PROCEEDINGS B

rspb.royalsocietypublishing.org

Research

Cite this article: Cleland TP, Schroeter ER, Schweitzer MH. 2015 Biologically and diagenetically derived peptide modifications in moa collagens. *Proc. R. Soc. B* **282**: 20150015. http://dx.doi.org/10.1098/rspb.2015.0015

Received: 4 January 2015
Accepted: 20 April 2015

Subject Areas:
biochemistry, palaeontology

Keywords:
moa, collagen, diagenesis, post-translational modifications

Author for correspondence:
Timothy P. Cleland
e-mail: clelat@rpi.edu

Electronic supplementary material is available at http://dx.doi.org/10.1098/rspb.2015.0015 or via http://rspb.royalsocietypublishing.org.

THE ROYAL SOCIETY PUBLISHING

Biologically and diagenetically derived peptide modifications in moa collagens

Timothy P. Cleland[1], Elena R. Schroeter[2] and Mary H. Schweitzer[2,3]

[1]Department of Biomedical Engineering, Rensselaer Polytechnic Institute, Troy, NY 12182, USA
[2]Department of Biological Sciences, North Carolina State University, Raleigh, NC 27695, USA
[3]North Carolina Museum of Natural Sciences, Raleigh, NC 27601, USA

The modifications that occur on proteins in natural environments over time are not well studied, yet characterizing them is vital to correctly interpret sequence data recovered from fossils. The recently extinct moa (Dinornithidae) is an excellent candidate for investigating the preservation of proteins, their post-translational modifications (PTMs) and diagenetic alterations during degradation. Moa protein extracts were analysed using mass spectrometry, and peptides from collagen I, collagen II and collagen V were identified. We also identified biologically derived PTMs (i.e. methylation, di-methylation, alkylation, hydroxylation, fucosylation) on amino acids at locations consistent with extant proteins. In addition to these *in vivo* modifications, we detected novel modifications that are probably diagenetically derived. These include loss of hydroxylation/glutamic semialdehyde, carboxymethyllysine and peptide backbone cleavage, as well as previously noted deamidation. Moa collagen sequences and modifications provide a baseline by which to evaluate proteomic studies of other fossils, and a framework for defining the molecular relationship of moa to other closely related taxa.

1. Introduction

The first investigations into biomolecular preservation in fossils focused on proteins [1–5]; however, polymerase chain reaction (PCR) and other technological advances [6–8] resulted in genomic studies superseding those of proteins. Now, recent advances in high-resolution mass spectrometry have resulted in a renewed interest in the utility of proteins preserved in fossils (reviewed in [9]). Ancient proteomic studies have extended the age for which biomolecules and phylogenetically informative molecular sequences can be recovered to well beyond the hypothesized limit for DNA preservation [8,10–17].

Proteomic studies on sequences from fossils directly characterize biologically and/or diagenetically derived post-translational modifications (PTMs). Because biologically derived PTMs originate from the organism itself, they inform on the ultimate protein function, phylogenetic changes or evolutionary adaptations at the molecular level that cannot be determined from DNA sequences alone [9]. In contrast to phylogenetic or physiological results of *in vivo* PTMs, diagenetically derived PTMs are the result of post-mortem decay. These modifications provide evidence that detected proteins are original to the fossil, and inform how proteins/amino acids degrade over geological time [11]. Historically, few biologically derived PTMs have been identified for extinct species (e.g. hydroxylation of proline, HYP; carboxylation of glutamic acid [16]) and even fewer diagenetically derived PTMs have been identified [11,13,14]. Although deamidation is the most common diagenetically derived PTM directly identified on peptides, non-enzymatic glycation has been hypothesized to be a major factor affecting preservation of ancient proteins because it results in highly cross-linked protein structures reducing protein solubility [17]. However, it has not been directly measured or detected in fossil taxa. Because few PTMs from ancient remains have been elucidated, information regarding which PTMs persist into the rock record and what types of modifications can occur diagenetically is limited.

Here, we investigate moa bone (MOR OST-255) for the preservation of protein and PTMs. Moa remains are of special interest because in addition to

bone, diverse elements (i.e. feathers, bones, mummified remains, egg shells) are attributed to them [18–20]. Because of the availability of exceptionally preserved specimens, DNA recovered from various moa specimens has been used to differentiate species [20,21] and estimate rates of DNA degradation [22]. However, few complementary studies of moa proteins have been attempted [3,23].

2. Material and methods

(a) Moa bone

Cortical bone fragments from an indeterminate moa specimen (MOR OST-255; 800–1000 years old [24]) were extracted as described by Cleland *et al.* [23]. Briefly, approximately 1 g of cortical bone was demineralized in 9 ml 0.6 M HCl for 4 h at room temperature; the pellet was washed with sterile, deionized water and fragments were further extracted using 50 mM ammonium bicarbonate at 65°C for 5 h. The HCl fraction was dialyzed against water (2000 MWCO Slide-A-Lyzer Cassettes; Pierce) for 4 days at 4°C, then lyophilized to completion. The ammonium bicarbonate fraction was dried completely without further dialysis using a speed vacuum. Resultant powders were stored at −80°C until use.

(b) Protein digestion and mass spectrometry

Each powder (1 mg for HCl, 1.5 mg for ammonium bicarbonate) was resuspended in 500 µl of 50 mM ammonium bicarbonate. 50 ul of each fraction were reduced using 10 mM dithiothreitol for 1 h at 37°C followed by alkylation with 20 mM iodoacetamide for 1 h in the dark at room temperature. After reduction and alkylation, the proteins were digested overnight at 37°C with 0.2 µg Promega modified trypsin. Digestion was subsequently stopped with 1 ul of 100% formic acid and stored at −80°C until mass spectrometry.

Without further sample processing, 10 ul of each extraction were injected onto a Waters nanoAcquity UPLC trap column (180 µm × 20 mm) with Symmetry C18 and washed for 5 min at 5 µl min^{-1}. Peptides were transferred to a Waters nanoAcquity UPLC (75 µm × 250 mm) BEH130C18 (1.7 µm particle size) analytical column and eluted at 300 nl min^{-1} on a Waters nanoAcquity with the following gradient: 2% B (99.9% acetonitrile, 0.1% formic acid) to 60% B at 30 min, 90% B at 32 min, 90% B at 35 min, 2% B at 37 min, 2% B at 60 min. Buffer A was 0.1% formic acid. Eluted peptides were analysed on a ThermoScientific Orbitrap XL with a scan range of 375–2000 m/z. The top five peaks for each precursor were fragmented using collision-induced dissociation, and dynamic exclusion was enabled with a repeat count of 1, exclusion duration of 10 s and repeat duration of 30 s.

(c) Data analysis

The resulting spectra were analysed using three different search engines: Mascot 2.3 [25] in Proteome Discoverer 1.2 (ThermoScientific), Sequest HT [26] in Proteome Discoverer 1.4 and PEAKS7 [27,28]. Each Mascot file was sequentially searched against several databases, namely Uniprot chicken, the common repository of adventitious proteins (The Global Proteome Machine), Uniprot osteocalcin, Uniprot bone, Uniprot collagen and Uniprot haemoglobin; all results were compiled and overall peptide and protein statistics calculated. The following parameters were used for searching on all databases unless noted in brackets: 10 ppm precursor tolerance; 0.5 Da fragment tolerance; static modification: carbamidomethyl cysteine (C); dynamic modifications: deamidated asparagine and glutamine (NQ), oxidation methionine (M), oxidation arginine and lysine (KP) for Uniprot chicken, Uniprot bone and Uniprot collagen, carboxy glutamic acid (E) for Uniprot osteocalcin. All peptides were filtered with a 5% FDR based on a decoy database.

Spectra were searched using Sequest HT against a Uniprot Archosauria + Testudinidae database and a Uniprot collagen database with the following parameters: 10 ppm precursor tolerance, 0.5 Da fragment tolerance, fixed modifications: none, variable modifications: carbamidomethyl (C), deamidated (NQ), oxidation (M), carboxymethyl (K). For the collagen database, oxidation (KP) was also added to variable modifications. This mass shift represents HYP and lysine but is limited only to collagen sequences. All peptides were filtered with a 5% FDR based on a decoy database.

Spectra were searched using PEAKS7 against a Uniprot Vertebrates database using: 10 ppm precursor tolerance, 0.5 Da fragment tolerance, fixed modifications: none, variable modifications: carbamidomethylation (C), deamidated (NQ), oxidation (M), oxidation or hydroxylation (RYFPNKD) or (G) at C-terminal, and Carboxymethyl (KW) or (X) at N-terminal. A maximum of five PTMs were allowed per peptide. Non-specific cleavage was allowed at both ends of the peptide, as well as a maximum of three missed cleavages. To find additional, unspecified PTMs and mutations, PEAKS PTM [28] and SPIDER searches were enabled. Results were filtered with the following parameters: peptides $-10\,\mathrm{lgP} \geq 15$ (FDR 0.6%) and proteins $-10\,\mathrm{lgP} \geq 20$ (FDR 0.0%).

For all searches, peptides were exported and those with overlapping sequences were culled. All peptides for all identified collagens were aligned against collagen I and II exemplars from Uniprot in Seaview 4. Consensus sequences were generated for each search algorithm using a basic majority rule method in Seaview.

3. Results and discussion

We report the first partial collagen I sequence (figures 1 and 2; electronic supplementary material, tables S1, S2, S5–S7), the first partial collagen II sequence (electronic supplementary material, figure S1 and tables S3, S5–S7; for coverage see electronic supplementary material, table S8) and several peptides from collagen V (electronic supplementary material, tables S4–S7; for coverage see electronic supplementary material, table S8) for moa. Because the collagen V sequences exhibited limited coverage, we did not align them; however, future analyses from this specimen and others will provide additional sequence information for each of the collagen V chains. Database searching using three algorithms resulted in some variation in sequence coverage in collagen Iα1 and α2 (figures 1 and 2) when compared with the mature sequence of chicken (Col1a1: PEAKS 73.7%, Mascot 77.8%, Sequest 70.9%, all combined 84.1%; Col1a2: PEAKS 63.3%, Mascot 47.0%, Sequest 50.7%, all combined 69.3%). This multi-algorithm approach facilitates identification of more complete sequences of other fossil taxa that may be missed when only one algorithm is applied.

This is an indeterminate specimen (i.e. species was unable to be determined, so only identified to a supraspecific level), but the collagen I sequences, and to a lesser extent the collagen II sequence, represent an important baseline for expected collagen sequences in closely related extinct and extant species, and provide critical data applicable to other undersampled palaeognath taxa. Additionally, these sequences can be used to refine phylogenetic hypotheses of other archosaurs (e.g. electronic supplementary material, figure S7), including extant and extinct crocodylians and extinct non-avian dinosaurs. The phylogeny resulting from collagen I sequences places moa in a clade with other palaeognath taxa as well as Galloansarae taxa (electronic supplementary material, figure S7). Unfortunately,

Figure 1. Collagen Iα1 sequence alignments for PEAKS, Mascot and Sequest peptides.

Figure 2. Collagen Iα2 sequence alignments for PEAKS, Mascot and Sequest peptides.

few collagen I sequences are known from palaeognath taxa, requiring additional sampling of bone from moa and other extant palaeognaths (e.g. emu, kiwi) to better elucidate relationships and lend support to hypothesized DNA-based phylogenies [29].

(a) *In vivo* post-translational modifications

We detected various biologically derived PTMs: methylation (figure 3*a*), di-methylation (electronic supplementary material, figure S2), alkylation (figure 3*a*; electronic supplementary material, figure S3), fucosylation (figure 3*b*; electronic supplementary material, figure S5) and hydroxylation; listed in table 2 and electronic supplementary material, tables S1–S4. With the exception of hydroxylated proline, few other *in vivo* PTMs have been identified from fossil remains. PEAKS PTM detected more PTMs than have been previously observed on ancient peptides without *a priori* knowledge of what may or may not preserve. This search strategy allowed us to detect enzymatic glycosylation for the first time in ancient bone proteins (figure 3; electronic supplementary material, figure S5).

Determining the endogeneity of fossil proteins and biological PTMs is critical for evaluating their biological function. To support PTM endogeneity, we compared the positions of fossil PTMs with those on proteins from extant eukaryotic taxa. We detected well-known methylation not only on lysine [30] but also on aspartic acid and glutamic acid (figure 3*a*; electronic supplementary material, tables S1–S3), only recently shown to occur in eukaryotic proteins [30]. These modifications support endogeneity. Additionally, we identified two peptides containing fucose on serine residues. Serine is one of several residues that is fucosylated in extant taxa [31]. Lastly, we identified one of the most common and potentially important positions of acetylation in extant proteins [32] on the moa lysine residues (figure 3*a*; electronic supplementary material, figure S3).

(b) Advanced-glycation end products and diagenetically modified peptides

In addition to *in vivo* PTMs, we were able to detect diagenetically derived protein modifications. Consistent with previous

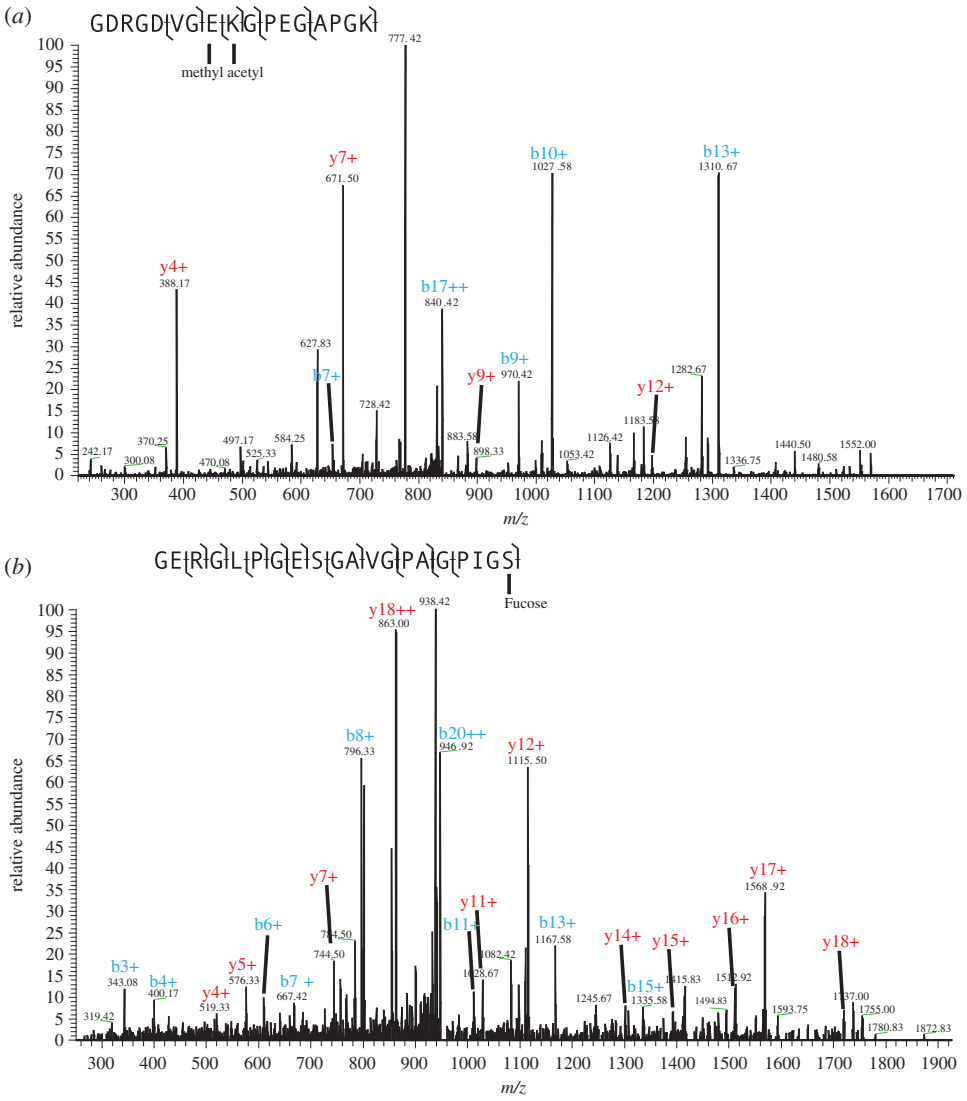

Figure 3. (*a*) Collagen IIα1 peptide (GDRGDVGEKGPEGAPGK) showing methylation and alkylation. (*b*) Collagen IIα1 peptide (GERGLPGESGAVGPAGPIGS) showing fucosylation. (Online version in colour.)

analyses of ancient bone [8,14,15], we observe deamidation for both glutamine and asparagine (figure 4; electronic supplementary material, figure S4) that is incomplete (e.g. electronic supplementary material, figure S4) and may be, in part, derived from sample preparation. We observed additional modifications we ascribe to diagenesis, including advanced glycation end-products (AGEs), backbone cleavage and variable HYP. It has been hypothesized that AGEs lead to preservation of proteins [17]; however, direct observation of AGEs on intact peptides has not been previously made from any ancient source. We observe carboxymethyllysine (CML) on several peptides, which leads to missed cleavages by trypsin (figure 4), and a potential backbone cleavage where CML is present at the C-terminus of the peptide (electronic supplementary material, figure S6). CML modification in the C-terminal position may not lead to enhanced preservation because it does not form cross-links [33] but instead modifies the side chain of lysine, resulting in the potential disruption of collagen tertiary structure and subsequent loss of fossil collagen. This may explain the selectivity of peptide preservation noted by San Antonio *et al.* [34]. Additionally, we observe heterogeneity in the presence or absence of CML on the same peptide (electronic supplementary material, table S1). Further research is necessary to detect and measure the types, amounts and effects of AGEs in ancient and fossil bones.

Cleavage of the protein backbone has been suggested as one of the major causes for loss of protein from bone [35]. Because trypsin is specific for digestion at arginine and lysine residues, we hypothesize that peptides that do not terminate with arginine, lysine or the end of the protein sequence represent backbone cleavages. We observe several peptides that may represent backbone cleavage on the collagen Iα1 chain (table 1). Two of the peptides (i.e. TGPPGPAGQDGR|G|PPGPPGAR and VGPPGPSGNIGL|PGPPGPAGK, where | represents potential backbone cleavage positions) show breakages at proline residues consistent with backbone cleavage by oxidation [36].

Finally, we hypothesize that variable HYP represents diagenetic change. In extant collagen, there is little variation in HYP percentage [37], but 49% (35 of 71) of moa collagen Iα1 peptides that contain HYP show variability (i.e. peptides with the same sequence have different numbers of HYP independent of hydroxylation position). On collagen Iα2 peptides, we observe 30% variability (12 of 39), and on collagen IIα1 peptides, we observe 56% variability (5 of 9 peptides). These variations are much higher than would be expected from a sample preparation artefact because HYP has been shown to persist in completely hydrolyzed samples and used to approximate collagen content [38]. Alternatively, this variation could represent oxidation of proline to glutamic

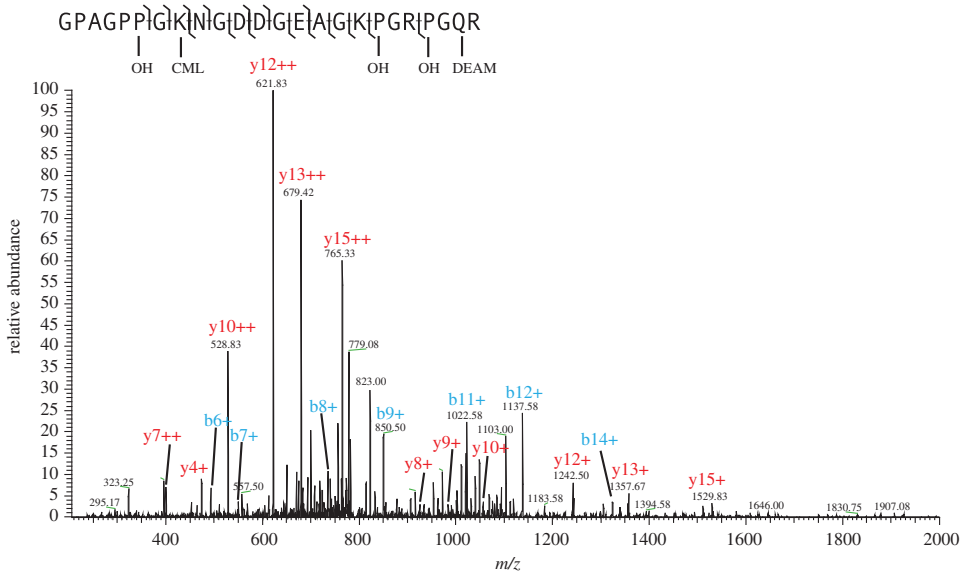

Figure 4. Collagen Iα1 peptide (GPAGPPGKNGDDGEAGKPGRPGQR) showing hydroxylation, carboxymethyllysine and deamidation. (Online version in colour.)

Table 1. Examples of hypothesized protein backbone cleavage detected for collagen I alpha 1 peptides.

EGAPGAEGAPGR	TGPPGPAGQDGR
EGAPGAEGAPGRD	TGPPGPAGQDGRPG
EGAPGAEGAPGRDG	TGPPGPAGQDGRPGPPGPPGAR
EGAPGAEGAPGRDGAAGPK	GAPGDRGEPGPPGPAGFAGPPGADGQPGAK
EGAPGAEGAPGRDGAAGPKGDR	GAPGDRGEPGPPGPAGFAGPPGADGQPGAKG
	VGPPGPSGNIGL
	VGPPGPSGNIGLPGPPGPAGK

Table 2. Protein and peptide modifications detected.

biologically derived	diagenetically derived
alkylation	backbone cleavage
dimethylation	carboxymethylation (advanced glycation end-product)
fucosylation	deamidation[a]
hydroxylation[a]	dehydroxylation/glutamic semialdehyde
methylation	

[a]Previously detected.

semialdehyde (i.e. an isobaric mass shift as hydroxyproline [11]). In either case, we observe large variations in modified proline residues that may be the result of diagenesis. The loss of hydroxylation may also explain the detection of serine/alanine substitutions in multiple peptides (i.e. loss of oxygen from the side chain of serine results in a transition of that residue to alanine) detected by database searching in multiple peptides (e.g. GP[S/A]GPPGKNGDDGEAGKPGR PG[Q/E]R). This sample peptide also shows a glutamine/glutamic acid residue difference that may only reflect deamidation but not an amino acid substitution. Alternatively, the serine/alanine transition is potentially an effect of algorithm differences in detection that need to be further resolved for all palaeoproteomic studies. These diagenetic changes to amino acid side groups, consistent with previously observed modifications to ancient DNA (e.g. deamination of cytosine [39]), may lead to incorrect assignments in phylogenetic analyses, and need to be considered when protein sequences are used to evaluate evolutionary relationships.

4. Conclusion

Protein sequences obtained from this moa provide a baseline against which peptides recovered from other fossils may be searched, and identify modifications that may occur in other fossils, allowing differentiation between inherent genetic change and diagenetic change of the original proteins. While others have reported several biologically and diagenetically derived PTMs [11,13], we identified four biological PTMs (table 2) out of five total modifications for the first time from fossil remains. We also detected three or potentially four novel diagenetic PTMs (table 2) that have not been previously detected from fossils. Delineating both *in vivo* and diagenetically derived PTMs in these extinct taxa will provide more robust hypotheses regarding the physiology [40] and/or phylogenies of these organisms, as well as the mechanisms leading to preservation or loss of proteins from bones of different ages or localities.

Data accessibility. All raw mass spectrometry data are available at Dryad (http://dx.doi.org/10.5061/dryad.q35s1).

Acknowledgements. We thank J. Horner and J. Scannella for access to and information on this moa specimen, J. Carlson and S. Baxter for access to the LTQ Orbitrap XL at David H. R. Murdoch Research Institute, and N. Kelleher and P. Thomas for access to Peaks7. We also thank the Willi Hennig Society for providing access to TNT, and three anonymous reviewers for helpful critiques for improving this manuscript.

Funding statement. This research was funded by NSF EAR 0541744 to M.H.S., NSF DGE-0750733 to T.P.C., the David and Lucile Packard Foundation to M.H.S. and NSF INSPIRE to M.H.S. and E.R.S.

Authors' contributions. All authors conceived and designed this study. T.P.C. acquired the data, T.P.C. and E.R.S. performed bioinformatics and all authors made interpretations. T.P.C. and E.R.S. wrote the manuscript and all authors revised it leading to the final version.

References

1. Lowenstein JM. 1980 Species-specific proteins in fossils. *Naturwissenschaften* **67**, 343–346. (doi:10.1007/bf01106588)
2. Rainey WE, Lowenstein JM, Sarich VM, Magor DM. 1984 Sirenian molecular systematics—including the extinct Steller's sea cow (*Hydrodamalis gigas*). *Naturwissenschaften* **71**, 586–588. (doi:10.1007/bf01189187)
3. Huq NL, Rambaud SM, Teh L-C, Davies AD, McCulloch B, Trotter MM, Chapman GE. 1985 Immunochemical detection and characterisation of osteocalcin from moa bone. *Biochem. Biophys. Res. Commun.* **129**, 714–720. (doi:10.1016/0006-291x(85)91950-3)
4. Wyckoff RWG, McCaughey WF, Doberenz AR. 1964 The amino acid composition of proteins from pleistocene bones. *Biochim. Biophys. Acta* **93**, 374–377. (doi:10.1016/0304-4165(64)90387-3)
5. Bada JL, Kvenvolden KA, Peterson E. 1973 Racemization of amino acids in bones. *Nature* **245**, 308–310. (doi:10.1038/245308a0)
6. Higuchi R, Bowman B, Freiberger M, Ryder OA, Wilson AC. 1984 DNA sequences from the quagga, an extinct member of the horse family. *Nature* **312**, 282–284. (doi:10.1038/312282a0)
7. Green RE *et al*. 2010 A draft sequence of the neandertal genome. *Science* **328**, 710–722. (doi:10.1126/science.1188021)
8. Orlando L *et al*. 2013 Recalibrating *Equus* evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* **499**, 74–78. (doi:10.1038/nature12323)
9. Cappellini E, Collins MJ, Gilbert MTP. 2014 Unlocking ancient protein palimpsests. *Science* **343**, 1320–1322. (doi:10.1126/science.1249274)
10. Schweitzer MH *et al*. 2009 Biomolecular characterization and protein sequences of the Campanian hadrosaur *B*. canadensis. *Science* **324**, 626–631. (doi:10.1126/science.1165069)
11. Cappellini E *et al*. 2012 Proteomic analysis of a pleistocene mammoth femur reveals more than one hundred ancient bone proteins. *J. Proteome Res.* **11**, 917–926. (doi:10.1021/pr200721u)
12. Buckley M. 2013 A molecular phylogeny of *Plesiorycteropus* reassigns the extinct mammalian order 'Bibymalagasia'. *PLoS ONE* **8**, e59614. (doi:10.1371/journal.pone.0059614)
13. Wadsworth C, Buckley M. 2014 Proteome degradation in fossils: investigating the longevity of protein survival in ancient bone. *Rapid Commun. Mass Spectrom.* **28**, 605–615. (doi:10.1002/rcm.6821)
14. van Doorn NL, Wilson J, Hollund H, Soressi M, Collins MJ. 2012 Site-specific deamidation of glutamine: a new marker of bone collagen deterioration. *Rapid Commun. Mass Spectrom.* **26**, 2319–2327. (doi:10.1002/rcm.6351)
15. Wilson J, van Doorn NL, Collins MJ. 2012 Assessing the extent of bone degradation using glutamine deamidation in collagen. *Anal. Chem.* **84**, 9041–9048. (doi:10.1021/ac301333t)
16. Nielsen-Marsh CM, Ostrom PH, Gandhi H, Shapiro B, Cooper A, Hauschka PV, Collins MJ. 2002 Sequence preservation of osteocalcin protein and mitochondrial DNA in bison bones older than 55 ka. *Geology* **30**, 1099–1102. (doi:10.1130/0091-7613(2002)030<1099:SPOOPA>2.0.CO;2)
17. Nielsen-Marsh CM, Richards MP, Hauschka PV, Thomas-Oates JE, Trinkaus E, Pettitt PB, Karavanić I, Poinar H, Collins MJ. 2005 Osteocalcin protein sequences of Neanderthals and modern primates. *Proc. Natl Acad. Sci. USA* **102**, 4409–4413. (doi:10.1073/pnas.0500450102)
18. Rawlence NJ, Wood JR, Armstrong KN, Cooper A. 2009 DNA content and distribution in ancient feathers and potential to reconstruct the plumage of extinct avian taxa. *Proc. R. Soc. B* **276**, 3395–3402. (doi:10.1098/rspb.2009.0755)
19. Oskam CL *et al*. 2010 Fossil avian eggshell preserves ancient DNA. *Proc. R. Soc. B* **277**, 1991–2000. (doi:10.1098/rspb.2009.2019)
20. Allentoft ME, Rawlence NJ. 2012 Moa's ark or volant ghosts of Gondwana? Insights from nineteen years of ancient DNA research on the extinct moa (Aves: Dinornithiformes) of New Zealand. *Ann. Anat.* **194**, 36–51. (doi:10.1016/j.aanat.2011.04.002)
21. Baker AJ, Huynen LJ, Haddrath O, Millar CD, Lambert DM. 2005 Reconstructing the tempo and mode of evolution in an extinct clade of birds with ancient DNA: the giant moas of New Zealand. *Proc. Natl Acad. Sci. USA* **102**, 8257–8262. (doi:10.1073/pnas.0409435102)
22. Allentoft ME *et al*. 2012 The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proc. R. Soc. B* **279**, 4724–4733. (doi:10.1098/rspb.2012.1745)
23. Cleland TP, Voegele K, Schweitzer MH. 2012 Empirical evaluation of bone extraction protocols. *PLoS ONE* **7**, e31443. (doi:10.1371/journal.pone.0031443)
24. Schweitzer MH, Wittmeyer JL, Horner JR. 2007 Soft tissue and cellular preservation in vertebrate skeletal elements from the Cretaceous to the present. *Proc. R. Soc. B* **274**, 183–197. (doi:10.1098/rspb.2006.3705)
25. Perkins DN, Pappin DJC, Creasy DM, Cottrell JS. 1999 Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567. (doi:10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2)
26. Eng JK, McCormack AL, Yates III JR. 1994 An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5**, 976–989. (doi:10.1016/1044-0305(94)80016-2)
27. Ma B, Zhang K, Hendrie C, Liang C, Li M, Doherty-Kirby A, Lajoie G. 2003 PEAKS: powerful software for peptide *de novo* sequencing by tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **17**, 2337–2342. (doi:10.1002/rcm.1196)
28. Han X, He L, Xin L, Shan B, Ma B. 2011 PeaksPTM: mass spectrometry-based identification of peptides with unspecified modifications. *J. Proteome Res.* **10**, 2930–2936. (doi:10.1021/pr200153k)
29. Mitchell KJ, Llamas B, Soubrier J, Rawlence NJ, Worthy TH, Wood J, Lee MSY, Cooper A. 2014 Ancient DNA reveals elephant birds and kiwi are sister taxa and clarifies ratite bird evolution. *Science* **344**, 898–900. (doi:10.1126/science.1251981)
30. Sprung R, Chen Y, Zhang K, Cheng D, Zhang T, Peng J, Zhao Y. 2008 Identification and validation of eukaryotic aspartate and glutamate methylation in proteins. *J. Proteome Res.* **7**, 1001–1006. (doi:10.1021/pr0705338)
31. Sharon N, Lis H. 2008 Glycoproteins: structure and function. In *Glycosciences: status and perspectives* (eds HJ Gabius, S. Gabius), pp. 133–162. New York, NY: Wiley-VCH Verlag GmbH.
32. Glozak MA, Sengupta N, Zhang X, Seto E. 2005 Acetylation and deacetylation of non-histone proteins. *Gene* **363**, 15–23. (doi:10.1016/j.gene.2005.09.010)
33. Shapiro BP, Owan TE, Mohammed SF, Meyer DM, Mills LD, Schalkwijk CG, Redfield MM. 2008 Advanced glycation end-products accumulate in vascular smooth muscle and modify vascular but not ventricular properties in elderly hypertensive canines. *Circulation* **118**, 1002–1010. (doi:10.1161/CIRCULATIONAHA.108.777326)
34. San Antonio JD, Schweitzer MH, Jensen ST, Kalluri R, Buckley M, Orgel JPRO. 2011 Dinosaur peptides suggest mechanisms of protein survival. *PLoS ONE* **6**, e20381. (doi:10.1371/journal.pone.0020381)
35. Schweitzer MH. 2004 Molecular paleontology: some current advances and problems. *Ann. Paléontol.* **90**, 81–102. (doi:10.1016/j.annpal.2004.02.001)
36. Berlett BS, Stadtman ER. 1997 Protein oxidation in aging, disease, and oxidative stress. *J. Biol. Chem.* **272**, 20 313–20 316. (doi:10.1074/jbc.272.33.20313)
37. Barnes MJ, Constable BJ, Morton LF, Royce PM. 1974 Age-related variations in hydroxylation of lysine and proline in collagen. *Biochem. J.* **139**, 461–468.
38. Sroga GE, Karim L, Colón W, Vashishth D. 2011 Biochemical characterization of major bone-matrix proteins using nanoscale-size bone samples and proteomics methodology. *Mol. Cell. Proteomics* **10**, M110.006718. (doi:10.1074/mcp.M110.006718)
39. Hofreiter M, Jaenicke V, Serre D, Haeseler AV, Pääbo S. 2001 DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. *Nucleic Acids Res.* **29**, 4793–4799. (doi:10.1093/nar/29.23.4793)
40. Seo J, Lee KJ. 2004 Post-translational modifications and their biological functions: proteomic analysis and systematic approaches. *J. Biochem. Mol. Biol.* **37**, 35–44. (doi:10.5483/BMBRep.2004.37.1.035)