Collagen fingerprinting for the species identification of archaeological amphibian remains

MICHAEL BUCKLEY and MARC CHEYLAN

Amphibians are well known as being one of the main groups of animals today most threatened by environmental changes but they are also some of the least well understood of all terrestrial vertebrates. This gap in knowledge is much greater as we look further back into the relatively recent past, despite representing an invaluable resource in archaeological and palaeontological assemblages that are more indicative of palaeoclimate conditions than most other vertebrate taxa. This in part stems from their remains being typically much less studied, partly due to the less common forms of expertise required for identifications based on skeletal morphology – the most frequently observed tissue that remains in ancient assemblages. Here we apply a method of biomolecular species identification by collagen peptide mass fingerprinting to the British Late Pleistocene assemblage of Pin Hole Cave (Creswell Crags, UK) as well as a range of relevant extant taxa for comparison. Our results demonstrate the ability to separate at the species level with all modern taxa investigated, allowing for the identification of these archaeological remains to the amphibian taxa known to exist in Late Pleistocene Britain. Analyses of the Pin Hole assemblage found a dominance of the two species previously known from the site (common frogs and toads: *Rana temporaria* and *Bufo bufo*, respectively) and also a small number of the rarer natterjack toad (*Epidaulea calamita*) not previously identified in the Creswell Crags region but known from other sites in the UK; additionally, one specimen appeared to yield the fingerprint of the moor frog (*R. arvalis*), now extinct in the UK. As such, collagen fingerprinting is here shown to widen the known palaeobiodiversity of taxa, and highlights the further potential to enhance our understanding of climate change in the past.

For over a century, archaeologists and palaeontologists have studied animal bones for their potential in recovering information relating to past environments (e.g. Klein & Cruz-UrIBE 1984; Fernández-Jalvo et al. 2011). Of all terrestrial vertebrates, amphibians have been considered one of the most useful taxonomic groups for studying palaeoenvironments due to their greater sensitivity to subtle changes in climate (Blaustein et al. 2010; While & Uller 2014) to the extent that they have been used to describe changing environments linked to the decline of Neanderthal populations in the Mediterranean during the Late Pleistocene (Blain et al. 2008). Their sensitivity to climate, similar to that of reptiles, relates to the fact that they are ectothermic, yet they remain typically much less well studied than most other vertebrate groups, particularly mammals and fish. The extant amphibians form the Lissamphibia, composed of three main groups, the Salientia (frogs and toads) making up the largest group, the Caudata (salamanders and newts) and the Gymnophiona (limbless caecilians). However, their dependency on the weather impacts upon their activity, particularly in cooler parts of the world where they hibernate through winters and become completely inactive at night (Jaeger & Hailman 1981; Hazel 1989; Storey 1990). As a result it is not surprising that the majority of the more than 6300 known species of amphibians live in the tropics, with very few species existing in northern Europe today (Duellman 1999).

In Britain, extant amphibian taxa include the common frog (*Rana temporaria*), common toad (*Bufo bufo*) and the much less common natterjack toad (*Epidaulea calamita*) along with the smooth, palmate and great crested newts (*Lissotriton vulgaris*, *L. helveticus* and *Triturus cristatus*, respectively). Although there appear to be many British Late Pleistocene and Early Holocene sites that contain herpetofaunal skeletal remains, these components of the assemblages in particular remain poorly studied compared with those from warmer climates, with only few that have been published in detail (e.g. Allison et al. 1952). Until relatively recently, species identification has been based on the analysis of osteological criteria, either by direct comparison with reference collections or by using morphometric or geometric morphometric methods (Gaudin 1974; Holman 1998; Bailon 1999; Gleed-Owen 2000; Blain et al. 2015). However, these anatomical methods require a high level of expertise and access to reference collections substantially less common than for other taxa. In addition, they pose difficulties for taxa with little differentiation (genus *Pelophylax*, *Hyla*, *Discoglossus*, etc.), or because of the osteological variability observed within the same taxon. The allocation of bones to a particular species is therefore sometimes questionable. More generally for other taxa, genetic data from DNA sequencing has been used for species identification in wildlife forensics (Dawney et al. 2007) and in some archaeological and palaeontological
assemblages (Newman et al. 2002; Sommer et al. 2009; Murray et al. 2013). However, ancient DNA (aDNA) studies can be relatively time consuming and have limited success rates, particularly with those dating back to the Pleistocene (Höss et al. 1996).

The small size of microfaunal remains places greater limitations on the number of different techniques that can be applied to them. For example, anuran scapulae are typically up to only ~30 mg, and even some of the largest bones (e.g. femur, tibiofibula or ilium) only tend to range up to ~70–80 mg (Gleed-Owen 1998). Although advances in technology do allow for a greater range of genetic information to be recovered from smaller samples, and there are several studies that have investigated microfauna (e.g. Fulton et al. 2013), the relative numbers of aDNA studies have been low whilst sampling still affects larger proportions of the skeletal remains (i.e. the complete bones in many of these cases). This continues to be a substantial sampling risk given that many analyses of ancient material have poor success rates. For example, aDNA analyses were carried out on various fenland sites but were not successful (Beebee 1997).

The development of collagen fingerprinting (also known as zooarchaeology by mass spectrometry, or ZooMS) over a decade ago provided a new alternative biomolecular approach to the species identification of fragmentary animal bone (Buckley et al. 2009). This approach is based on the utilization of type 1 collagen, the dominant protein in bone, which is extracted into solution and enzymatically digested into peptides that are then measured by soft-ionization mass spectrometry, most commonly matrix assisted laser desorption ionization time of flight (MALDI-ToF) mass spectrometry. This collagen fingerprinting approach has in the past been applied to the study of domesticated animals (Buckley et al. 2009) as well as wild fauna (Buckley & Kansa 2011), including marine mammals (Buckley et al. 2014), rodents (Buckley et al. 2016, 2018), bats (Buckley & Herman 2019), and fish (Harvey et al. 2018). Here we present new markers for the species determination of amphibians, concentrating on those present in Britain during the Late Pleistocene and Early Holocene. As a case study, the archaeofaunal assemblage from Pin Hole Cave (Creswell Crags, Derbyshire, UK) was explored to evaluate whether or not improvements could be made upon previous morphological interpretations.

Pin Hole Cave, Creswell Crags, UK

Pin Hole Cave (SK533742) is a narrow cave situated in the northern outcrop of the Creswell Gorge. It is a particularly important archaeological site in the British Isles because it acts as the type site for the characteristic mammal assemblage of Marine Isotope Stage 3 fauna, so its thorough understanding is of high value in palaeo-biostratigraphy (Currant & Jacobi 2001). The Creswell Crags are also known for Britain’s earliest cave art (Bahn et al. 2003), yielding insights into one of the three known phases of ancient human occupations in the area, starting with Neanderthals. Pin Hole Cave has an entrance in the north side of the gorge and measures ~31 m long by only ~1–2 m wide, but with sediment build-up that likely originates from being washed down into the cave through small fissures in the limestone, mainly in Devensian times (~50 to 10 ka; Armstrong 1932). Excavations in the late 19th century and early 20th century revealed at least two principal sediment bodies dating to the Pleistocene period, an upper red cave earth and a lower yellow cave earth but with faunal remains and lithic artefacts found throughout both (Armstrong 1932). In the late 20th century, excavations were once again carried out in two small areas approximately 30 m into the cave. One of these was ~1.5×1.0 m at the top of the sequence, and the other ~1.0×0.5 m investigating much earlier deposits at the base (Jenkinson 1989) to more carefully obtain microfaunal remains. These remains recovered in the 1980s were the source of the collagen fingerprints analysed here, concentrating mainly on specimens that would not be identifiable from their morphological characteristics.

Only a limited number of analyses have been carried out previously on the amphibian remains from Pin Hole Cave (Gleed-Owen 1998). These were on a small number of specimens from the ‘Armstrong spoil heap’, which recovered a single smooth newt bone, a handful of common toad elements (n = 2) and a dominance of common frog (n = 20). However, given that a greater range of species are known from contemporary sites of the region (Gleed-Owen 1998) the aims of this study were to investigate the amphibian component of the archaeofaunal assemblage collected from Pin Hole Cave in the 1980s (which was dominated by microfaunal remains) using collagen fingerprinting as a means of getting a better understanding of the contemporary herpetofauna of this key site.

Material and methods

This study analyses the collagen fingerprints of 6823 specimens previously collected as part of an earlier study targeting megafaunal bone fragments using a relatively non-destructive approach to the morphology of the bone specimens (which itself derived from 12 317 specimens originally analysed; Buckley et al. 2017). In brief, this involved the incubation of the bone specimens with 0.3 M hydrochloric acid (HCl; Fluka, UK) for 3 h prior to the solution being removed and filtered using 10 kDa molecular weight cut-off ultrafilters into 50 mM ammonium bicarbonate (ABC; Sigma, UK) which was then digested with the enzyme trypsin (Promega, UK) overnight for 18 h. After digestion, 2 μL of peptide solution was mixed with alpha-cyano hydroxycinnamic acid matrix (Sigma, UK), spotted onto a stainless steel target plate and allowed to air dry. The spots were then analysed
Table 1. Collagen peptide mass biomarkers (rounded down) for the amphibian taxa considered in this study (sequences are given in Table S1 where possible; labels following Buckley (2016) with a complete label map for X. laevis given in Table S2).* Tandem mass spectrum also shown here for the variant with one fewer hydroxylated proline for further confidence in placement of the P-V amino acid substitution; superscripted S numbers relate to the Supporting Information figures that best show the closest inferred sequences.

| Species                          | 1t47 | 1t28 | 1t62 | 1t21 | 2t45 | 2t75/76 |
|----------------------------------|------|------|------|------|------|---------|
| B. bufo                          | 11343 | 1367 | 14577310 | 1572515 |       |         |
| E. calamita                      | 1134 | 1393 | 142743 | 1572 |       |         |
| R. temporaria                     | 1119 | 1395586-57 | 145571 | 1588 |       |         |
| R. arvalis                       | 1119 | 1393 | 145571 | 1588 | 1588814 | 19366517 |
| R. dalmatina                     | 111953 | 139355 | 144551121512 | 1588 |       |         |
| P. ridibundus                    | 1119 | 1393 | 145369 | 1588 |       |         |
| T. cristatus                     | 1361 |       | 14275545 | 1572 | 15801596 | 19666518 |
| L. vulgaris                      | 1393 |       | 142743 | 1588 | 15941610 | 19666518 |
| L. helveticus                    | 1393 |       | 142743 | 1572 | 15941610 |         |

using a Bruker Ultraflex II MALDI-ToF mass spectrometer (Buckley et al. 2017). Modern samples were also analysed following this approach. Bones of common toad (B. bufo) and common frog (R. temporaria) were obtained from the National Museums Scotland, great crested newt (T. cristatus), smooth newt (L. vulgaris) and palmate newt (L. helveticus) were obtained from the University of Sheffield Department of Archaeology, bones of natterjack toad (E. calamita), moor frog (R. arvalis) and agile frog (R. dalmatina) were acquired from the collection of ‘Ecole Pratique des Hautes Etudes’ housed in the CEFE-CNRS in Montpellier, France, and a specimen of marsh frog (Pelophylax ridibundus) was sampled from the Natural History Museum of Vienna (Austria), the latter being a representative of a genus for which it is unclear whether it was native to the UK or not (Beebee et al. 2005). The peptide digest aliquot of one each of these species (except for P. ridibundus) was also subject to LC-Orbitrap Elite tandem mass spectrometry following a modified method of Buckley et al. (2015) in order to assist with peptide sequence identification (the gradient was run over 30 min instead of 60, with peptide standards and blanks run between each analysis), which relies upon probability based matching of peptide fragment (tandem mass spectrometry) spectra against a given database of sequences. For this the COL1A1 and COL1A2 sequences from the western clawed frog (Xenopus tropicalis; KAE8575433.1 & NP_001072718.1) and the African clawed frog (X. laevis; BAA94972.1 & XP_018121777.1), as well as those of the American bullfrog (Lithobates catesbeianus; BAA29028.1 & PIO33778.1), the Tibetan frog (Nanorana parkeri; XP_018408224.1 & XP_018416270.1), and the COL1A1 sequence of the Japanese fire-bellied newt (Cynops pyrrhogaster; Q9YIB4.1) and COL1A2 sequence of the axolotl (Ambystoma mexicanum; ADZ04657.1) were retrieved from a protein BLAST search of the X. laevis sequence from UniProt and used to assist sequence interpretation; for collagen variation studies, the two-lined caecilian (Rhinatrema bivittatum; XP_029428289.1 & XP_029444760.1) and the tiny Cayenne caecilian (Microacaelia unicolor; XP_030076719.1 & XP_030056470.1) were also included. The mass spectrometry proteomics data have been deposited with the ProteomeXchange Consortium via the PRIDE partner repository with the data set identifier PXD017143.

Results

Taxon discrimination

As could be inferred from sequence analysis alone (of X. tropicalis, X. laevis and N. parkeri), the peptide marker that we consider highly conserved across most terrestrial mammals at m/z 1105 (GVQGPPGPAGPR where here and hereafter the underlining indicates hydroxylation; 1147) appears to be species-specific in some taxa as it has the sequence GAQGPPGPQPR in X. tropicalis (at m/z 1134) but GAQGPPGPQGAR in X. laevis (at m/z 1108; also see Harvey et al. 2019). The only currently known mammals to deviate from the GVQGPPGPAGPR sequence are the cetaceans (with sequence GVQGPPGPAGPR at m/z 1079 (Buckley et al. 2014); note the greater change in mass than would be expected by sequence alone caused due to the change in the hydroxylated residue, creating a further loss of 16 Da). In Lithobates catesbeianus the sequence is GAIGPPQGPR (m/z 1119) and in N. parkeri is GALGPPPQGPR (also m/z 1119).

The collagen peptide mass fingerprints acquired from the analysis of bone from different reference species yielded some taxonomically informative biomarkers. For example, both the common and natterjack toads (i.e. members of the Bufonidae family) yielded a peak observed at m/z 1134 (i.e. peptide 1t147 – see Tables S1, S2), whereas all frogs studied here (Rana and Pelophylax of the Ranidae family) had a homologous peak at m/z 1119 (Table 1; its homologue in the newts studied here is unclear from the fingerprints but in the fire-bellied newt (C. pyrrhogaster) it has the sequence GGQGA-PGAQGPR, predicted to yield a signal at m/z 1052). Genus-level markers were also observed (e.g. Epidalea at m/z 1443 and Bufo at m/z 1473, or Pelophylax at m/z 1469
and Rana at m/z 1471) but notably in both cases where multiple species of the same genus were present (i.e. Rana and Lissotriton), several variations could be observed (e.g. Fig. 1).

Archaeological amphibian remains

In the archaeological samples from the 1980s excavations at Pin Hole Cave analysed here, at least 146 could be identified as common frog (Rana temporaria) through the combination of the markers at m/z 1119, 1395 and 1455/71. Furthermore, at least 139 contained both the m/z 1134 marker and m/z 1367, indicative of common toad (B. bufo) whereas eight contained m/z 1393 and m/z 1427/43 in addition to the toad marker at m/z 1134, therefore indicative of the natterjack toad (E. calamita; Fig. 2).

Interestingly, a single specimen yielded a fingerprint indicative of moor frog (R. arvalis; Fig. 3).

Discussion

Collagen type 1 variation in amphibians

The manually selected biomarkers, based on relative signal intensity and variation across the taxa of interest, unexpectedly derived more from the COL1A1 chain than the COL1A2 chain, despite the latter being considered much more variable in mammalian type 1 collagen (e.g. Buckley et al. 2016). However, comparison of the number of differences observed in the sequence information (Table 2) clarifies that even in amphibians the COL1A2 chain remains substantially more variable,
twice as much in many instances. Therefore, although it is possible that some peptide markers may have been given less preference here due to difficulties in assigning sequence, it appears as though they are preferentially not being observed in the MALDI-ToF mass spectrometric analysis. This phenomenon has been noted before (Buckley 2016), but not to the extreme that we see a greater dominance of more taxonomically informative COL1A1 markers over those of COL1A2 within the MALDI-ToF peptide mass fingerprint.

In the fingerprints there are also many more potential biomarkers present that could be utilized. Those presented here (Table 1) are what we consider the minimum number of peptide species biomarkers required for distinguishing between this particular suite of taxa; machine learning approaches (e.g. Gu & Buckley 2018) could more confidently inform of greater combinations of markers, as well as with a greater range of taxa, should a greater database of known modern and archaeological specimens be incorporated. Adoption of such a machine learning approach could also more readily utilize markers observed in this study but not selected due to commonly observed mass shifts, e.g. *m/z* 1282 and *m/z* 1298 as further means of support to separate *T. cristatus* and *L. helveticus* from *L. vulgaris*; it is important to note that proteins undergo biological (e.g. oxidation of proline (P) residues, causing a +16 Da shift) as well as ‘diagenetic’ changes (e.g. oxidation of methionines (M), also +16 Da shift, and deamidations of asparagine (N) and glutamine (Q) residues (resulting in a +1 Da shift)). However, these mass shifts have so far always been observed in addition to their peaks representing their non-deamidated forms (e.g. Chowdhury et al. 2019), whereby the ratios of modified to non-modified forms have even been used to support endogeneity of ancient protein (e.g. Rybczynski et al. 2013), in some cases identifying intrusive bone (Buckley et al. 2017).

### Capacity for palaeoenvironmental inferences

The ability to make objective identifications of amphibian skeletal remains to the species level is clearly of significant importance in the study of archaeo- and palaeofaunal assemblages. Recovery of such information from a number of Pleistocene sites is in particular of great interest for inferring palaeotemperatures (Blain et al. 2012, 2014), but also useful for palaeobiostratigraphy in the case of Britain due to the scope for evaluating past land connections (Schreve 2001). There is also greater significance in the absence of some species in particular. For example, the presence of *E. calamita* is sometimes thought more indicative of salt marshes, and being thermophilous is considered as indicative of mean minimum temperatures being above ~15 °C (Beebee 1976). Further evaluation of this particular species could also lead to insights into whether or not the currently inferred scattered distribution is a relic from what was once a more continuous occurrence, with the earliest known dates for this species in Britain at ~12 992 cal. a BP (Gleed-Owen 1998). However, given that the absence of particular newt species, such as *L. vulgaris*, can be considered indicative of climatic conditions it is highly

---

**Table 2.** Number of amino acid differences between known amphibian COL1A1 (unshaded) and COL1A2 (shaded) sequences (excludes 39 C-terminal residues in the latter across all taxa, but includes *A. mexicanum* less 31 N-terminal residues and one central residue, *N. parkeri* missing 88 central residues, *Lithobates catesbeianus* missing 18 central residues and *M. unicolor* missing 11 central residues, all from COL1A2; ‘central’ is here specified as more likely to be conserved for structural reasons within the helical part of the protein).

| X. tropicalis | X. laevis | Lithobates catesbeianus | N. parkeri | C. pyrrhogaster | A. mexicanum | Rhinatrema bivittatum | M. unicolor |
|---------------|-----------|-------------------------|------------|----------------|-------------|----------------------|------------|
| X. tropicalis | 88        | 41                      | 89         | 110            | 132         | –                    | 145        | 148        |
| X. laevis     | 88        | –                       | 108        | 122            | 128         | –                    | 151        | 162        |
| Lithobates    | 188       | 192                     | 69         | 130            | –           | 153                   | 162        |
| catesbeianus  | –         | –                       | –          | –              | –           | –                    | –          |
| N. parkeri    | 258       | 266                     | 200        | 145            | –           | 168                   | 169        |
| C. pyrrhogaster | 231      | –                       | –          | 317            | –           | 146                   | 149        |
| A. mexicanum  | 232       | 229                     | 261        | 325            | 236         | –                    | 97         |
| Rhinatrema    | –         | –                       | –          | –              | –           | –                    | –          |
| bivittatum    | –         | –                       | –          | –              | –           | –                    | –          |
| M. unicolor   | 255       | 257                     | 282        | 344            | 260         | 182                   | –          |
beneficial to have a method that can screen as much of an assemblage as possible to reduce potential sampling biases towards skeletal elements more likely to survive (which could involve species-specific biases).

Overcoming the ambiguities

As for any taxon, some skeletal elements are more difficult to identify to the species level than others; for example, there is a much greater difficulty in identifying ilial remains in anurans. There are many cases where such difficulties in identification are reported, such as from the Bronze Age barrow site at Deeping St. Nicholas (Lincolnshire, UK) where there were several complete elements (e.g. phalanges, radioulnae and fibulare) as well as numerous fragments that could not be determined below Anura (Gleed-Owen 1998). This is despite years of expertise in morphological identifications of amphibian remains – itself a rare specialism in zooarchaeology and palaeontology. The taxonomic resolution presented here for collagen fingerprinting is particularly promising in its potential application to previous reports of ambiguities, e.g. separating ‘R. arvalis/dalmatina’ from both the Saxon sites of Gosberton Choppike Drove (Lincoln) and Terrington St Clement (Norfolk), where species identification was not achieved (Gleed-Owen 1998). This could yield important inferences relating to the palaeobiogeography of this species, which is no longer present in
Britain. Such difficulties in species level identifications similarly apply to the newts, for example distinguishing *L. vulgaris* from *L. helveticus* at the Pleistocene Potter’s Cave, Dyfed (Gleed-Owen 1998). In addition to the identification of obvious ambiguities, there are numerous examples in the literature of misidentifications. For example, at the Holocene deposits in Whitemoor Channel (Bosley, east Cheshire), *E. calamita* was initially identified by Holman & Stuart (1991) but later re-examination re-identified these specimens as *B. bufo* (Gleed-Owen 1998). Likewise for the newts, *L. vulgaris* was identified at Boxgrove (Holman 1992) but later believed to be another misidentification (Gleed-Owen 1998).

The potential discovery of *R. arvalis* at Pin Hole Cave is particularly informative (although based on a poorly resolved peak), being found in a range of environments but usually associated with much wetter habitats than *R. temporaria*, which is considered strictly associated with humid conditions (Necas *et al.* 1997). Not only is it a species known to prefer slightly acidic bodies of pH 5–6 (Ischenko 1997), but it is also considered an indicator of healthy and unpolluted moorland pools (Corbett 1989). Although no longer present in the UK, remains of *R. arvalis* have been reported from interglacial sites in southeast England (Holman 1987; Holman & Clayden 1988; Holman *et al.* 1988, 1990; Ashton *et al.* 1994) as well as the ambiguous specimens from the later Holocene mentioned above. Therefore, the discovery here represents the most northerly site in the UK for this species, and the only one from the Late Pleistocene/Early Holocene period, somewhat bridging a substantial temporal gap.

**Conclusions**

The species identification of animal bone using traditional techniques has been practised for many decades but has largely been dominated by investigations into megafaunal studies. This is despite the greater palaeoenvironmental inferences that can be made using microfauna, in particular the remains of amphibians. This study presents a morphologically non-destructive approach to the high-throughput study of archaeological amphibian remains that can be used to rapidly make inferences from assemblages that number in the thousands of specimens. Most importantly, the taxonomic resolution that can be achieved is apparently at the species level, which could  be used to not only place confidence in ambiguous identifications, but effectively assess near complete assemblages. This overcomes one of the most troublesome taphonomic biases for the study of archaeological assemblages, which in the case of amphibian remains is vital given interests in making environmental inferences based on the absence of particular species.

**Acknowledgements.** – We gratefully acknowledge funding from the NERC (NE/H015132/1) for the original study and the Royal Society for fellowship (UF120473) funding and thank the Creswell Crags Heritage Centre for access to the original archaeological specimens, as well as the University of Sheffield Department of Archaeology, Jerry Herman at the Natural History Museum of Vienna for access to additional reference specimens. We also thank Chris Gleed-Owen and several anonymous reviewers for comments on the manuscript at different stages of its preparation.

**Author contributions.** – MB conceived the study design, carried out the laboratory analyses and wrote the manuscript; MC provided reference materials and contributed to writing the manuscript.

**References**

Allison, J., Godwin, H., Warren, S. & Hinton, M. A. C. 1952: Late-glacial deposits at Nazinge in the Lea valley, North London. *Philosophical Transactions of the Royal Society of London, Series B* 236, 169–240.

Ashton, N. M., Bowen, D. Q., Holman, J. A., Hunt, C. O., Irving, B. G., Kemp, R. A., Lewis, S. G., McNabb, J., Parfitt, S. & Seddon, M. B. 1994: Excavations at the Lower Palaeolithic site at East Farm, Barnham, Suffolk 1989-92. Journal of the Geological Society of London 151, 599–605.

Armstrong, A. L. 1932: Excavations in the Pin Hole Cave, Creswell Crags, Derbyshire. Proceedings of the Prehistoric Society of East Anglia 6, 330–334.

Bahn, P., Petitt, P. & Ripoll, S. 2003: Discovery of Palaeolithic cave art in Britain. *Antiquity* 77, 227–231.

Bailon, S. 1999: Différenciation ostéologique des anoures (Amphibia, Anura) de France. *In Desse, J. & Desse-Berset, N. (eds.): Fiches d’ostéologie animale pour l’archéologie. Série C: varia, 1–38. Centre de Recherches Archéologiques du CNRS, Valbonne."

Beebe, T. 1976: The natterjack toad (*Bufo calamita*) in the British Isles: a study of past and present status. *British Journal of Herpetology* 5, 515–521.

Beebe, T. J., Buckley, J., Evans, I., Foster, J. P., Gent, A. H., Gleed-Owen, C. P., Kelly, G., Rowe, G., Snell, C., Wycherley, J. T. & Zeisset, I. 2005: Neglected native or undesirale alien? Resolution of a conservation dilemma concerning the pool frog *Rana lessonae*. *Biodiversity & Conservation* 14, 1607–1626.

Beebe, T. J. C. 1997: Extraction and analysis of DNA from sub-fossil frog bones. Report for English Nature, Peterborough. https://doi.org/10.1186/s40555-014-0094-3.

Blain, H.-A., Bailon, S. & Cuenca-Bescós, G. 2008: The Early-Middle Pleistocene palaeoenvironmental change based on the squamate reptile and amphibian proxies at the Gran Dolina site, Atapuerca, Spain. *Palaeogeography, Palaeoclimatology, Palaeoecology* 261, 177–192.

Blain, H.-A., Lózano-Fernández, I. & Böhme, G. 2015: Variation in the ilium of central European water frogs *Pelophylax* (Amphibia, Ranidae) and its implications for species-level identification of fragmentary anuran fossils. *Zoological Studies* 54, 4. https://doi.org/10.1186/s40555-014-0094-3.

Blain, H.-A., Panera, J., Uribeportete, D., Rubio-Jara, S. & Pérez-González, A. 2012: Characterization of a rapid climate shift at the MIS 8/7 transition in central Spain (Valdociarros II, Autonomous Region of Madrid) by means of the herpetological assemblages. *Quaternary Science Reviews* 47, 73–81.

Blain, H.-A., Santonja, M., Pérez-González, A., Panera, J. & Rubio-Jara, S. 2014: Climate and environments during Marine Isotope Stage 11 in the central Iberian Peninsula: the herpetofaunal assemblage from the Acheulean site of Aridos-1, Madrid. *Quaternary Science Reviews* 94, 7–21.

Blaustein, A. R., Walls, S. C., Bancroft, B. A., Lawler, J. J., Searle, C. L. & Gervasi, S. S. 2010: Direct and indirect effects of climate change on amphibian populations. *Diversity* 2, 281–313.

Buckley, M., Farina, R. A., Lawless, C., Tambussi, P. S., Varela, L., Carlini, A. A., Powell, J. E. & Martinez, J. G. 2015: Collagen sequence analysis of the extinct giant ground sloths Lestodon and Megath-
**Fig. S1.** Collagen peptide mass fingerprints from modern *Bufo bufo* and *Rana temporaria*.

**Fig. S2.** Collagen peptide mass fingerprints from the three known newts in the UK.

**Fig. S3.** Tandem mass spectrum of peptide at *m/z* 1119 (from *R. dalmatina*).

**Fig. S4.** Tandem mass spectrum of peptide at *m/z* 1134 (from *B. bufo*).

**Fig. S5.** Tandem mass spectrum of peptide at *m/z* 1393 (from *R. dalmatina*).

**Fig. S6.** Tandem mass spectrum of peptide at *m/z* 1379 (from *R. lessonae*).

**Fig. S7.** Tandem mass spectrum of peptide at *m/z* 1395 (from *R. lessonae*).

**Fig. S8.** Tandem mass spectrum of peptide at *m/z* 1427 (from *T. cristatus*).

**Fig. S9.** Tandem mass spectrum of peptide at *m/z* 1443 (from *T. cristatus*).

**Fig. S10.** Tandem mass spectrum of peptide at *m/z* 1473 (from *B. bufo*).

**Fig. S11.** Tandem mass spectrum of peptide at *m/z* 1455 (from *R. dalmatina*).

**Fig. S12.** Tandem mass spectrum of peptide at *m/z* 1471 (from *R. dalmatina*).

**Fig. S13.** Tandem mass spectrum of peptide at *m/z* 1572 (from *B. bufo*).

**Fig. S14.** Tandem mass spectrum of peptide at *m/z* 1588 (from *R. arvalis*).

**Fig. S15.** Tandem mass spectrum of peptide at *m/z* 1596 (from *T. cristatus*).

**Fig. S16.** Tandem mass spectrum of peptide at *m/z* 1610 (from *L. vulgaris*).

**Fig. S17.** Tandem mass spectrum of peptide at *m/z* 1936 (from *R. arvalis*).

**Fig. S18.** Tandem mass spectrum of peptide at *m/z* 1966 (from *R. dalmatina*).

*Table S1.* Sequences of biomarkers as determined via LC-Orbitrap Elite mass spectrometry.

*Table S2.* Peptide label map of *Xenopus laevis*.

*Data S1.* Supplementary COL1 sequences.