



Medieval fish remains on the Newport ship identified by ZooMS collagen peptide mass fingerprinting

Michael Buckley¹ · Virginia L. Harvey¹ · David Petiffer² · Hannah Russ³ · Wim Wouters⁴ · Wim Van Neer⁴

Received: 22 June 2020 / Accepted: 25 November 2021 / Published online: 11 February 2022
© The Author(s) 2022

Abstract

Fish represent a key economic, social and ecological group of species that humans have exploited for tens of thousands of years. However, as many fish stocks are going into decline and with little known about the anthropogenic impacts on the health of the marine ecosystem pre-Industrial Revolution, understanding historical and archaeological exploitation of fish species is key to accurately modelling these changes. Here, we explore the potential of collagen peptide mass fingerprinting (also known as Zooarchaeology by Mass Spectrometry, or ZooMS) for identifying fish remains from the Medieval (fifteenth century) Newport ship wreck (Wales, UK), and in doing so we establish a set of biomarkers we consider useful in discriminating between European fish taxa through the inclusion of over 50 reference taxa. The archaeological results identified nine distinct taxonomic groups, dominated by ling (> 40%), and a substantial amount of cod (> 20%) and hake (~20%). The vast majority of samples (> 70%) were identified to species level, and the inability to identify the remaining taxonomic groups with confidence using ZooMS was due to the fact that the reference collection, despite being relatively large in comparison to those presented in mammalian studies, reflects only a small proportion of fish biodiversity from this region. Although the results clearly demonstrate the potential for ZooMS as a means of fish bone identification, the sheer number of different fish species that potentially make up ichthyoarchaeological assemblages leads to obvious requirements for the analysis on much greater numbers of modern reference specimens, or the acquisition of collagen sequences.

Keywords Medieval fish trade · ZooMS · Collagen peptide mass fingerprinting · Species identification

Introduction

With the health of modern day marine populations being of constant concern, it is particularly important to have accurate information of the composition and diversity of species from the past (Costello et al. 2010). Although historical records can provide some information to this effect, the data

is often lacking in breadth and scope (Costello et al. 2006). Given that it was during the ‘Middle Ages’ when demand for fishing rapidly increased in line with rising human populations (Hoffmann 1996), improving our understanding of the species being traded in the past is of particular interest for evaluating the overall anthropogenic impacts on these fish populations. It is during this time that we see a shift from the use of limited local inland fish resources to marine resources, particularly from the eleventh century. This phenomenon has been documented in several western European countries through morphological (Harland et al. 2016; Van Neer & Ervynck 2016) as well as stable isotopic analysis (Müldner 2016). However, it is also during this time that substantial environmental changes were occurring, such as the climatic changes that saw temperatures drop ~1 °C from the ‘Medieval Warm Period’ of the tenth to the twelfth centuries to the ‘Little Ice Age’ of the fourteenth to the nineteenth centuries. Ecological consequences of developing fisheries, perhaps combined with fluctuations in temperature, can be recognised through the changes in relative abundance of keystone

✉ Michael Buckley
m.buckley@manchester.ac.uk

¹ School of Natural Sciences, Manchester Institute of Biotechnology, University of Manchester, 131 Princess Street, Manchester M1 7DN, UK

² School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

³ School of Archaeology, History and Anthropology, University of Wales Trinity Saint David, College Street, Lampeter SA48 7ED, UK

⁴ Royal Belgian Institute of Natural Sciences, Vautierstraat 29, B-1000 Brussel, Belgium

species, such as the decline of salmon and sturgeon (e.g. Hoffman 1996), or the movement/increase of species such as eel and carp (e.g. Hoffman 2005). By accurately identifying securely dated fish remains from archaeological sites, we are more able to track population changes due to over-exploitation (Barrett 2019), such as those that occurred with the herring during the twelfth to the fourteenth centuries among other locally exploited species (Hoffman 2005). Ultimately, all around Europe during the later Medieval period, many fisheries that were barely touched or purely for subsistence became subject to commercial exploitation for distant consumers, species that included cod, hake, pike and tuna (e.g. Barrett et al. 2011).

The Medieval period in Europe saw a dramatic change in the procurement, trade and consumption of marine fish, with the origins and expansion of the preserved cod and herring trade (e.g. Barrett and Orton 2016; Barrett et al. 2004, 2008, 2011; Orton et al. 2011; Serjeantson and Woolgar 2006). Religious practices during this period have been linked to an increase in fish consumption related to the avoidance of terrestrial meat consumption on Fridays (e.g. Woolgar 2000), and other religious days, as well as during certain religious periods or festivals, which meant that meat could not be eaten for around half of the year under Christian law. The eleventh century sees the beginning of the preserved fish trade in northern Europe, whereby large Gadiformes fish (often Atlantic cod and ling) were processed and dried to produce ‘stockfish’, while herring were preserved in barrels, both traded on a large-scale basis (Barrett et al. 2004). While the expansion of the preserved fish trade was underway, it is clear that a wide variety of fresh marine and freshwater fish were caught and then consumed at both coastal and inland locations in Britain (Brown et al. 2010; Kowaleski 2003; Serjeantson and Woolgar 2006; Starky et al. 2000), though this aspect of fish consumption is less well understood. If the data presented by Serjeantson and Woolgar (2006, pp. 110–114) are considered, despite the recovery of some large medieval fish bone assemblages ($n \geq 4000$), the identified portion and species diversity are extremely low. Therefore we could potentially be missing crucial information in the understanding of past environment, climate, fish populations and fishing practices due to limitations in the identification of fish remains.

The Newport Ship is the most substantial Late Medieval vessel excavated and recovered in Britain. The ship was discovered on the west bank of the River Usk in Newport, South Wales, in 2002. More than 23 m of the clinker-built ship was recovered, along with significant artefact and environmental assemblages, including fish remains. Finds point to strong Iberian connections during the active life of the ship, which arrived in Newport, in the Severn Estuary, after the spring of AD 1468 (Nayling and Jones 2014). The fish remains from the excavations of the Newport Ship (Russ

2012) provide further evidence for the nature of fish consumption and trade during this period, but only 161 of the 659 bones (<25%) could be identified, albeit at differing taxonomic levels (Russ 2012); therefore, the use of biomolecular techniques to explore species-level information is of particular interest in ichthyoarchaeology.

Species identification of archaeological remains

Notwithstanding preservation biases, archaeological remains can offer a direct source of information on marine species compositions throughout human history and can thus be highly informative of the marine ecosystem prior to the advent of the fishing industry and the keeping of records. In order for such archaeological deposits to be of any benefit, accurate taxonomic identification is required. This is traditionally carried out using morphological comparison of archaeological specimens with modern reference skeletons and remains the first line of consideration for zooarchaeologists (Driver 1992). While some skeletal elements are species diagnostic, many can only be reliably identified at higher taxonomic levels (Family, Order, Class level). This becomes particularly problematic when the faunal remains are fragmentary and/or highly degraded, impacting upon morphological criteria available to the analyst; the more abundant skeletal elements, such as fin rays, ribs, branchiostegals and pterygiophores, are also those less readily identified by morphology (Radu 2005: p. 12). In fish, these combined issues are highly relevant, as bones are often small and have low levels of mineralisation, making accurate identification difficult when relying on morphology alone (Olson and Walther 2007).

Recent decades have seen the development of biomolecular techniques for species identification of archaeological remains, the most common of which is DNA analysis (e.g. Ludwig et al. 2009), often considered a gold standard in the role of identifying ancient specimens. However, despite the fact that next-generation sequencing technology is advancing rapidly and has clear value in retrieving population-level information (for review see Oosting et al. 2019), DNA sequencing remains a time-consuming process that is susceptible to cross-contamination, where mishandling during excavation and retrieval can contaminate samples relatively easily (Llamas et al. 2017; Malmstrom et al. 2005). Furthermore, DNA is easily degraded, limiting its usefulness when specimens are particularly old or damaged as seen in fossils and archaeological samples (note the very low success rates for the study on Neolithic caprinae bones by Kahila Bar-Gal et al. (2003)).

Collagen peptide mass fingerprinting, also referred to as ‘Zooarchaeology by Mass Spectrometry’ or ZooMS for short (Buckley et al. 2009), represents a potential molecular alternative to DNA sequencing with regard to the identification of archaeological bone specimens. The quaternary

structure of collagen, the dominant protein in bone, is such that it imparts properties that allow it to survive greater temperatures for longer (Lozano et al. 2002). It is a triple helix formed from two $\alpha 1(I)$ peptide chains and a third $\alpha 2(I)$ chain in most vertebrate types (Brodsky & Ramshaw 1997). However, in some fish, there is the notable replacement of one of these $\alpha 1(I)$ chains with a genetically distinct $\alpha 3(I)$ chain derived from a duplication of the COL1A1 gene (Morvan-Dubois et al. 2003). The presence of this chain also appears to have tissue-specific abundances (Kimura and Ohno 1987); for example, it is present in greater concentrations in scale and bone than it is in muscle or skin (Kimura et al. 1991). Most importantly for this study, the $\alpha 3(I)$ chain appears to show the highest levels of sequence variation (Buckley 2018; Harvey et al. 2021).

The amino acids which comprise these chains consist of glycine-Xaa-Yaa tripeptide repeats, with any other amino acids substituting for Xaa or Yaa. The most common tripeptide conformation of the peptide chain is glycine-proline-hydroxyproline, whereby hydrogen-bonding between the proline and hydroxyproline residues increases thermal and structural stability (Ramshaw et al. 1998; Rich and Crick 1961). This stability allows collagen to persist in bone at least several millions of years old (e.g. Rybczynski et al. 2013), whereas DNA degrades much faster than this, with variable survival rates related to environmental factors (Allentoft et al. 2012). We have shown that peptide mass fingerprinting, which uses MALDI-TOF mass spectrometric analysis of digested collagen to produce fingerprints, can be a useful tool for distinguishing between taxa, and can be specific to various levels of taxonomic grouping. Importantly, whilst the taxonomic resolution of collagen peptide mass fingerprint spectra is typically limited to the genus for most fauna (e.g. Buckley et al. 2017; 2019) with some current exceptions within the cervids (e.g. Buckley and Collins 2011; Buckley et al. 2017), species level discrimination can often be obtained with microfauna (Buckley et al. 2016; Buckley and Herman 2019), and with fish remains (Harvey et al. 2018; Guiry et al. 2020). Here, we further this work by widening our reference material to those beyond a given region with the aims of discriminating various common European fish species found aboard a Medieval fish trade vessel. In particular, we aimed to identify unknown archaeological specimens recovered from a fifteenth century shipwreck found at Newport, made of timbers from the Basque region of northern Spain and thought to have largely sailed between the UK and Portugal (Nayling and Jones 2014). It is thought some of the assemblage could have been potential trade foods, such as the larger gadids, e.g. Atlantic cod and ling, as well as herring. However, there are also other species represented that are perishable and not typically preserved through salting/drying, and therefore considered shorter term food stuffs. These were more likely crew provisions

either taken from point of origin or even caught at sea during the voyage (Russ 2012).

Materials and methods

Modern samples of 67 species were taken from the University of Sheffield and from personal collections deriving of 61 genera from 30 families spanning 13 orders (Supplementary Table S1). The species were selected in such a way as to obtain a good coverage of the main taxa that have been of importance as food fish in past and present fisheries, in both maritime and continental environments.

Of the archaeological samples gathered from the remains of the fifteenth century ship discovered buried by the banks of the River Usk in Newport, South Wales, 77 specimens that were considered unidentifiable from a morphological point of view were selected for ZooMS analysis (498 of 659 bones from this assemblage were not initially identified beyond 'fish' (Pisces), 303 of which were ribs, 104 scales, and 91 not identifiable to element (Supplementary Table S2)). In order to preserve some for future research, a ~20% subsample was deliberately chosen from across as many contexts as possible). Trifluoroacetic acid (TFA), formic acid (FA), ammonium bicarbonate (ABC), dithiothreitol (DTT), iodoacetamide (IAM) and α -cyano-4-hydroxycinnamic acid were acquired from Sigma-Aldrich (UK), acetonitrile (ACN) and hydrochloric acid (HCl) were purchased from BDH (UK), sequencing-grade trypsin from Promega (UK) and C18 ZipTips were bought from Varian (UK).

Collagen was extracted and digested following modified methods of van der Sluis et al. (2014). Modern reference samples were defatted with 100% hexane at room temperature (with an initial short step of ~15 min followed by one for ~2 h) and the bone samples left to dry in a fume hood overnight. A total of 0.5 mL of 0.6 M HCl was then added overnight in order to demineralise the specimens and extract acid-soluble collagen. This solution was then added to 10-kDa molecular weight cutoff (MWCO) ultrafilter spin columns and centrifuged for 15 min at 14,000 rpm, washed twice with 50 mM ABC (each time centrifuged as above) and 100 μ L ABC used to collect the retentate. The collagen from the archaeological specimens was extracted using the same method as the extraction of the acid-soluble fraction written above but without defatting and in 96-well microtiter plates following Buckley et al. (2016).

In order to purify the peptides from the modern reference samples, C18 ZipTip® solid phase extraction (SPE) pipette tips were used following Buckley et al. (2009), eluted with 50% ACN/0.1% TFA, lyophilised and resuspended with 10 μ L 0.1% TFA. One-tenth was spotted onto a stainless steel Bruker MALDI-TOF target plate along with 1 μ L of

matrix solution (10 mg/mL α -cyano-4-hydroxycinnamic acid in 50% ACN/0.1% TFA). Analysis of each fraction was accomplished using a calibrated Bruker Ultraflex II MALDI TOF/TOF mass spectrometer over the m/z range 700–3700 (Buckley et al. 2009). Markers for the identification of the archaeological specimens were selected based on those that were taxon-specific, but peaks over m/z 3000 were ignored as few of the archaeological samples showed any beyond this value, most likely as a consequence of taphonomic processes acting on the larger peptide chains. We also explored overall similarity between the peaks observed in each archaeological group (the spectrum with most peaks per group) and its nearest modern reference match by identifying the percentage of the 100 most intense peaks for which there is an equivalent peak m/z (within 0.5 mass units; marked with '1' in Supplementary Tables S3–S13).

In order to confirm homology between some of the key biomarkers in this study, further sequencing analyses were carried out on each of the archaeological groups identified at least to genus level, whereby a sub-aliquot of the collagen digests was subject to LC–MS/MS (Buckley et al. 2015; using a Waters nanoAcquity UPLC system coupled to a Thermo Scientific Orbitrap Elite MS). After concentration on a 20 mm \times 180 μ m pre-column, digested peptides were separated on a 1.7- μ m Waters nanoAcquity Ethylene Bridged Hybrid C18 analytical column of (75 mm \times 250 μ m

i.d.) using a gradient beginning at 99% buffer A/1% buffer B and finishing at 75% buffer A (0.1% FA in H₂O)/25% buffer B (0.1% FA in ACN). Resulting data files were searched against a custom database of fish collagen (I) sequences (see Supplementary Table S14 for NCBI accession numbers) using Error Tolerant searches with Mascot v2.5 allowing for one missed cleavage, peptide tolerance of ± 0.5 Da, MS/MS fragment ion mass value tolerance of 0.5 Da, variable oxidation of proline (P) and lysine (K) modifications (mass shift = + 15.99 Da; equivalent mass to the process of hydroxylation). Mascot search results were compared to the collagen fingerprints of the matched groups to identify peptides that could be assigned to a specific taxonomic level (e.g. Harvey et al. 2018). We also used the sequence information to aid the identification of peptide biomarker locations based against the *Gadus morhua* sequence and fingerprint (see Fig. 1, Supplementary Table S15 and Supplementary Figs. S1–S14).

Results and discussion

Taxonomic resolution in modern fish bone

Several collagen peptide peaks appeared potentially useful with discriminating at the taxonomic level of order or family

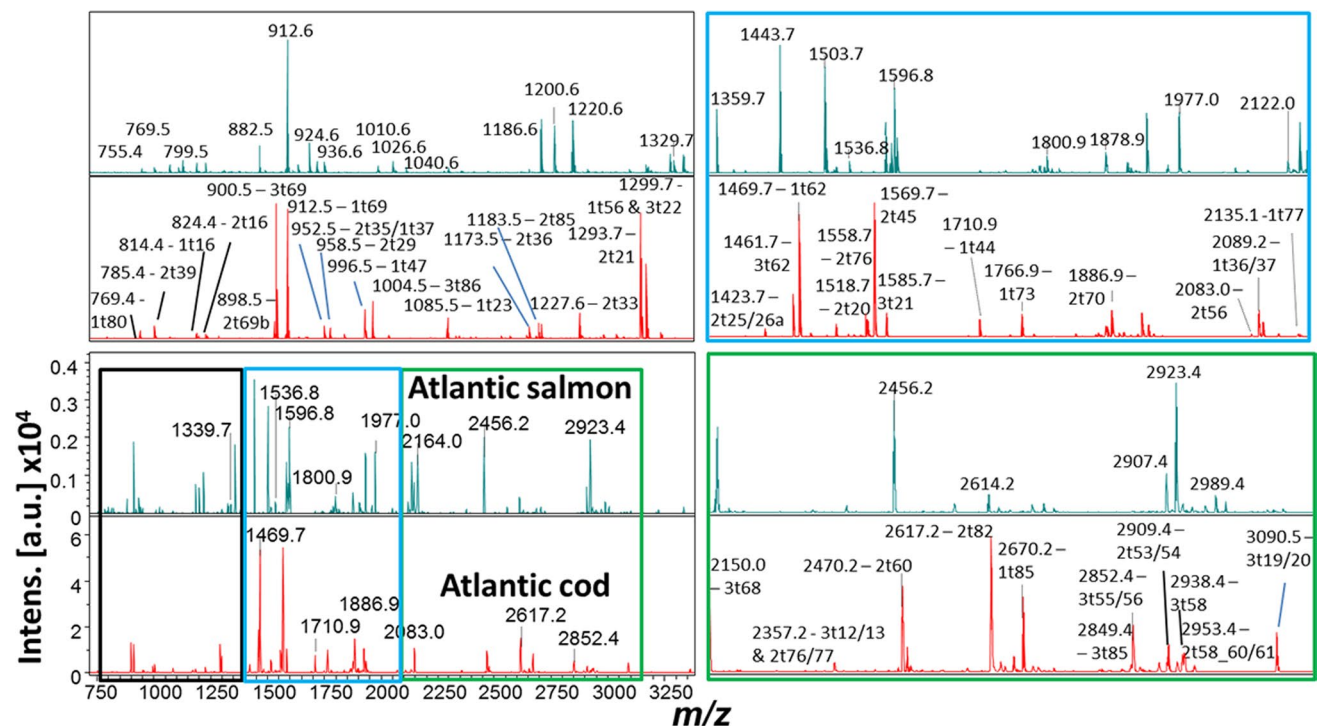


Fig. 1 Modern reference spectra with zoomed in portions (inset) of collagen digests analysed by MALDI-TOF mass spectrometry and annotated according to sequence information from LC–MS/MS

search of *Gadus morhua* (see Supplementary Table S15 for label sequence information) in comparison to that of *Salmo salar*

(Supplementary Table S1). For example, a single peak at m/z 900 (3t69; note m/z values throughout have been rounded down) is apparently exclusive to, but not present in all, the Gadiformes represented in 17 of the 20 modern reference species for this group. Unfortunately, biomarkers were not as readily observed throughout the higher-taxon rankings where, conversely, the greater the range of reference taxa acquired, the less likely these are to be retained (i.e. the more gadiformes within our reference database, the less likely there are to be biomarkers unique to the whole group). Another example, albeit on only two specimens, is that of the clupeids (and potentially Clupeiformes as a whole) having a peak at m/z 1471 (distinct from the isotopic envelope stemming from the peak at m/z 1469 seen across many fish taxa). Some of these groups of taxa can be recognised by small sets of peptide biomarkers, e.g. the combinations of peaks at m/z 1429 and 1443 are only seen in the four cyprinids of this study, whereas the combination of peaks at m/z 1445, 1469 and 1558 is only seen in sparids (whereas lack of the latter could also reflect the mugilids). The majority of Scorpaeniformes all appear to have a peak at m/z 1480 (except *Myoxocephalus*; Supplementary Fig. S1), and there appears to be a peak at m/z 1612 common to all the Pleuronectiformes (Pleuronectidae; *L. limanda*, *P. platessa*, *P. flesus* and *H. hippoglossus*, albeit not *M. kitt*), but most of the remaining taxonomic orders appear to yield highly complex variations in spectra amongst their species, particularly where we have larger numbers of species in the reference collection, such as for the Perciformes (Supplementary Table S1). Furthermore, our attempts at creating a scoring mechanism for comparing whole fingerprints resulted in substantial variation, ranging from only 72% of shared peaks for two modern specimens of *Pleuronectes platessa* to 86% overlap for two specimens of *Platichthys flesus* (although we have seen this range higher for other modern taxa not within this study).

Species identification of the archaeological fish bones

Of the 77 archaeological specimens, 71 yielded distinct fingerprints initially categorised as groups (Groups 1–9; Fig. 2, Supplementary Fig. S14, S19 and Supplementary Table S2) with only six yielding poor fingerprints (following the criteria presented in Harvey et al. 2016). Spectra were processed similarly to modern samples by peak picking with a signal to noise threshold > 5 and the m/z values compared against the biomarker table derived from the modern reference material (Supplementary Table S1).

In addition to other dominant markers (e.g. particularly those within the region m/z 1400–1600), the peak at m/z 900 (3t69) was only observed in the reference taxa of Gadiformes and found in archaeological Groups 1, 4 and 5 (Fig. 3, and Supplementary Figs. S15–20), which also

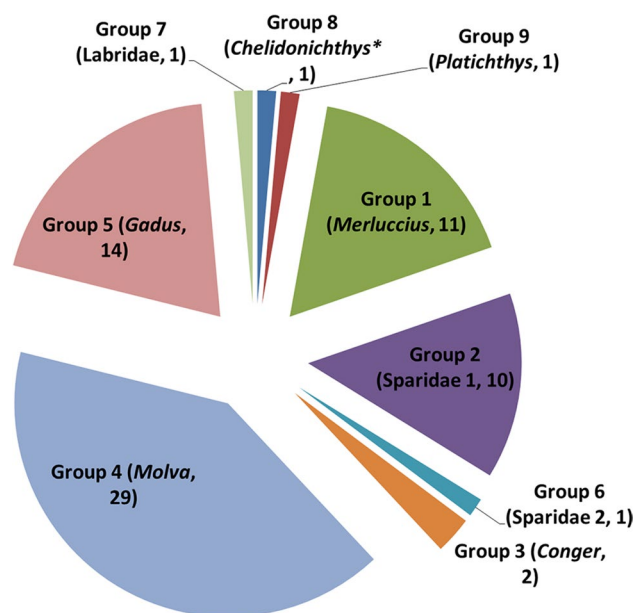
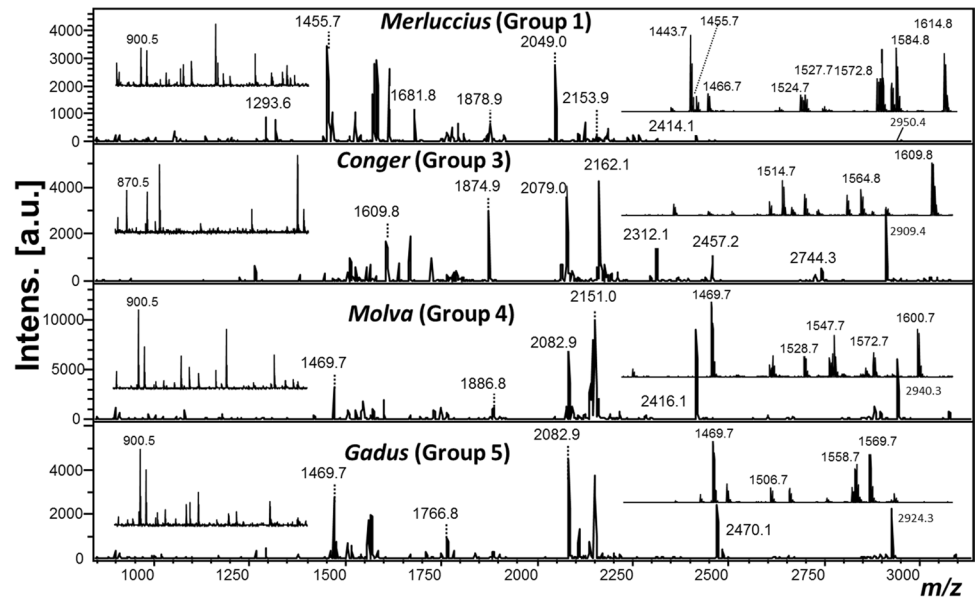


Fig. 2 Composition of the taxonomic groups identified from the subsample of fish remains from the Newport ship wreck, analysed by ZooMS (*more specifically closer to *C. lucerna*)

dominated the assemblage (Fig. 2). More specifically, the peaks at m/z 900.5, 1443, 1572 and 1614 (among many other more speculative markers compared with these that are supported with tandem mass spectral interpretation; Supplementary Table S1) imply that Group 1 ($n = 11$) most closely matches *Merluccius merluccius*; however, we do note some minor differences (this yielded a lower than expected peak overlap match of 52% given our biomarker matches, but there are no other relatives expected in these regions closer than other species we have included in this study). Numerous peaks in Group 4 ($n = 29$) spectra, including m/z 1469, 1528, 1600 and 2416, imply that they derive from the common ling (*Molva molva*) and are clearly distinct from its closest relative the blue ling (*Molva dypterygia*); the overlap score for this group was 68%. Group 5 ($n = 14$) showed a major peak at m/z 14612 indicative of *G. morhua*, the Atlantic cod, with other corroborating peaks (e.g. m/z 1469 and 1558), and a good overlap score of 73%. Groups 2 ($n = 10$) and 6 ($n = 1$) proved to be more problematic to identify to species, but we did find that peaks in the m/z 1400–1650 region (e.g. m/z 1445, 1469 and 1533) were most useful at placing both within the Sparidae family of the Perciformes (Supplementary Fig. S21–24). However, they were too dissimilar from all studied here to suggest an accurate match to any of the eight genera represented and so unsurprisingly the overlap scores to the closest sparid (*Pagellus* based on biomarker matches) were our lowest at 33% and 46% respectively. Group 7 ($n = 2$) appears to closely match the reference ballan wrasse (*Labrus bergylta*), with peaks such as

Fig. 3 Example collagen peptide mass fingerprints of four of the nine groups (see Supplementary Fig. S19 for the remaining five groups)



at m/z 1455, 1500 and 1548, but given the number of peak differences and our limited reference material (although a good overlap score of 74%), this group could potentially derive from another member of the Labridae (Supplementary Fig. S25–26). Group 8 ($n=1$) more closely matches that of the tub gurnard (*Chelidonichthys lucerna*; Supplementary Figs. 27–28) with peaks at m/z 1445, 1480 and 1555 although with a poor overlap score (39%). Group 3 ($n=2$) appeared to match closely the reference spectra for conger eel (*Conger conger*) through peaks at m/z 1445, 1467, and 1514, and also has no other relative that it could derive from closer than what we have tested in this study (*Anguilla*); the overlap score was somewhat also mediocre at 61%. Lastly, a peak at m/z 1630 suggests that Group 9 ($n=1$) appears to derive from a member within the Pleuronectidae family (more specifically only seen in *Platichthys*, *Pleuronectes* and *Microstomus*), with peaks of m/z 1469 and 2499 more specifically indicative of the European flounder (*Platichthys flesus*; Supplementary Fig. S24–25); the overlap score was also highest amongst these groups at 83%.

Methodological considerations

Analysis of the archaeological spectra indicates that identification of various taxonomic ranks is possible through collagen fingerprinting, which in turn is able to yield meaningful relative proportions of taxa represented within the assemblages (Fig. 2). Although we could not identify three of the six groups beyond family level, we were able to determine the three most dominant groups to genus/species level, showing the dominance of ling (> 40%) and substantial amounts of both cod (> 20%) and hake (~20%). Two of the groups that we could not identify beyond family level due

to limitations in the reference library were from the sparids (one of which was the fourth most abundant group), but despite eight sparid genera being included in this study, this reflects only two-thirds of known from this region (missing *Spicara*, *Pagrus*, *Diplodus* and *Centracanthus*). A small number of other taxa were also identified, including conger eel, wrasse family, grey gurnard and European flounder.

However, although confident matches were possible for some groups (1, 4 and 5), for others it was hindered by limitations in the extent of reference material, such as potentially the blackspot seabream (*Pagellus bogaraveo*) as discussed above (Supplementary Table S1). In order to overcome these, and allow for ZooMS collagen fingerprinting to be used more widely for fish taxa, firstly, a larger set of modern reference species is required, simply to allow more species to be identified. Although there is a steady increase in published fish reference biomarkers, these are most commonly done within limited taxonomic groups (e.g. Richter et al. 2011; Rick et al. 2019; Guiry et al. 2020). Here, we attempted to choose a wide range of taxa most common for the European region, with a reference set almost as large as the archaeological assemblage itself. Yet, this is still too small to completely resolve all identifications to species level, as is the theoretical level of identification possible for the ancient samples. Secondly, whilst this study is a preliminary attempt at identifying putative discriminatory sets of peptide markers, it is important to note that there is a lack of biological replicates for each individual sample. Before ZooMS can be confidently used for identification across a suitably wide range of fish taxa, many further samples are needed for analyses that ideally include biological replicates. Furthermore, with no two samples belonging to the same genus in many instances, it is not possible to discriminate

any of the identified spectra at a species level. Thus, the biomarkers represented here can only indicate to which genus the archaeological spectra might belong. For example, Group 1 showed indications of being *M. merluccius*; however without a much greater range of related taxa being represented in our reference material, it is not yet possible to place high confidence in the species designation.

A comparison with morphological analysis of the larger assemblage

The results of identification by skeletal morphological characteristics and those achieved using ZooMS are not directly comparable because they represent two distinct series of bone specimens. However, a number of observations can be made in considering the results achieved through identification of fish remains based on morphological characteristics (Russ 2012) and those achieved using ZooMS (Table 1).

Firstly, of the 71 remains identified by ZooMS (Fig. 2), 69 of the resulting identifications were consistent with the morphological analysis of the full Newport Ship assemblage, despite these ZooMS identifications not being carried out on the same bones (with biases against such overlap). The 161 specimens identified to at least sub-class level on morphological grounds (by HR) included taxa that were not identified by the ZooMS analysis; for the most part, these were species that formed only a minor proportion of the assemblage (by count), including Atlantic salmon (*Salmo salar*), tusk (*Brosme brosme*) and shark/skate/ray (Elasmobranchii). However, one species that was frequently identified in the archaeological material by morphology (HR),

but surprisingly absent in the ZooMS identifications, was the Atlantic herring (*Clupea harengus*). This anomaly likely results from the skeletal element bias introduced during sampling of material for ZooMS, which focussed on morphologically unidentifiable components of the assemblage, compared with the bias of morphologically identifiable material towards particular skeletal elements of particular taxa; preservational biases of archaeological bone have been well known for decades (Brewer 1992; Kidwell and Flessa 1996), even with studies of fish remains (Nagaoka 2005).

Both methods indicated that gadiform fish (cod order) were well represented, and that within this order members of the Merlucciidae, Gadidae and Lotidae families were present. Both approaches identified ling; morphological analysis did not distinguish between the blue and common ling, while ZooMS specifically identified the latter. European hake (*Merluccius merluccius*) was identified by morphological analysis for which ZooMS analysis also matched 11 specimens but with some peak differences (see above; Fig. 3). Atlantic cod was also identified both by morphology (by HR) and by ZooMS, while tusk (*B. brosme*), albeit only a single bone, was identified by morphology but not in the ZooMS samples. Outside of the Gadiformes order, both methods identified small numbers of conger eel (*Conger conger*), sparids (but see below) and flatfish (to the species level by ZooMS and to the family level by morphology, but see below).

Both methods provided instances where more precise identifications were possible; morphological analysis of the larger assemblage (by HR) identified the blackspot seabream as a Sparidae family fish present in the assemblage and that

Table 1 A comparison of identifications for fish remains from the Newport Ship by Russ using morphological characteristics (Russ 2012) and a morphologically unidentifiable sub-sample analysed by ZooMS, number of identified specimens (NISP)

Common name	Species	Taxonomic Group	Morphology	ZooMS
Atlantic herring	<i>Clupea harengus</i>	Clupeidae	98	0
Atlantic Salmon	<i>Salmo salar</i>	Salmonidae	1	0
European hake	<i>Merluccius merluccius</i>	Merlucciidae	20	12
Atlantic cod	<i>Gadus morhua</i>	Gadidae	7	14
Ling ¹	<i>Molva molvaldypterygia</i>	Lotidae	13	29
Tusk	<i>Brosme brosme</i>	Lotidae	1	0
Cod order		Gadiformes	8	0
Blackspot seabream ²	<i>Pagellus bogaraveo</i>	Sparidae 1	2	10
Seabream family		Sparidae 2	1	1
European conger	<i>Conger conger</i>	Congridae	2	2
Flatfish family ³		Pleuronectidae	2	1
Shark/skate/ray		Elasmobranchii	6	0
Wrasse family		Labridae	0	1
Tub gurnard	<i>Chelidonichthys lucerna</i>	Triglidae	0	1
		Total	161	71

¹ Specifically *Molva molva* for ZooMS.

² 'Sparidae 1' for ZooMS.

³ Specifically *Platichthys flesus* for ZooMS.

potentially two Sparidae family fish were present in total (Russ, 2012); ZooMS identified two Sparidae family fish, but in this case it was not possible to narrow the identification any lower than family level due to the aforementioned restrictions of the size of the reference database. The blackspot seabream, however, was not included in the modern sample; therefore, it is possible that the method would have identified one of the sparids to this species had the data for a modern specimen been included. Interestingly, morphological analysis of the larger assemblage (by HR) identified two specimens to the righteye flounder family Pleuronectidae, which includes a number of species that would be viable candidates for the specimens, such as European plaice, European flounder and dab (to mention only a few). The ZooMS analysis identified one specimen as European flounder (Fig. 3), further demonstrating the capability of this method to provide species-level identification where morphological approaches cannot. It is also clear that ZooMS provides the potential to identify a much wider range of species than morphology alone, but requires further development of the reference libraries.

Although both methods identified some of the more frequently occurring taxa, both also indicated the presence of fish families and species not identified by the other. Morphological analysis identified numerous remains of Atlantic herring ($n=98$), as well as small numbers of Atlantic salmon ($n=1$) and shark/skate/ray ($n=6$), none of which was identified in the ZooMS sample. Conversely, tub gurnard ($n=1$) and a wrasse family fish ($n=1$), possibly ballan wrasse, were identified by ZooMS but not during morphological identification.

A ‘traditional’ analysis of fish remains from archaeological sites, however, considers far more than taxonomic identification, and must always be a precursor to any form of analysis that may be destructive and/or only provide biomolecular data, such as ZooMS and isotopic analyses. In the case of the Newport Ship, ZooMS increased the number of identified taxa within the assemblage by two (grey gurnard and wrasse) and improved the taxonomic level of identification in two cases: from *Molva molva/dypterygia* to *Molva molva* and from Pleuronectidae to *Platichthys flesus*. ZooMS was applied in this case to assess its potential to increase species diversity in an assemblage, and to explore whether or not the standard interpretation had been missing some taxa during fish bone identification as a result of fragmentation and/or biases in the frequency and survival of species diagnostic elements. The results of this study do indicate the possibility that studies using morphological identification alone may be providing a biased and inaccurate view of past fish populations and human fishing practices. In the case of the Newport Ship, the ZooMS results increased species diversity when considered alongside the morphological results, but in terms of quantification only contributed to

increasing the number of identified specimens (NISP). When evaluating the contribution fish made to human diet in the past, it is more common for the minimum number of individuals (MNI) to be considered. The potential for ZooMS to contribute to increasing MNI for any given assemblage is currently restricted to identifying additional taxa, rather than increasing the number of fish represented within an already identified taxon.

Increasing species diversity and providing more precise taxonomic identification has the potential to make an extremely valuable contribution to the understanding of past environments, climate, fish populations and human fishing practices. As fisheries industries adapt to government regulations aimed at protecting fish species that are endangered by the impacts of fishing as well as environmental and climatic change, it is critical that an accurate understanding of past fish populations is available. While the ZooMS analysis carried out here is only a small step in evaluating its applications in the study of fish remains, it has demonstrated that our knowledge of fishes in the past is incomplete. The results suggest that both methods can provide results that are representative of the main fish taxa within an archaeological assemblage, but potentially with some loss of information for species, especially those that form a minor component. This is especially true for a ‘traditional’ morphological approach, which in this case considered all fish remains recovered from the Newport Ship (where only 161 of 659 were identified to some taxonomic level at least at the level of fish), unlike the ZooMS study that considered only a ~20% sub-sample of the ‘unidentified’ portion of the assemblage. Therefore, it is very likely that some of the species that were morphologically identified (HR) were not represented in the sub-sample to begin with, rather than ZooMS failing to identify them (given the high (> 90%) success rates). However, the results here could be used to suggest that further consideration needs to be taken into account with regards to subsampling strategies if aiming to infer a more accurate assemblage composition.

Conclusion

This study has been able to show the existence of putative marker sets that allow for the differentiation of one fish species from another through collagen fingerprinting, spanning a wide taxonomic range. Archaeological remains were analysed and inferences about the specimens were made using the marker reference table developed here with some nine distinct groups identified, showing a dominance of ling, and substantial amounts of cod, hake and an unidentified sparid. Whilst the confidence in the specificity of such inferences was limited due to a lack of biological replicates with species and genera, the study has highlighted potential markers

for several taxonomic levels, including the Gadiformes order and the Pleuronectidae family.

There is some evidence that applying ZooMS in the analysis of fish remains from archaeological sites may increase the levels of identifiable material and the range of species represented. Furthermore, it is possible that ZooMS may change the frequency that different fish species are represented, which may work to alter our understanding of fish exploitation, trade and consumption. To test this hypothesis, a larger archaeological assemblage needs to be analysed using both methods, and against a wider range of spectra. Fishes remain a particular technical challenge for collagen fingerprinting due to the vast size and diversity of the group, and the possible structural and functional variation within the collagen of these species. However, it is possible that geographical constraints be considered part of this process. This would be of particular use when considering orders such as the Perciformes, the largest order of vertebrates on the planet (Nelson, 2006). A group this size would prove technically and logistically challenging to analyse for taxonomic discrimination and consequently would require focusing on smaller groups within, such as families and genera. Given the current lack of available sequences for most of these species and the difficulties this caused in confirming homology for likely biomarkers across the range of species included here (Supplementary Table S1), we consider the simplest route forward would be to utilise machine learning based approaches of peptide biomarker determination (e.g. Gu & Buckley 2018). This approach is likely to be more useful than the attempted ‘overlap’ scoring approach presented here, but we also recognise that these machine learning approaches require multiple specimens of each reference taxon. Overcoming such challenges will be a key consideration for continuing research in this area, at least until further genetic data becomes available, but should come about naturally with ever-expanding sets of analyses.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12520-021-01478-y>.

Acknowledgements Many thanks are due to the Newport Medieval Ship Project, with access to and permission to carry out analyses on the material from the Newport Ship granted by curator Dr Toby Jones. We also thank Umberto Albarella for access to University of Sheffield reference material.

Funding The authors received support of the Royal Society for funding to MB and the University of Manchester’s Faculty of Science and Engineering for a Dean’s award to VLH. Royal Society, UF120473, Michael Buckley

Availability of data and material In addition to the supplementary files, and supplementary figures, peak lists have been provided for the best of each archaeological group are provided. Sequencing data for archaeological materials are available via ProteomeXchange with identifier PXD021207.

Declarations

Conflict of interest The authors declare no competing interests.

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

References

- Allentoft ME, Collins M, Harker D, Haile J, Oskam CL, Hale ML, Campos PF, Samaniego JA, Gilbert MTP, Willerslev E, Zhang G, Scofield RP, Holdaway RN, Bunce M (2012) The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proc Roy Soc B* 279:4724–4733
- Barrett JH, Locker AM, Roberts CM (2004) “Dark Age Economics” Revisited: The English Fishbone Evidence AD 600–1600. *Antiquity* 78:618–636
- Barrett JH, Johnstone C, Harland J, Van Neer W, Eryvynck A, Makowiecki M, Heinrich D, Hufthammer AK, ABEnghoff, Amundsen C, Christiansen JS, Jones AKG, Locker A, Hamilton-Dyer S, Jonsson L, Lõugas L, Roberts C, Richards M (2008) Detecting the medieval cod trade: a new method and first results. *J Archaeol Sci* 35:850–861
- Barrett JH, Orton DC, Johnstone C, Harland J, Van Neer W, Eryvynck A, Roberts C, Locker A, Amundsen C, Enghoff IB, Hamilton-Dyer S, Heinrich D, Hufthammer AK, Jones AKG, Jonsson L, Makowiecki D, Pope P, O’Connell TC, de Roo T, Richards M (2011) Interpreting the expansion of sea fishing in medieval Europe using stable isotope analysis of archaeological cod bones. *J Archaeol Sci* 38:1516–1524
- Barrett JH (2019) An environmental (pre)history of European fishing: past and future archaeological contributions to sustainable fisheries. *J Fish Biol* 2019:1–12
- Barrett JH, Orton DC (2016) Cod and herring. The archaeology and history of medieval sea fishing. Oxbow, Oxford
- Brewer DJ (1992) Zooarchaeology: method, theory, and goals. *Archaeological Method and Theory* 4:195–244
- Brodsky B, Ramshaw JA (1997) The collagen triple-helix structure. *Matrix Biol* 15(8–9):545–554
- Brown A, Turner R, Pearson C (2010) Medieval fishing structures and baskets at Sudbrook Point, Severn Estuary, Wales. *Mediev Archaeol* 54:346–361
- 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
- Buckley M, Fariña RA, Lawless C, Tambusso PS, Varela L, Carlini AA, Powell JE, Martinez JG (2015) Collagen sequence analysis of the extinct giant ground sloths *Lestodon* and *Megatherium*. *PLoS ONE* 10:11
- Buckley M, Gu M, Shameer S, Patel S, Chamberlain AT (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
- 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
- Buckley M, Recabarren OP, Lawless C, García N, Pino M (2019) A molecular phylogeny of the extinct South American gomphothere through collagen sequence analysis. *Quat Sci Rev* 224:105882
- 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
- Buckley M, Herman J (2019) Species identification of Late Pleistocene bat bones using collagen fingerprinting. *Int J Osteoarchaeology* 29:1051–1059
- Buckley M (2018) Zooarchaeology by mass spectrometry (ZooMS) collagen fingerprinting for the species identification of archaeological bone fragments. In *Zooarchaeology in practice* (pp. 227–247). Springer, Cham.
- Costello MJ, Bouchet P, Emblow CS, Legakis A (2006) European marine biodiversity inventory and taxonomic resources: state of the art and gaps in knowledge. *Mar Ecol Prog Ser* 316:257–268
- Costello MJ, Coll M, Danovaro R, Halpin P, Ojaveer H, Miloslavich P (2010) A census of marine biodiversity knowledge, resources, and future challenges. *PLoS ONE* 5:e12110
- Driver JC (1992) Identification, classification and zooarchaeology. *Circaea* 9:35–47
- Guiry EJ, Buckley M, Orchard TJ, Hawkins AL, Needs-Howarth S, Holm E, Szpak P (2020) Deforestation caused abrupt shift in Great Lakes nitrogen cycle. *Limnol Oceanogr*. <https://doi.org/10.1002/lno.11428>
- Harland JF, Jones AKG, Orton DC, Barrett JH (2016) Fishing and fish trade in Medieval York: the zooarchaeological evidence. In: Barrett JH, Orton DC (eds.) *Cod and herring. The archaeology and history of Medieval sea fishing*. Oxford: Oxbow. Pp: 172–204.
- Harvey VL, Egerton VM, Chamberlain AT, Manning PL, Buckley M (2016) Collagen fingerprinting: a new screening technique for radiocarbon dating ancient bone. *PLoS ONE* 11:3
- Harvey VL, Daugnora L, Buckley M (2018) Species identification of ancient Lithuanian fish remains using collagen fingerprinting. *J Archaeol Sci* 98:102–111
- 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(8):201955
- Hoffmann RC (1996) Economic development and aquatic ecosystems in medieval Europe. *Am Hist Rev* 101:631–669
- Hoffmann RC (2005) A brief history of aquatic resource use in medieval Europe. *Helgoland Mar Res* 59:22–30
- 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 International Journal of Osteoarchaeology* 13(3):120–131
- Kidwell SM, Flessa KW (1996) The quality of the fossil record: populations, species, and communities. *Annu Rev Earth Pl Sci* 24:433–464
- Kimura S, Miyauchi Y, Uchida N (1991) Scale and bone type I collagens of carp (*Cyprinus carpio*). *Comparative biochemistry and physiology*. *Comp. Biochem* 99(2):473
- Kimura S, Ohno Y (1987) Fish type I collagen: tissue-specific existence of two molecular forms, ($\alpha 1$)₂ $\alpha 2$ and $\alpha 1$, $\alpha 2$, $\alpha 3$, in Allaska pollack. *Comp Biochem Phys B* 88:409–413
- Kowaleski M (2003) The commercialization of the sea fisheries in Medieval England and Wales. *Int J Marit Hist* 15(2):177–231. <https://doi.org/10.1177/084387140301500212>
- Llamas B, Valverde G, Fehren-Schmitz L, Weyrich LS, Cooper A, Haak W (2017) From the field to the laboratory: controlling DNA contamination in human ancient DNA research in the high-throughput sequencing era. *STAR: Science & Technology of Archaeological Research* 3(1):1–14
- Lozano LF, Peña-Rico MA, Jang-Cho H, Heredia A, Villarreal E, Ocotlán-Flores J, Gomez-Cortes AL, Aranda-Manteca FJ, Orozco E, Bucio L (2002). Thermal properties of mineralized and non mineralized type I collagen in bone. *MRS Online Proceedings Library Archive*, 724.
- Ludwig A, Arndt U, Debus L, Rosello E, Morales A (2009) Ancient mitochondrial DNA analyses of Iberian sturgeons. *J Appl Ichthyol* 25(1):5–9
- Malmström H, Storå J, Dalén L, Holmlund G, Götherström A (2005) Extensive human DNA contamination in extracts from ancient dog bones and teeth. *Mol Biol Evol* 22(10):2040–2047
- Morvan-Dubois G, Le Guellec D, Garrone R, Zylberberg L, Bonnaud L (2003) Phylogenetic analysis of vertebrate fibrillar collagen locates the position of zebrafish $\alpha 3$ (I) and suggests an evolutionary link between collagen α chains and Hox clusters. *J Mol Evol* 57(5):501–514
- Müldner G (2016) Marine fish consumption in medieval Britain: the isotope perspective from human skeletal remains. In: Barrett JH, Orton DC (eds.) *Cod and herring. The archaeology and history of Medieval sea fishing*. Oxford: Oxbow. Pp: 239–249.
- Nagaoka L (2005) Differential recovery of Pacific Island fish remains. *J Archaeol Sci* 32:941–955
- Nayling N, Jones T (2014) The Newport Medieval Ship, Wales, United Kingdom. *Int J Nautical Archaeol* 43:239–278
- Nelson JS (2006) *Fishes of the world. Fourth edition.*, i–xix, 1–601.
- Olson C, Walther Y (2007) Neolithic cod (*Gadus morhua*) and herring (*Clupea harengus*) fisheries in the Baltic Sea, in the light of fine-mesh sieving: a comparative study of subfossil fishbone from the late Stone Age sites at Ajvide, Gotland, Sweden and Jettböle, Åland, Finland. *Environ Archaeol* 12:175–185
- Oosting T, Star B, Barrett JH, Wellenreuther M, Ritchie PA, Rawlence NJ (2019) Unlocking the potential of ancient fish DNA in the genomic era. *Evol Appl* 12(8):1513–1522
- Orton DC, Makowiecki D, de Roo T, Johnstone C, Harland J, Jonsson L, Heinrich D, Enghoff IB, Løugas L, Van Neer W, Ervynck A (2011) Stable isotope evidence for late medieval (14th–15th C) origins of the eastern Baltic cod (*Gadus morhua*) fishery. *PLoS ONE* 6(11):e27568
- Radu V (2005) *Atlas for the identification of bony fish bones from archaeological sites*. Asociația Română de Arheologie, Studii de Preistorie Supplementum 1. București : Contrast.
- Ramshaw JA, Shah NK, Brodsky B (1998) Gly-X-Y tripeptide frequencies in collagen: a context for host-guest triple-helical peptides. *J Struct Biol* 122:86–91
- Rich A, Crick FHC (1961) The molecular structure of collagen. *J Mol Biol* 3:483–IN4
- Richter KK, Wilson J, Jones AKG, Buckley M, van Doorn N, Collins MJ (2011) Fish ‘n chips: ZooMS peptide mass fingerprinting in a 96 well plate format to identify fish bone fragments. *J Archaeol Sci* 38:1502–1510
- 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
- Russ H (2012) Newport Medieval Ship Project. Specialist Report: Fish Remains. https://archaeologydataservice.ac.uk/archiveDS/archiveDownload?t=arch-1563-2/dissemination/pdf/Newport_Medieval_Ship_Specialist_Report_Fish_Remains.pdf
- Rybczynski N, Gosse JC, Harington CR, Wogelius RA, Hidy AJ, Buckley M (2013) Mid-Pliocene warm-period deposits in the High Arctic yield insight into camel evolution. *Nat Commun* 4:1–9
- Starkey DJ, Reid C, Ashcroft N (2000) Sea fisheries: the commercial sea fisheries of England and Wales since 1300. Chatham, London

- Serjeantson D, Woolgar, CM (2006) Fish consumption in Medieval England. In: Woolgar CM, Serjeantson D, Waldon T (eds) *Food in Medieval England; diet and nutrition*. Oxford University Press, Oxford, pp 102–130
- van der Sluis LG, Hollund HI, Buckley M, De Louw PG, Rijdsdijk KF, Kars H (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* 416:80–91
- Van Neer W, Ervynck A (2016) The rise of sea-fish consumption in inland Flanders, Belgium. In: Barrett JH, Orton DC (eds.) *Cod and herring. The archaeology and history of Medieval sea fishing*. Oxford: Oxbow. Pp: 156–171
- Woolgar C (2000) Take this penance now, and afterwards the fare will improve: seafood and late medieval diet. In: Starkey DJ, Reid C, Ashcroft N (eds) *England's sea fisheries: the commercial sea fisheries of England and Wales since 1300*. Chatham, London, pp 36–44

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.