Comparing ancient DNA survival and proteome content in 69 archaeological cattle tooth and bone samples from multiple European sites

Caroline Wadsworth [a], Noemi Procopio [a], Cecilia Anderung [b], José-Miguel Carretero [c,d], Eneko Iriarte [c], Cristina Valdiosera [d,e], Rengert Elburg [f], Kirsty Penkman [g], Michael Buckley [a,*]

[a] Manchester Institute of Biotechnology, The University of Manchester, Manchester M1 7DN, UK
[b] Department of Ecology and Genetics, Uppsala University, Norbyvägen 18D, SE-752 36, Uppsala, Sweden
[c] Laboratorio de Evolución Humana, Universidad de Burgos, Edificio I+D+i Plaza de Misael Bañuelos s/n, 09001, Burgos, Spain
[d] Department of Archaeology and History, La Trobe University, Melbourne, VIC 3086, Australia
[e] Centro Mixto, Universidad Complutense de Madrid, Instituto de Salud Carlos III de Evolución y Comportamiento Humanos, 28029 Madrid, Spain
[f] Archaeological Heritage Office Saxony, Zur Wetterwarte 7, 01109 Dresden, Germany
[g] Department of Chemistry, University of York, York, YO10 5DD, UK

ARTICLE INFO

ABSTRACT

Ancient DNA (aDNA) is the most informative biomolecule extracted from skeletal remains at archaeological sites, but its survival is unpredictable and its extraction and analysis is time consuming, expensive and often fails. Several proposed methods for better understanding aDNA survival are based upon the characterisation of some aspect of protein survival, but these are typically non-specific; proteomic analyses may offer an attractive method for understanding preservation processes. In this study, in-depth proteomic (LC-Orbitrap-MS/MS) analyses were carried out on 69 archaeological bovine bone and dentine samples from multiple European archaeological sites and compared with mitochondrial aDNA and amino acid racemisation (AAR) data. Comparisons of these data, including estimations of the relative abundances for seven selected non-collagenous proteins, indicate that the survival of aDNA in bone or dentine may correlate with the survival of some proteins, and that proteome complexity is a more useful predictor of aDNA survival than protein abundance or AAR. The lack of a strong correlation between the recovery of aDNA and the proteome abundance may indicate that the survival of aDNA is more closely linked to its ability to associate with bone hydroxyapatite crystals rather than to associate with proteins.

Significance: Ancient biomolecule survival remains poorly understood, even with great advancements in 'omics' technologies, both in genomics and proteomics. This study investigates the survival of ancient DNA in relation to that of proteins, taking into account proteome complexity and the relative protein abundances to improve our understanding of survival mechanisms. The results show that although protein abundance is not necessarily directly related to aDNA survival, proteome complexity appears to be.

© 2017 Published by Elsevier B.V.

1. Introduction

In the last two decades the field of biomolecular archaeology has rapidly expanded, and nowadays genomics and proteomics as well as other omics techniques are frequently applied to archaeology studies [1,2]. Analysis of ancient DNA (aDNA) is increasingly used worldwide, partly due to the advent of Next Generation Sequencing (NGS) technologies that have overcome some of the former limitations of aDNA (such as fragmentation, low copy number, damage and contamination). There are now numerous detailed and rigorous experimental and bioinformatic guidelines [3–7] set out for the successful extraction, sequencing and analysis of aDNA from a wide range of geographic and environmental areas, with recovery of aDNA as far back as 700k years in permafrost [8] and ca. 500 K in non-permafrost environments [7,9,10]. However, despite the recent technological revolution of NGS, the reality is that the routine use of these technologies is limited to only a few laboratories in the world, given the high costs of these procedures and the difficulties handling massive amounts of sequence data. There are therefore still a large number of aDNA studies that do not use NGS [11,12], and so improving our understanding of aDNA survival, as well as the development of a reliable screening method to determine the presence of aDNA in a sample would be a powerful tool, allowing only those bones most likely to contain aDNA to be selected for further analysis. Improved understanding in ancient biomolecule survival and improved screening techniques would be useful in preventing unnecessary sample destruction, minimising resource wastage and increasing the chances of successful extractions. Furthermore,

* Corresponding author.
E-mail address: m.buckley@manchester.ac.uk (M. Buckley).

understanding the molecular taphonomy of ancient bones and teeth through biomolecules (and particularly proteins given their role in bone structure) will build our knowledge of how degradation/preservation of organic material occurs within the archaeological record. Here we present for the first time a direct comparison of the ancient proteome and aDNA in bones from four different environments in a total of 69 ancient cattle samples as a means to investigate whether ancient protein survival is a useful biomarker for aDNA preservation.

1.1. Methods comparing aDNA survival with bone preservation

1.1.1. Histology

One of the earliest and simplest proposed methods of evaluating archaeological bone for its preservation, to the extent that it could screen for the presence of aDNA, involved detailed examination of the histology of the sample [13]. After the death and burial of an organism, its tissues undergo a number of diagenetic processes, e.g. the removal of collagen fibrils by saprophytic bacteria, (see Smith et al. [14] and Nielsen-Marsh et al. [15] for reviews) which can affect the microstructure of the bone [16,17]. The histological integrity of archaeological bone can be measured using BSE-SEM (back-scattered electron microscopy) and generally good bone preservation is correlated with highly ordered bone histology with well-defined physical features [18,19]. Good structural preservation is therefore correlated with improved biomolecule survival, with several studies indicating that better preserved bones are more likely to contain amplifiable aDNA [13] and other biomolecules [20,21]. Assessment of bone histology has also been combined with measurement of the total nitrogen content of the sample (as a proxy for total protein content and therefore an indication of biomolecule preservation), where bones with good histological preservation and high nitrogen content are considered most likely to have good biomolecule preservation [22,23]. However, histological structure is not inextricably linked with good biomolecule preservation, with some specimens showing poor histological preservations that were successful for aDNA analyses [13].

1.1.2. Amino acid racemisation

Amino acids with a single chiral centre exist in one of two isomeric forms – the D-enantiomer and the L-enantiomer. In life amino acids exist exclusively in the L- form, but after the death of the organism they begin to racemise to the D- form, to ultimately form an equilibrium of L- and D- isomers in roughly equally proportions, known as a racemic mixture. The rate of amino acid racemisation (AAR) is determined predominantly by time and temperature [24], but also by environmental factors. The racemisation of aspartic acid (Asp) has been proposed as a screening method for the presence of aDNA ([21], but see also [37]), measured as the Asx DL ratio as it also includes the extent of asparagine hydrolysis. AAR values were suggested as a good indicator of the likelihood of DNA survival as the kinetics of racemisation appeared to mirror the rates of DNA depurination [25–27], which is thought to be the major limiting factor in aDNA survival.

1.1.3. Proteomic analyses of ancient bone

Early investigations into biomolecule survival in archaeological remains indicated that some small bone proteins, such as osteocalcin (OC), may have had much greater survival rates than DNA in ancient bone due to their functionally important mineral-binding properties [28]. OC was particularly ideal for analysis because it was easy to isolate through application directly to a solid phase extraction cartridge [29] and was therefore potentially useful as a biomarker for the presence of aDNA. However, analysis of intact OC in archaeological bone using 'top-down' proteomic methods showed this not to be the case, with fewer samples containing intact OC than were successful for aDNA [30].

In recent years, several studies have been carried out on archaeological bone using bottom-up proteomic techniques [31,32], which have proved effective even at the low protein levels present in ancient bone. The advantage over top-down approaches is that it is possible to evaluate the presence of numerous non-collagenous bone proteins (NCPs) within a single analysis. The sensitivity, high success rate and relatively low running costs of such analyses in comparison to aDNA analysis means that proteomic analysis of archaeological remains is becoming an attractive tool for the characterisation and/or species identification of complex tissues [33,34]. However, they may also yield insights into aDNA survival and potentially offer an alternate screening method for bones most likely to contain aDNA. In this study, the success or failure of aDNA extractions from the same series of archaeological cattle bones and teeth was compared with their corresponding Asx DL values and bone proteomes in an attempt to identify potential protein biomarkers for the presence of aDNA and to improve our understanding of aDNA survival.

2. Materials and methods

We expand upon the findings presented in Buckley et al. [30] wherein the amino acid racemisation values, amino acid concentrations and the success or failure of aDNA and OC extractions were reported for 34 archaeological cattle bones. Herein, a total of 69 tooth and bone samples from archaeological cattle (including some of the 34 from [30]) were analysed by in-depth proteomic analysis, all of which had previously undergone aDNA extraction and analysis carried out by Anderung [35]. Of these 69 samples, 36 were teeth and 33 were bones (see Supplementary Table S1 for skeletal information). Samples were recovered from several different sites with different mean annual temperatures (MAT) and effective burial temperatures (T_{eff}): the Bronze Age site Asine (15 samples) and the preclassical site Lerna (four samples) both in Greece (T_{eff} 18 °C; ~2000–1500 BCE for Asine and ~500–800 BCE for Lerna); Zauschwitz (10 samples) and Dresden-Cotta (three samples) in Saxony, Germany (these 13 samples were grouped together and considered "Saxony"; T_{eff} 12.5 °C including a temporal range from 5500 to 600 BCE); and Bronze Age material from El Portalón cave in northern Spain (37 samples; T_{eff} 9.5 °C) (Supplementary Table S1). Of the 37 samples from El Portalón, 19 were taken from a collection that had been stored in a museum for several years and 18 were more recently excavated.

Mitochondrial DNA extraction, purification and analysis were carried out by Anderung [35] using hybridisation and magnetic bead separation following a modified method from [36]. Amino acid racemisation analyses were carried out as described in Buckley et al. [30]. Proteomes were obtained from ~50 mg bone powder per sample following Wadsworth & Buckley [32]. Peptide masses obtained via LC-MS/MS analysis were searched against the SwissProt database for matches to primary protein sequences using the Mascot search engine. Each search included the fixed carbamidomethyl modification of cysteine and the variable modifications for deamidation, pyroglutamate formation, and oxidation of lysine, proline and methionine residues. Enzyme specificity was limited to trypsin with up to 2 missed cleavages allowed, mass tolerances were set at 5 ppm for the precursor ions and 0.5 Da for the fragment ions and all spectra were considered as having either 2 + or 3 + precursors. Proteome complexity was calculated by manually examining peptide matches from Mascot searches for the relevant genus (*Bos* or *Bison*) which have an ion score above 20, and proteins were included in the count only if they had at least three unique high-confidence peptides matches.

Relative abundances were calculated using a label-free quantitation method carried out by Progenesis QI software. Three analyses on this dataset were run on the Progenesis QI software; one comparing the relative abundances of proteins in samples from the four different 'environmental groupings', one comparing protein abundances collectively in tooth samples vs bone samples and the final analysis comparing protein abundances between the museum and recently excavated samples from El Portalón cave site. Principal Component Analysis (PCA) was performed on normalized abundances exported from Progenesis using R

Software with the package FactoMineR, and plots were produced using the Python programming language (Python Software Foundation, https://www.python.org/). This facilitated a visual separation according to the variations in abundance of proteins in each sample, and helped to detect clustered data and/or outliers. To clean the data in order to leave out unreliable proteins, we also excluded from the study proteins for which the number of unique peptides was ≤1. We also used a protein score cut-off to remove proteins which score was smaller than the median score of the whole dataset (score = 122).

3. Results

Of the 69 samples tested, 56 yielded aDNA (31 teeth and 25 bone) and 13 samples failed to yield aDNA (5 teeth and 8 bone). NCPs were successfully identified in 66 samples (samples AS3, AS9 and LE4 yielded no identifiable proteins), with 49 of those samples containing 10 or more NCPs. As there are many different NCPs found in all samples, this study will focus on seven selected proteins: fetuin-A, prothrombin, pigment epithelium-derived factor (PEDF), lumican, chondroadherin (CHAD), secreted phosphoprotein 24 (SPP24) and matrix metalloproteinase-20 (MMP20), which have been commonly identified in ancient bone in other studies [31,32] and which have different biological functions and properties. In particular, three of the proteins of interest are known to be collagen-binding proteins (PEDF [37], lumican [38] and CHAD [39]) and three are mineral-binding proteins (fetuin-A [40], prothrombin [41] and SPP24 [42]); the choice of these specific proteins was also intended to compare their survival related to their specific localisation within the bone tissue (Table 1).

3.1. Variations in proteome complexity across different archaeological sites

The samples analysed in this study came from multiple different sites across Europe, which included different site types (e.g. open air, cave) and different MATs and T_{eff}. Collagens alpha 1 (type I) and alpha 2 (type I) (hereafter collagen $\alpha1(I)$ and $\alpha2(I)$ respectively) were identified in every sample as the highest scoring proteins (i.e., combined peptide scores) in Mascot database searches and had high relative abundances according to the Progenesis QI analysis. Four other collagen types were commonly identified but were not as ubiquitous as collagens $\alpha1(I)$ and $\alpha2(I)$, these are collagens $\alpha1(II)$, $\alpha1(III)$ and $\alpha1(XI)$ (in 53, 54, and 34 samples respectively) and collagen $\alpha2(XI)$ (in 66 samples).

When considering the NCPs, proteome complexity varies greatly between sites, and as expected the samples from sites with warmer climates which have a higher MAT and T_{eff} (Lerna and Asine) have poorer proteomes (less complex) than the samples from sites with lower T_{eff} and MATs (Saxony and El Portalón) even though some of the latter were older. The four samples from Lerna had the poorest protein complexities with one sample containing no NCPs (but five collagen types) and the other three samples containing one, four and six NCPs. The 15 samples from Asine had relatively poor protein complexities with ≤10 NCPs identified in all but one sample (which contained 18 NCPs). Samples recovered from Saxony had relatively well preserved proteins, with eight out of 13 samples containing ≥10 NCPs and three of these having over 20 identified NCPs. The 37 samples from El Portalón cave site have the richest proteomes; no sample has <10 identified NCPs, 12 have between 15 and 20 NCPs in total and 25 have over 20 identified NCPs (up to a maximum of 37 NCPs) (Fig. 1).

Two of the seven proteins of interest (fetuin-A and PEDF) were found in at least one sample from each site, with fetuin-A being identified in 62 samples and PEDF in 55 samples. Prothrombin, lumican and MMP20 are found in at least one sample from each site except Lerna (54, 45 and 32 samples respectively); SPP24 is not found in any samples from Asine or Lerna (but in 37 samples overall; mostly those from El Portalón cave; Table 1).

With regards to the amino acid racemisation data, there was no discernible correlation between proteome complexity and the Asx DL values obtained for each sample in this study. There was a weak correlation between Asx DL values and the MAT and T_{eff} in that five of the samples from El Portalón (the site with the lowest MAT and T_{eff} in the study) have the lowest recorded Asx DL values of 0.09 and one of the samples from Lerna, (the site with the highest T_{eff} in the study), has one of the highest Asx DL values of 0.24 (Fig. 1).

The number of NCPs was correlated with the MAT and T_{eff} of a site, where samples from the colder site (El Portalón) showed the highest proteome complexity, followed by samples from Saxony (with intermediate MAT and T_{eff}) and finally followed by samples from the warmer sites in Greece. There was also an interesting correlation between the age of different specimens from the same site and the number of NCPs, where the oldest samples from Saxony (DD4, DD11, DD75, DD76 and DD9) showed the lowest proteome complexity within all samples from Saxony (with the only exception for DD29, see Supplementary Table S1) (Fig. 1). However, there seemed to be no clear correlations between Asx DL and the number of NCPs detected (Fig. 1).

Ancient DNA survival in the samples analysed was generally good, with 56 of the 69 samples yielding aDNA. The success or failure of the aDNA analysis appears to be highly dependent on the site the samples were recovered from. The samples from El Portalón cave were the most successful in terms of aDNA extractions, as only one of the 37 samples did not yield any aDNA. Three of the 13 samples failed for aDNA retrieval from Saxony, with six failures out of 15 samples from Asine and Lerna was the worst in terms of preservation, with only one of the four samples being successful for aDNA analysis (Fig. 1). We noticed a correlation between the number of NCPs identified in a sample and the presence of aDNA; in fact, when the number of NCPs was equal or bigger than 15, usually aDNA recovery was successful (only two out of 13 samples with number of NCPs = 17 and 18 failed the aDNA analysis, see Supplementary Table S1) and that when the NCPs number was higher than 18, aDNA analyses were always successful.

3.2 Protein abundances across different archaeological sites

Principal component analysis of the relative abundances of each protein matched in each of the 69 samples indicated that not only are bone and tooth dentine proteomes distinct, but that there are signals specific

Table 1
Summary table showing seven selected proteins of interest and the number of samples they were identified in by site, sample type and whether they were successful or failed for aDNA analysis.

| Protein	Asine (15)	Lerna (4)	Saxony (13)	El Portalón (37)	Bone (33)	Tooth (36)	aDNA success (56)	aDNA failure (13)	Protein function
Fetuin-A	10	3	12	37	29	33	50	12	Osteogenesis, bone mineralization and resorption
Prothrombin	9	0	8	37	21	30	47	7	Associated with osteoclasts, coagulation factor
PEDF	9	1	11	34	25	31	49	9	Regulation of osteoblasts and osteoclasts differentiation
Lumican	5	0	7	33	23	23	41	5	Regulation of collagen fibril formation
CHAD	5	0	8	8	24	1	20	5	Regulation of chondrocytes growth and proliferation
SPP24	0	0	3	34	11	26	35	1	Involved in bone turnover and repair
MMP20	3	1	2	26	1	30	28	4	Tooth-specific MMP, functions in dental enamel formation
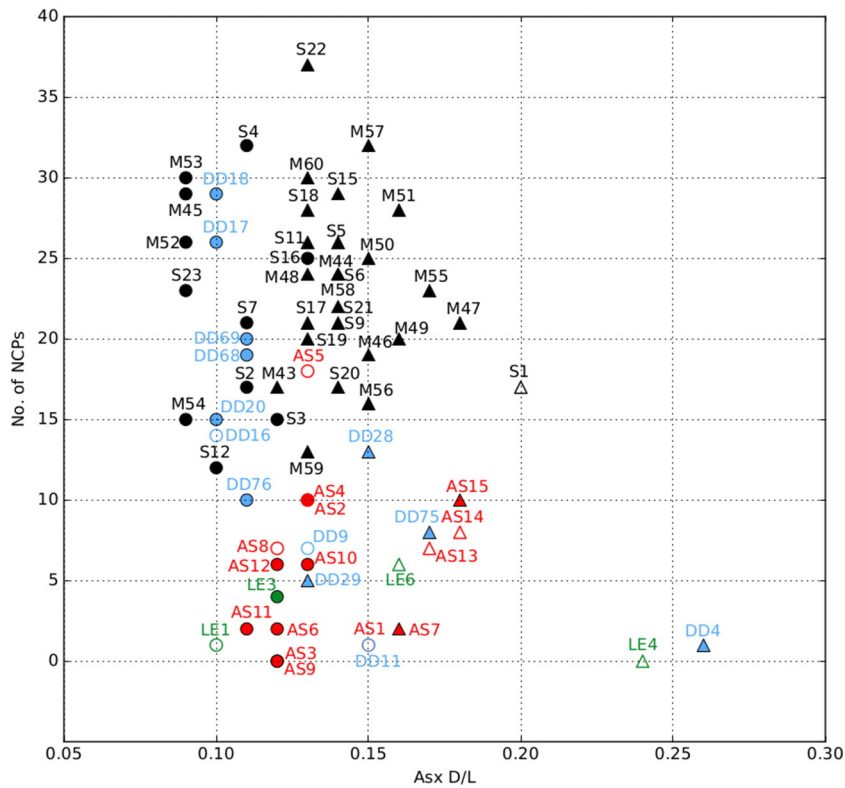
Fig. 1. Plot of the number of NCPs identified versus the Asx DL ratio in the 69 bone and teeth samples. Circles represent bone, whereas triangles represent tooth dentine samples; hollow symbols failed for aDNA recovery. Different colours indicate different sites (green = Lerna, black = El Portalón, red = Asine, blue = Saxony). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

to the archaeological site (Fig. 2), as well as to the approximate levels of degradation and 'thermal ages' of the samples [43]. The majority of NCPs detected in the archaeological samples were at least one order of

magnitude lower in abundance compared with the dominant collagen (collagen α2(I)). The only exception (not including albumin due to potential laboratory contamination issues) was fetuin-A, which was the
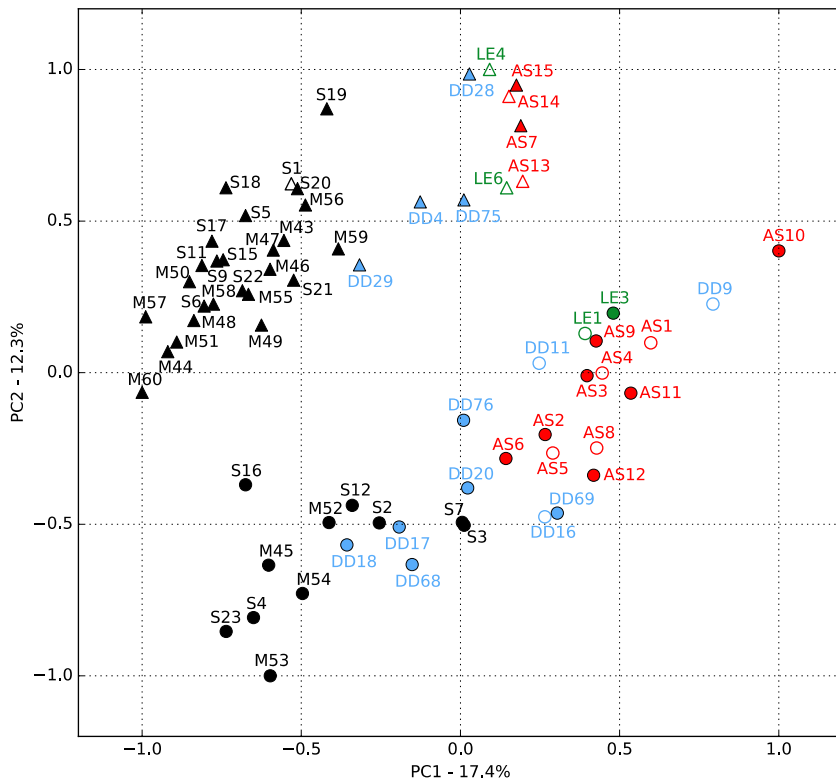

Fig. 2. PCA obtained using the complete quantitative proteomic data showing clusters originated from bone samples (circles) and tooth samples (triangles) from different sites (green = Lerna, black = El Portalón, red = Asine, blue = Saxony); hollow symbols represent those that failed aDNA.

most abundant protein recovered in all samples apart from those from Lerna, in which prothrombin was the most abundant (Fig. 3). Interestingly, fetuin-A and prothrombin are both mineral-binding proteins, and this suggests a potential better preservation of these proteins within the specimens compared with the collagen-binding proteins. MMP20 was abundant in the El Portalón dentine samples, whereas the dentine samples from the other sites showed low abundance values but when present they were noticeably greater than bone (Fig. 3). Similar to that observed for the proteome complexity and amino acid racemisation levels, the relative abundances do not show any particular correlation with AAR values (Supplementary Fig. S1) where many of the lowest Asx DL had some of the lowest relative abundances for many proteins.

3.3 Proteome changes during museum storage

Of the 37 samples from El Portalón cave site, 19 samples had been selected for analysis from a museum collection ('S' samples) and 18 ('M' samples) had been more recently excavated. Ancient DNA preservation was very good in the samples from El Portalón cave with only one sample (S1) from the museum samples failing to yield aDNA. Both the museum and recently excavated samples had similar ranges of proteome complexities (Fig. 1 and Fig. 2).

The average relative abundances of the seven proteins of interest are also very similar in the museum and recently excavated samples (see Supplementary Fig. S2) although the museum samples have slightly higher relative abundances for chondroadherin, lumican, PEDF, fetuin-A and prothrombin than the recently excavated samples, whereas the opposite is true for MMP20 and SPP24.

3.4 Comparison of bone and dentine proteomes

Of the 69 samples analysed in this study, 36 were tooth (dentine) samples and 33 were bone samples; the ranges of proteins recovered from the two types of sample were similar. However, of the seven selected proteins of interest focused on in this study, chondroadherin (thought to be cartilage-specific) appears to be more commonly (though not exclusively) identified in bone samples, and with greater relative abundance than in dentine samples. Interestingly, SPP24 and MMP20 were more commonly identified (and with greater relative abundances) in dentine samples; in particular, MMP20 is identified in 30 of the 36 tooth samples and only in one bone sample that originated from a mandible; this is in accordance with other studies which describe the role of MMP20 in dental enamel formation [44,45] (Table 1 and Fig. 4), where its observation in the one bone sample in this study may be due to contamination during the sampling process.

The success rate for aDNA extractions appears to be marginally higher in the tooth samples than bone samples, as only five of the 36 tooth samples failed to yield aDNA, compared to eight of the 33 bone samples. However, 26 of the 36 tooth samples were from El Portalón
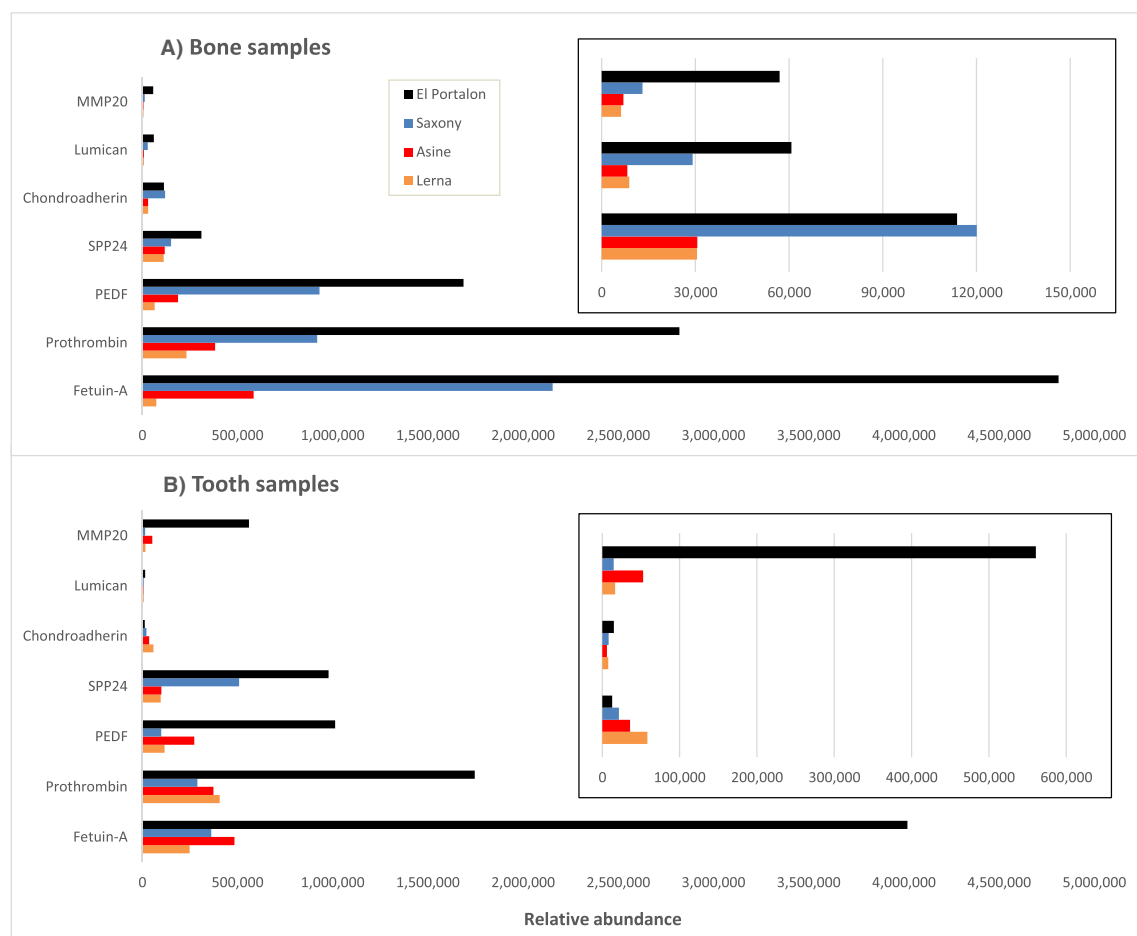

Fig. 3. Average relative abundances for seven selected identified proteins in samples from different sites for (A) bone and (B) tooth dentine; by comparison the relative abundances for COL1A2 in bone are 1.39×10^7, 1.19×10^7, 1.42×10^7 and 1.0×10^7 and in teeth are 1.19×10^7, 1.20×10^7, 1.36×10^7 and 1.21×10^7 for Lerna, Asine, Saxony and El Portalón respectively. Inset: zoomed in view of the uppermost three proteins (MMP20, lumican and chondroadherin). Note the difference in the scale within the two insets.

cave which appears to have the best conditions for biomolecular preservation of the sites studied here, so this increased success rate may be more related to the site conditions rather than whether the sample was from a tooth or bone.

4. Discussion

There are several studies which have suggested that proteins are more likely to have better survival rates than aDNA [46], as DNA is a more labile molecule than most proteins. In this study the samples which were successful for aDNA analysis often had richer proteomes than the samples which were unsuccessful, but there were also some exceptions (there were some samples which either had relatively high NCP counts and failed aDNA analysis or which had no identifiable NCPs but were successful for aDNA analysis).

Despite suggestions that the storage of archaeological skeletal materials in museums may promote biomolecular degradation [47], in this study no significant differences in either the frequency of successful aDNA extractions nor the total NCP counts were identified between the recently excavated and museum-stored samples from El Portalón cave; indeed the one sample from this site which failed the aDNA analysis (S1) had been recently excavated. However, it is important to note that both the museum-stored and more recently excavated samples tested here are exceptionally well preserved, and that more poorly preserved samples may degrade faster in a museum environment and therefore display a greater difference in biomolecule preservation between museum-stored and recently excavated samples.

It appears that protein and aDNA survival is heavily dependent on the MAT and T_{eff} of the site that they were recovered from (Fig. 1); although other factors such as hydrological activity and environmental pH could not be as easily evaluated. However, of the sites analysed here, the samples from sites with a higher T_{eff} (Lerna and Asine, T_{eff} 18 °C for both sites) had much poorer proteomes and a higher number of aDNA failures than the samples from sites with lower T_{eff} (Saxony and El Portalón, T_{eff} 12.5 °C and 9.5 °C respectively). It is therefore likely that aDNA and overall protein survival in ancient bone are relatively independent of one another and that both are more dependent on the site MAT and T_{eff} of the site. The relationship between aDNA and protein survival is not well understood and remains unclarified with this study; this may therefore limit the use of proteomics as a screening method for the presence of aDNA.

Of the proteins identified here, there are no proteins which are present solely in samples containing aDNA or vice versa, and this was essentially due to the presence of two samples which failed aDNA extraction but which had a good number of NCPs (17 and 18). Our previous suggestion that fetuin-A could be a promising source of phylogenetic information in archaeological bone and tooth samples [32] due to its abundance, survival and high level of sequence variability between taxa [48] was supported here; fetuin-A was identified in 58 of the 69 samples tested here, including some of the most poorly preserved samples. Furthermore, we also observed a 2-fold higher relative abundance of fetuin-A in bone and tooth samples successful for aDNA extraction than the second most abundant NCP considered in the study (prothrombin). It is therefore possible that such relative abundances of selected proteins could also be a potential indicator of aDNA extraction success. It should be noted that OC peptides were observed in most samples but excluded from this report (Table S1) because of our stringent requirement for three unique peptides even though in the case of OC, as one of the smallest proteins only approximately 49 amino acids long this would equate to most of the protein.

The relative abundances of the seven selected proteins between samples, which do or do not contain aDNA, are not significantly different and therefore are unlikely to be useful in determining the likelihood of a successful aDNA extraction. Asx DL values also do not appear to correlate directly with whether a sample contains aDNA or not, nor with the number of NCPs a sample contains or the relative abundances of the selected NCPs, although there is some indication that threshold values could be indicative. It is therefore probable that Asx DL values are not wholly representative of the state of NCP or collagen preservation in a sample, but more likely to reflect the amount of soluble collagen retained by the sample, as suggested by Collins et al. [49] and Dobberstein et al. [50]. Amino acid racemisation values were generally relatively high (>0.08), and although samples with low DLs (<0.12) were highly likely to yield successful aDNA while samples with DL >0.15 did not, there was no correlation with proteome complexity. Interestingly, we were able to extract aDNA from samples with higher than 0.08 AAR values in contrast with previous findings [25].

A reliable screening method would ideally focus on a protein or a group of proteins that are always identified in the ancient bone proteome when the aDNA analysis is successful, and that are missing when the sample is failing aDNA analysis. Unfortunately we were not able to find any proteins matching these criteria, nor did we find any NCPs consistently identified in every sample; for this reason, we were not able to identify protein biomarkers useful to estimate the presence of aDNA in a sample. Although the search for a protein marker failed, we still noticed an interesting correlation between the number of recovered NCPs and the outcome of the aDNA extraction. In particular, the number of NCPs could be for a useful approach to evaluating the presence or absence of aDNA in the specimen where in this case samples with 18 or more NCPs were positive for aDNA analysis. However, additional work is
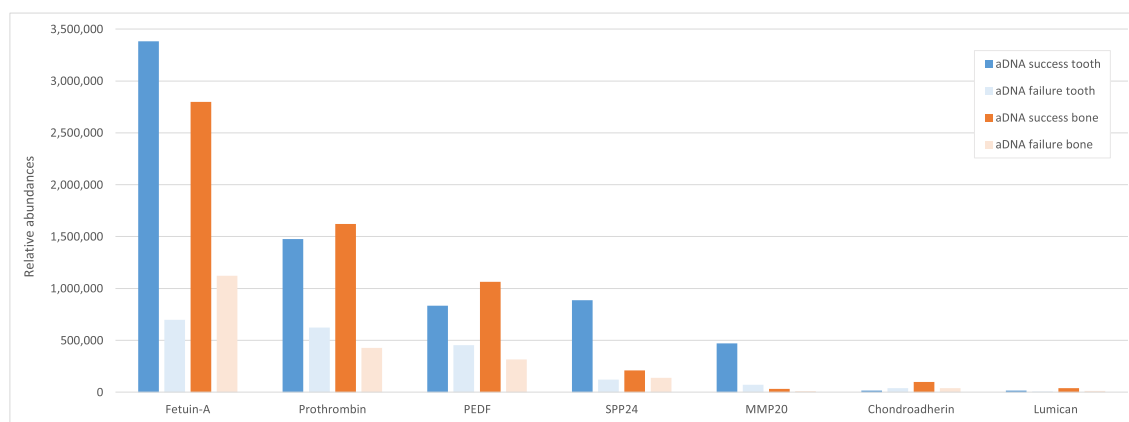

Fig. 4. Average relative abundances for seven commonly identified proteins in tooth and in bone samples which were successful for or failed aDNA analysis. Tooth samples have higher relative abundances of fetuin-A, SPP24 and MMP20 whereas bone samples have higher prothrombin, PEDF, chrondroadherin and lumican. By comparison, the relative abundance of COL1A2 was 1.13×10^7 and 1.42×10^7 for bone samples with and without aDNA respectively and 1.27×10^7 and 1.11×10^7 for dentine samples with and without aDNA respectively.

required to refine this property, which may be able to identify particular biomarker proteins, in order to be able to apply this methodology in the future.

4.1. Ancient DNA survival mechanisms

The lack of correlation between NCP and aDNA survival suggests that each biomolecule may rely on a different preservation mechanism. Many of the NCPs which are commonly identified in ancient bone are extracellular matrix (ECM) proteins, therefore likely to be of high abundance in bone matrix, or are known to associate with collagen which is highly abundant and resistant to decay [15,43]. However, DNA is known to adsorb to hydroxyapatite [51], where they are both thought to be preserved under the same conditions (a neutral or slightly alkaline pH; [52]). Hydroxyapatite-bound DNA is also more resistant to hydrolytic depurination [53] and may also be resistant to spontaneous decay as well as the enzymatic actions of DNase [54]. It therefore seems likely that DNA survival in ancient bone would be better screened for using methods that evaluate the state of preservation of hydroxyapatite crystals such as FTIR [55] or X-ray diffraction based approaches [56].

5. Conclusions

Proteomics can be a useful tool in the analysis of ancient bone and teeth, however this study has proved that this technique, in its current form, is unlikely to be useful as a screening method to determine whether a sample is likely to contain aDNA. Despite the identification of several proteins across the 69 samples analysed, there are none which fit the criteria to be a suitable biomarker for the presence or absence of aDNA. Additionally, the Asx DL values measured do not appear to correlate with the survival of NCPs in a sample or indicate which samples may be more likely to contain viable aDNA, but rather appear to reflect the state of collagen preservation [24].

However, even though we were not able to propose a precise protein (or group of proteins) as a suitable biomarker for the presence of aDNA in the ancient bone sample we did observe that samples yielding a high number of NCPs (>18) were always successful for the aDNA analysis. Therefore, the number of NCPs in a sample could potentially be used as a new way to evaluate the likely presence of aDNA in archaeological samples, although further studies of this type including a higher number of ancient samples would be needed to validate this preliminary finding. Although this study was unsuccessful in identifying potential biomarkers for aDNA, it does improve our understanding of aDNA survival mechanisms, potentially highlighting its stronger association to the inorganic phase than the organic phase of bone.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jprot.2017.01.004.

Transparency document

The Transparency document associated with this article can be found, in online version.

References

[1] G.D. Pérez, I. Jiménez, I. Rey, F.J.S. Sánchez-Cañete, F. Luque, A.M. Muñiz, M. Gálvez, J.S. Galdós, A. Sánchez, T.E.R. Tham, Genomics and proteomics in bioarchaeology: review, Archaeobios 4–17 (2013).

[2] T.A. Brown, K. Brown, Biomolecular Archaeology: An Introduction, John Wiley & Sons, 2011.

[3] P. Skoglund, H. Malmström, M. Raghavan, J. Storå, P. Hall, E. Willerslev, M.T.P. Gilbert, A. Götherström, M. Jakobsson, Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe, Science 336 (80) (2012) 466–469.

[4] R.E. Green, J. Krause, A.W. Briggs, T. Maricic, U. Stenzel, M. Kircher, N. Patterson, H. Li, W. Zhai, M.H.-Y. Fritz, A draft sequence of the Neandertal genome, Science 328 (80) (2010) 710–722.

[5] M. Meyer, M. Kircher, Illumina sequencing library preparation for highly multiplexed target capture and sequencing, Cold Spring Harb. Protoc. (2010) t5448.

[6] I. Mathieson, I. Lazaridis, N. Rohland, S. Mallick, N. Patterson, S.A. Roodenberg, E. Harney, K. Stewardson, D. Fernandes, M. Novak, Genome-wide patterns of selection in 230 ancient Eurasians, Nature 528 (2015) 499–503.

[7] J. Dabney, M. Meyer, Length and GC-biases during sequencing library amplification: a comparison of various polymerase-buffer systems with ancient and modern DNA sequencing libraries, BioTechniques 52 (2012) 87–94.

[8] L. Orlando, A. Ginolhac, G. Zhang, D. Froese, A. Albrechtsen, M. Stiller, M. Schubert, E. Cappellini, B. Petersen, I. Moltke, Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse, Nature 499 (2013) 74–78.

[9] M. Meyer, Q. Fu, A. Aximu-Petri, I. Glocke, B. Nickel, J.-L. Arsuaga, I. Martínez, A. Gracia, J.M.B. de Castro, E. Carbonell, A mitochondrial genome sequence of a hominin from Sima de los Huesos, Nature 505 (2014) 403–406.

[10] M. Meyer, J.-L. Arsuaga, C. de Filippo, S. Nagel, A. Aximu-Petri, B. Nickel, I. Martínez, A. Gracia, J.M.B. de Castro, E. Carbonell, Nuclear DNA sequences from the Middle Pleistocene Sima de los Huesos hominins, Nature 531 (2016) 504–507.

[11] D. Haouchar, C. Pacioni, J. Haile, M.C. McDowell, A. Baynes, M.J. Phillips, J.J. Austin, L.C. Pope, M. Bunce, Ancient DNA reveals complexity in the evolutionary history and taxonomy of the endangered Australian brush-tailed bettongs (Bettongia: Marsupialia: Macropodidae: Potoroinae), Biodivers. Conserv. 1–21 (2016).

[12] M.-H.S. Sinding, J. Arneborg, G. Nyegaard, M.T.P. Gilbert, Ancient DNA unravels the truth behind the controversial GUS Greenlandic Norse fur samples: the bison was a horse, and the muskox and bears were goats, J. Archaeol. Sci. 53 (2015) 297–303.

[13] S. Haynes, J.B. Searle, A. Bretman, K.M. Dobney, Bone preservation and ancient DNA: the application of screening methods for predicting DNA survival, J. Archaeol. Sci. 29 (2002) 585–592.

[14] C.I. Smith, C.M. Nielsen-Marsh, M.M.E. Jans, M.J. Collins, Bone diagenesis in the European Holocene I: patterns and mechanisms, J. Archaeol. Sci. 34 (2007) 1485–1493.

[15] C.M. Nielsen-Marsh, C.I. Smith, M.M.E. Jans, A. Nord, H. Kars, M.J. Collins, Bone diagenesis in the European Holocene II: taphonomic and environmental considerations, J. Archaeol. Sci. 34 (2007) 1523–1531.

[16] M.M.E. Jans, C.M. Nielsen-Marsh, C.I. Smith, M.J. Collins, H. Kars, Characterisation of microbial attack on archaeological bone, J. Archaeol. Sci. 31 (2004) 87–95.

[17] G. Turner-Walker, C.M. Nielsen-Marsh, H. Syversen, H. Kars, M.J. Collins, Sub-micron spongiform porosity is the major ultra-structural alteration occurring in archaeological bone, Int. J. Osteoarchaeol. 12 (2002) 407–414.

[18] L.S. Bell, Palaeopathology and diagenesis: an SEM evaluation of structural changes using backscattered electron imaging, J. Archaeol. Sci. 17 (1990) 85–102.

[19] R.E.M. Hedges, Bone diagenesis: an overview of processes, Archaeometry 44 (2002) 319–328.

[20] P.R. Smith, M.T. Wilson, Detection of haemoglobin in human skeletal remains by ELISA, J. Archaeol. Sci. 17 (1990) 255–268.

[21] C. Cattaneo, K. Gelsthorpe, P. Phillips, R.J. Sokol, Differential survival of albumin in ancient bone, J. Archaeol. Sci. 22 (1995) 271–276.

[22] J.F. Bailey, M.B. Richards, V.A. Macaulay, I.B. Colson, I.T. James, D.G. Bradley, R.E.M. Hedges, B.C. Sykes, Ancient DNA suggests a recent expansion of European cattle from a diverse wild progenitor species, Proc. R. Soc. Lond. B Biol. Sci. 263 (1996) 1467–1473.

[23] A. Götherström, M.J. Collins, A. Angerbjörn, K. Lidén, Bone preservation and DNA amplification, Archaeometry 44 (2002) 395–404.

[24] M.J. Collins, E.R. Waite, A.C.T. Van Duin, Predicting protein decomposition: the case of aspartic–acid racemization kinetics, Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 354 (1999) 51–64.

[25] H.N. Poinar, M. Höss, J.L. Bada, S. Pääbo, Amino acid racemization and the preservation of ancient DNA, Science 272 (80) (1996) 864–866.

[26] J.L. Bada, X.S. Wang, H. Hamilton, Preservation of key biomolecules in the fossil record: current knowledge and future challenges, Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 354 (1999) 77–87.

[27] H.N. Poinar, B.A. Stankiewicz, Protein preservation and DNA retrieval from ancient tissues, Proc. Natl. Acad. Sci. 96 (1999) 8426–8431.

[28] M.J. Collins, A.M. Gernaey, C.M. Nielsen-Marsh, C. Vermeer, P. Westbroek, Osteocalcin in fossil bones: evidence of very slow rates of decomposition from laboratory studies, Geology 28 (2000) 1139–1142.

[29] G. Colombo, P. Fanti, C. Yao, H.H. Malluche, Isolation and complete amino acid sequence of osteocalcin from canine bone, J. Bone Miner. Res. 8 (1993) 733–743.

[30] M. Buckley, C. Anderung, K. Penkman, B.J. Raney, A. Götherström, J. Thomas-Oates, M.J. Collins, Comparing the survival of osteocalcin and mtDNA in archaeological bone from four European sites, J. Archaeol. Sci. 35 (2008) 1756–1764.

[31] E. Cappellini, L.J. Jensen, D. Szklarczyk, A. Ginolhac, R.A.R. da Fonseca, T.W. Stafford Jr., S.R. Holen, M.J. Collins, L. Orlando, E. Willerslev, Proteomic analysis of a

pleistocene mammoth femur reveals more than one hundred ancient bone proteins, J. Proteome Res. 11 (2011) 917–926.

[32] C. Wadsworth, M. Buckley, Proteome degradation in fossils: investigating the longevity of protein survival in ancient bone, Rapid Commun. Mass Spectrom. 28 (2014) 605–615.

[33] M. Buckley, A molecular phylogeny of Plesiorycteropus reassigns the extinct mammalian order "Bibymalagasia", PLoS One 8 (2013), e59614. .

[34] M. Buckley, Species identification of bovine, ovine and porcine type 1 collagen; comparing peptide mass fingerprinting and LC-based proteomics methods, Int. J. Mol. Sci. 17 (2016) 445.

[35] C. Anderung, Genetic Analyses of Bovid Remains and the Origin of Early European Cattle, Uppsala University, 2006.

[36] C. Anderung, A. Bouwman, P. Persson, J.M. Carretero, A.I. Ortega, R. Elburg, C. Smith, J.L. Arsuaga, H. Ellegren, A. Götherström, Prehistoric contacts over the Straits of Gibraltar indicated by genetic analysis of Iberian Bronze Age cattle, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 8431–8435.

[37] A. Sekiya, H. Okano-Kosugi, C.M. Yamazaki, T. Koide, Pigment epithelium-derived factor (PEDF) shares binding sites in collagen with heparin/heparan sulfate proteoglycans, J. Biol. Chem. 286 (2011) 26364–26374.

[38] H. Alimohamad, T. Habijanac, H. Larjava, L. Häkkinen, Colocalization of the collagen-binding proteoglycans decorin, biglycan, fibromodulin and lumican with different cells in human gingiva, J. Periodontal Res. 40 (2005) 73–86.

[39] B. Månsson, C. Wenglén, M. Mörgelin, T. Saxne, D. Heinegård, Association of chondroadherin with collagen type II, J. Biol. Chem. 276 (2001) 32883–32888.

[40] W. Jahnen-Dechent, A. Heiss, C. Schäfer, M. Ketteler, Fetuin-A regulation of calcified matrix metabolism, Circ. Res. 108 (2011) 1494–1509.

[41] P.A. Price, A.A. Otsuka, J.W. Poser, J. Kristaponis, N. Raman, Characterization of a gamma-carboxyglutamic acid-containing protein from bone, Proc. Natl. Acad. Sci. 73 (1976) 1447–1451.

[42] E.J. Brochmann, K. Behnam, S.S. Murray, Bone morphogenetic protein–2 activity is regulated by secreted phosphoprotein–24 kd, an extracellular pseudoreceptor, the gene for which maps to a region of the human genome important for bone quality, Metabolism 58 (2009) 644–650.

[43] C.I. Smith, A.T. Chamberlain, M.S. Riley, C. Stringer, M.J. Collins, The thermal history of human fossils and the likelihood of successful DNA amplification, J. Hum. Evol. 45 (2003) 203–217.

[44] Y. Lu, P. Papagerakis, Y. Yamakoshi, J.C.-C. Hu, J.D. Bartlett, J.P. Simmer, Functions of KLK4 and MMP-20 in dental enamel formation, Biol. Chem. 389 (2008) 695–700.

[45] M. Sulkala, M. Larmas, T. Sorsa, T. Salo, L. Tjäderhane, The localization of matrix metalloproteinase-20 (MMP-20, enamelysin) in mature human teeth, J. Dent. Res. 81 (2002) 603–607.

[46] C. Nielsen-Marsh, Biomolecules in fossil remains-multidisciplinary approach to endurance, Biochemistry (London) 24 (2002) 12–14.

[47] M. Pruvost, R. Schwarz, V.B. Correia, S. Champlot, S. Braguier, N. Morel, Y. Fernandez-Jalvo, T. Grange, E.-M. Geigl, Freshly excavated fossil bones are best for amplification of ancient DNA, Proc. Natl. Acad. Sci. 104 (2007) 739–744.

[48] M. Buckley, C. Wadsworth, Proteome degradation in ancient bone: diagenesis and phylogenetic potential, Palaeogeogr. Palaeoclimatol. Palaeoecol. 416 (2014) 69–79.

[49] M.J. Collins, K.E.H. Penkman, N. Rohland, B. Shapiro, R.C. Dobberstein, S. Ritz-Timme, M. Hofreiter, Is amino acid racemization a useful tool for screening for ancient DNA in bone? Proc. R. Soc. Lond. B Biol. Sci. (2009) (rspb20090563).

[50] R.C. Dobberstein, M.J. Collins, O.E. Craig, G. Taylor, K.E.H. Penkman, S. Ritz-Timme, Archaeological collagen: why worry about collagen diagenesis? Archaeol. Anthropol. Sci. 1 (2009) 31–42.

[51] G. Bernardi, Chromatography of nucleic acids on hydroxyapatite, Nature 206 (1965) 779–783.

[52] J. Burger, S. Hummel, B. Herrmann, W. Henke, DNA preservation: a microsatellite-DNA study on ancient skeletal remains, Electrophoresis 20 (1999) 1722–1728.

[53] T. Lindahl, Instability and decay of the primary structure of DNA, Nature 362 (1993) 709–715.

[54] M. Brundin, D. Figdor, G. Sundqvist, U. Sjögren, DNA binding to hydroxyapatite: a potential mechanism for preservation of microbial DNA, J. Endod. 39 (2013) 211–216.

[55] C. Schwarz, R. Debruyne, M. Kuch, E. McNally, H. Schwarcz, A.D. Aubrey, J. Bada, H. Poinar, New insights from old bones: DNA preservation and degradation in permafrost preserved mammoth remains, Nucleic Acids Res. (2009) gkp159.

[56] A. Bartsiokas, A.P. Middleton, Characterization and dating of recent and fossil bone by X-ray diffraction, J. Archaeol. Sci. 19 (1) (1992) 63–72.