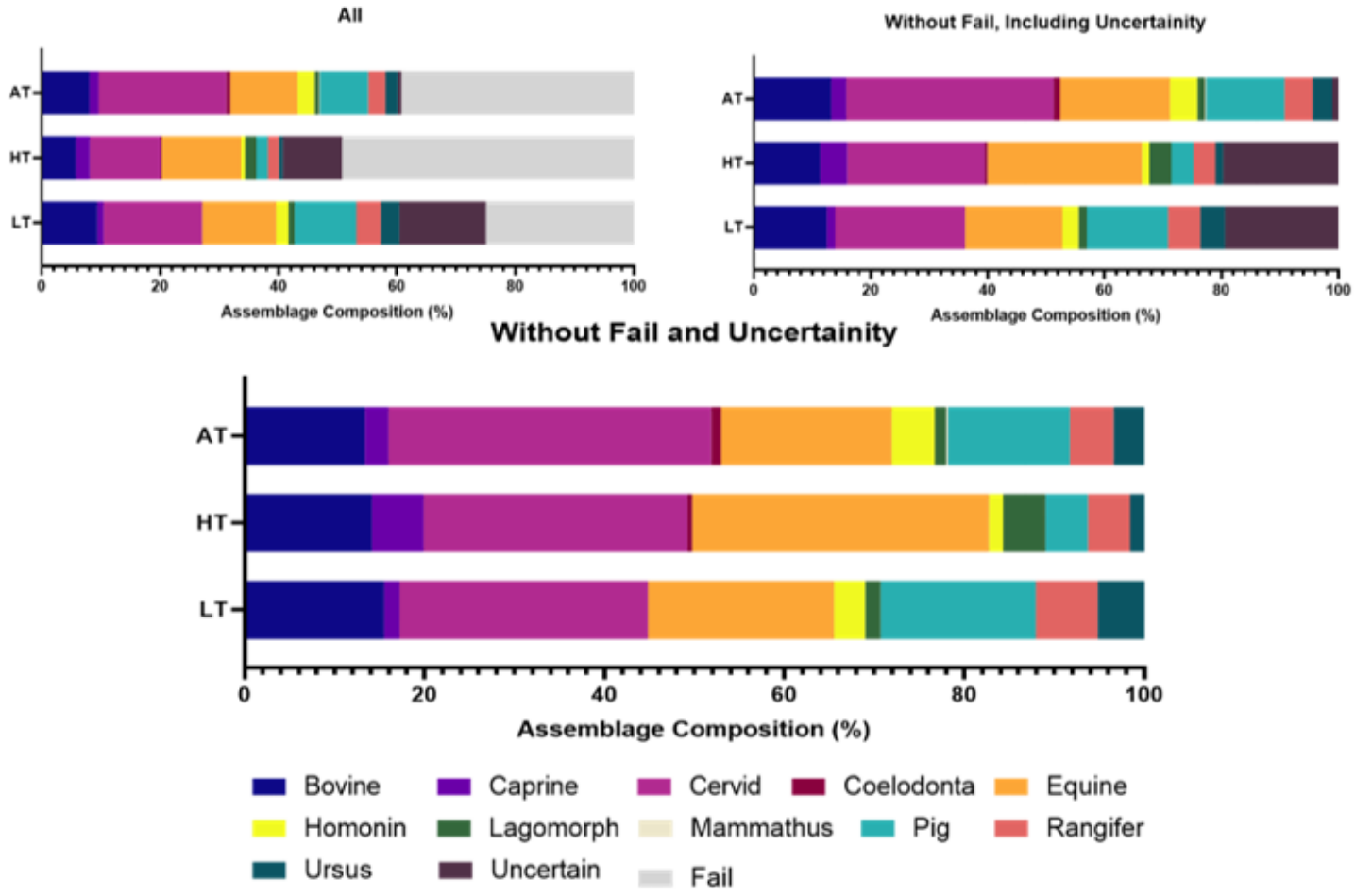

Figure 3

Species Abundance of Layer B2, Grotte Mandrin as measured by LT, HT and AutoZooMS.

Assemblage composition of Grotte Mandrin via LT, HT and AutoZooMS. LT n=96, HT n=469 and AutoZooMS n = 1707. All includes samples that produced successful species identifications, samples that produced some successful collagen fingerprinting but uncertain identifications and fails. Without fail, including uncertainty includes all samples that yielded collagen. Without fail and uncertainty includes samples with successful species identification spectra only.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial.docx