



OPEN

Palaeoproteomics gives new insight into early southern African pastoralism

Louise Le Meillour^{1,2}, Séverine Zirah², Antoine Zazzo¹, Sophie Cersey³, Florent Déroit⁴, Emma Imalwa⁵, Matthieu Lebon⁴, Alma Nankela⁶, Olivier Tombret^{1,4}, David Pleurdeau⁴ & Joséphine Lesur¹

The advent of domestication is a major step that transformed the subsistence strategies of past human societies. In Africa, domestic caprines (sheep and goat) were introduced in the north-eastern part of the continent from the Near East more than 9000 years ago. However, their diffusion southwards was slow. They are thought to have made their first appearance in the southern part of the continent *ca.* 2000 years ago, at a few Later Stone Age sites, including Leopard Cave (Erongo region, Namibia), which provided the oldest directly dated remains assigned to sheep or goat on the basis of morphology of bones and teeth. However, similarities in morphology, not only between these two domesticated caprine species, but also between them and the small wild antelopes, raised questions about the morphological species attribution of these remains. Additionally, the high fragmentation of the site's osteological remains makes it difficult to achieve species-level taxonomic identification by comparative anatomy. In this paper, we report molecular species identification of the Leopard Cave remains using palaeoproteomics, a method that uses protein markers in bone and tooth collagen to achieve taxonomic identification of archaeological remains. We also report new direct radiocarbon dates. Wild antelope remains from museum collections were used to enrich the available protein record and propose *de novo* type I collagen sequences. Our results demonstrate that the remains morphologically described as domesticates actually belong to a wild antelope species and that domestic caprines first appeared at Leopard Cave 1500 years later than previously thought. This study illustrates that the use of palaeoproteomics coupled with direct radiocarbon dates is particularly suited to complement classic zooarchaeological studies, in this case concerning the arrival of the first herding practices in arid environments.

Understanding how past human populations interacted with their environment, and particularly with other animals, allows understanding large parts of societies, their organisation and their economy. Climatic change is one of the factors that may have pressured human populations to adapt their subsistence strategy and diet. With Saharan and Sahelian aridification starting at the end of the African Humid Period (*ca.* 14,800–5500 years before present [BP]¹), Africa experienced its last major climatic change² and humans had to adapt to new environments³. Thus, the inception of domestication may have been constrained by fewer food resources and their increased unpredictability⁴. Given that they have no wild representatives on the continent, it is now commonly accepted that domestic sheep (*Ovis aries*) and goat (*Capra hircus*) were introduced from the Near East^{5,6}. The first archaeozoological evidence of these species dates from the end of the 9th and the beginning of the 8th millennium BP, in the eastern part of the Sahara^{7–13}. Yet, their diffusion across the continent was slow: the first

¹UMR 7209 Archéozoologie, Archéobotanique: Sociétés, Pratiques et Environnements (AASPE), Muséum national d'Histoire naturelle, CNRS, 55 rue Buffon, 75005 Paris, France. ²UMR 7245 Molécules de Communication et Adaptations des Microorganismes (MCAM), Muséum national d'Histoire naturelle, CNRS, 63 rue Buffon, 75005 Paris, France. ³USR 3224 Centre de Recherche sur la Conservation (CRCC), Muséum national d'Histoire naturelle, CNRS, Ministère de la Culture, 36 rue Geoffroy Saint Hilaire, 75005 Paris, France. ⁴UMR 7194 Histoire naturelle de l'Homme Préhistorique (HNHP), Muséum national d'Histoire naturelle, CNRS, UPVD, 1 rue René Panhard, 75013 Paris, France. ⁵University of Namibia, Windhoek, Namibia. ⁶National Heritage Council of Namibia, 153 Dr. AB May and Rev. Michael Scott streets, Aussenplanplatz, Windhoek, Namibia. ✉email: llemeillour@mnhn.fr; jolesur@mnhn.fr



Figure 1. Map of southern Africa synthesizing all Later Stone Age sites from which caprine remains have been identified, including Leopard Cave (red dot), sites with remains attributed morphologically to caprines (after ref.¹⁴) (green dots), and sites with directly radiocarbon dated remains attributed morphologically to caprines (after ref.¹⁵) (orange dot). Site names circled in red refer to the three archaeological sites where caprine remains were molecularly identified (ref.¹⁶ and this study for Toteng and Leopard Cave; ref.¹⁷ for Blydefontein). Associated dates are given to the right of the map. The map was generated by Louise Le Meillour and David Pleurdeau using Google Maps (<https://www.google.fr/maps>) for country and river edges and created using Adobe Illustrator CC 2017 (<https://www.adobe.com/>).

domesticated caprines appear in southern Africa around 2000 BP (Fig. 1), and their presence is scarce within the faunal assemblages.

The modalities of the possible routes, timing, and diffusion of caprine herding practices are the subject of a long-standing debate amongst the scientific community^{18–21}. Genetic^{22–25} and linguistic^{26–28} studies point toward an east African origin of the human populations and the herds, both of which could have reached the southernmost tip of the continent *ca.* 2300 years ago (see ref.^{20,29} for a recent review). The hypothesis of a migration of ‘proto-Khoekhoe’ populations is supported both by allelic similarities between Nama people from Namibia and Maasai people from east Africa^{30–32} and by the linguistic merging of click languages^{28,33,34}. The presence of caprine species breeds (fat-tailed sheep, for example^{35,36}) that are only represented in eastern and southern Africa is a supplementary argument in favour of strong networks of exchange between the two regions. Alternatively, the scarcity of remains of domesticated livestock in African Later Stone Age (LSA) archaeological sites has been argued to support the hypothesis of livestock diffusion from the northern part of Africa to the southern tip of the subcontinent by gradual infiltration, as has been demonstrated for other regions³⁷. According to this model, domestic animals were introduced progressively into small, local hunter-gatherer groups through cultural percolation from livestock farmers from the north, implying a ‘low-intensity’ herding economy without any profound social or economic changes²⁰. In this model, major cultural changes (technical or subsistence strategy shifts) that can be archaeologically attested arose considerably later and were linked to a more massive arrival of the Khoekhoe population, around 1000 BP³⁸.

The western margins of the Kgalagadi basin and the Namib Desert, in Namibia, feature prominently in the question whether herding practices originated earlier, from the North, or later, from the South^{19,39,40}. In central western Namibia, only a few archaeological sites have yielded early remains attributed to domesticated caprines, namely Geduld⁴¹, Mirabib⁴², Stripped Giraffe⁴³, Oruwanje 95/1⁴⁴ and Charé^{45,46}. At the LSA site of Leopard Cave, in the Erongo region of the country¹⁵, remains identified morphologically as sheep or goat and directly radiocarbon dated (i.e. dated on the remains themselves) to 2300–2100 years cal. BP have been reported and thus were considered to be the oldest evidence of caprine herding. Extensive fieldwork has been ongoing at the

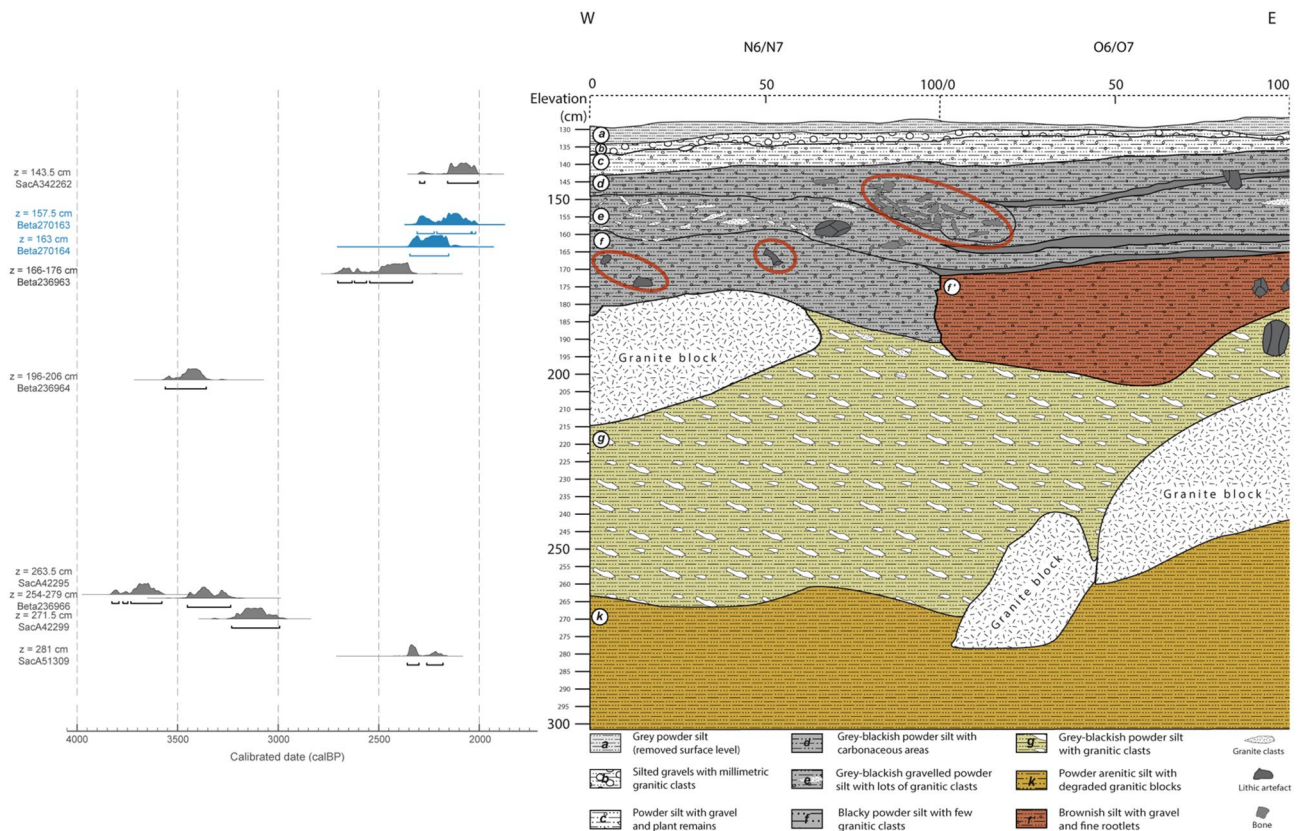


Figure 2. Stratigraphy of squares N6/N7 and O6/O7 in Leopard Cave. The stratigraphic position of the remains included in the study is indicated by red circles. All radiocarbon dates obtained within the sequence are presented by depth. The two dates corresponding to the direct dating of the teeth morphologically identified as caprine are shown in blue.

site, mainly focused on the lower levels and rock art pigments⁴⁷. In addition to the two dental remains previously described¹⁵, the excavations have yielded one mandible from the same layer and 25 remains from a layer higher up in the cave's stratigraphic sequence (Fig. 2 and Table S1), all morphologically identified as possible caprines. The excavation campaigns also allowed refining of the chronostratigraphy of the rock shelter, confirming the integrity of the sequence, particularly for the upper layers, where the remains identified morphologically as caprines were recovered (Fig. 2 and Table S3).

Although the Leopard Cave remains fuelled a debate^{14,18} regarding the timing and process of the arrival of the first herding practices, the secure taxonomic identification of the first introduced caprine species in southern Africa (i.e. whether sheep or goat) is still pending^{48–53}, even though nearly all the archaeological remains in this part of the continent have been attributed to *O. aries*¹⁹ based on comparative anatomy. Sheep and goat are morphologically close⁵⁴; therefore, distinguishing between them is often a real challenge, especially given the high degree of bone fragmentation often observed in African archaeological contexts. The presence of mid-sized wild bovids that also share strong morphological similarities with domestic caprines makes morphological identification of bones and teeth particularly uncertain⁴⁹. This problem is illustrated, for instance, by a recent morphological reassessment of a few bones of mid-sized bovid initially attributed to wild antelope species from the sites of Falls Rock Shelter and Snake Rock, on the Daures (also known as the Brandberg), Namibia, in layers dating from 2000 BP, which were reassigned to sheep⁵⁵. Similarly, ancient DNA studies showed that remains morphologically identified as domestic caprines from Blydefontein rock shelter (South Africa) actually belong to several wild antelope species, including the springbok (*Antidorcas marsupialis*)⁴⁹.

In this context, the use of molecular tools can be of great interest because ancient biomolecules preserved in archaeological remains can reveal features impossible to assess with traditional, morphological methods. In addition to ancient DNA studies, the past few years have witnessed a multiplicity of analyses based on the extraction and characterisation of long-term persistent biomolecules: proteins^{56–61}. Since their amino acid sequence is coded for by the DNA, proteins carry genetically derived features that allow researchers to propose phylogenetic relationships of extinct species^{62–65}, sex attribution^{66,67} or even characterisation of palaeopathogens⁶⁸. The major protein of bone and teeth, type I collagen (constituting 90% of the organic phase of these materials), is composed of amino acid sequences that carry taxon-related differences in their order, allowing species discrimination⁶⁹. Shifts in amino acids of type I collagen allow distinction between domestic caprine species, and palaeoproteomics is therefore particularly suited to the study of the diffusion of domesticates in southern Africa^{16,57}. For the purpose of addressing this question in relation to Leopard Cave, we conducted palaeoproteomic analyses on all






















| DP# | Position in the sequence (Accession number used to align) | Peptide sequence | Species in which the peptide is present | Other ruminant species in which the peptide is present | Reference |
|--------------------------------------|--|--|---|--|--------------------------------|
| Alpha 1 chain type I collagen | | | | | |
| #1 | 919-934 (<i>C. hircus</i> A0A452FHU9) | ... PGEV GG PPGPPG PAGEK ... |    | <i>Bos indicus</i> , <i>Bos mutus</i> , <i>Bos taurus</i> , <i>Bubalus bubalis</i> , <i>Muntiacus muntjak</i> , <i>Muntiacus reevesi</i> , <i>Odocoileus virginianus texanus</i> | Le Meillour <i>et al.</i> 2018 |
| | 927-942 (<i>O. aries</i> W5P481) | ... AGEV GG PPGPPG PAGEK ... |  | | |
| Alpha 2 chain type I collagen | | | | | |
| #2 | 590-607 (<i>O. aries</i> W5NT17) | ... GPPGESGAAG P AG P IGSR ... |   | <i>Bos indicus</i> , <i>Bos mutus</i> , <i>Bos taurus</i> , <i>Bubalus bubalis</i> , <i>Odocoileus virginianus texanus</i> | This study |
| | | ... GPPGESGAAG P AG P IGSR ... |   | <i>Muntiacus muntjak</i> , <i>Muntiacus reevesi</i> | |
| #3 | 692-706 (<i>O. aries</i> W5NT17) | ... GEAG P AG P AG P AGSR ... |    | <i>Bos indicus</i> , <i>Bos mutus</i> , <i>Bos taurus</i> , <i>Bubalus bubalis</i> , <i>Muntiacus muntjak</i> , <i>Muntiacus reevesi</i> , <i>Odocoileus virginianus texanus</i> | This study |
| | | ... GEAG P AG P AG P AGSR ... |  | | |
| #4 | 746-775 (<i>O. aries</i> W5NT17) | ... GEN G PV G PT G P A GAAG P SG P NG P PG P AGSR ... |    | <i>Bos indicus</i> , <i>Bos mutus</i> , <i>Bos taurus</i> , <i>Bubalus bubalis</i> , <i>Muntiacus muntjak</i> , <i>Muntiacus reevesi</i> , <i>Odocoileus virginianus texanus</i> | This study |
| | | ... GEN G PV G PT G P A GAAG P SG P NG P PG P AGSR ... |  | | |
| #5 | 845-877 (<i>O. aries</i> W5NT17) | ... GPSG E PGTAG P PG T PG P Q G LL G AP G FL G LP G SR ... |    | <i>Bos indicus</i> , <i>Bos mutus</i> , <i>Bos taurus</i> , <i>Muntiacus muntjak</i> , <i>Muntiacus reevesi</i> | Buckley <i>et al.</i> 2010 |
| | 623-655 (<i>C. hircus</i> A0A452G3V6) | ... GPSG E PGTAG P PG T PG P Q G FL G PP G FL G LP G SR ... |   | | |

Table 1. Alpha 1 and alpha 2 chains of type I collagen discriminating peptides of *Aepyceros melampus* and *Antidorcas marsupialis*. Specific amino acids are represented in boldface red. Sheep is represented in grey; goat, in brown; springbok, in orange; and impala, in turquoise. Sequences were constructed using the PEAKS database-assisted de novo tool and aligned using Geneious software.

the remains morphologically identified as possible caprines among the mid-sized bovids, including ones identified below the taxonomic level of family.

Results

Palaeoproteomic analyses were carried out on 28 remains from different layers of Leopard Cave. Two of them are dental remains from layer F identified as caprines¹⁵, one is a mandible half from layer E, and the other 25 are various skeletal elements from layer D (Fig. 2). The remains from layer E and D were morphologically identified as either caprines (sheep or goat) or mid-sized bovids (a category that includes sheep, goat, springbok and impala, *Aepyceros melampus*) (Table S1). Given the morphological proximity between these taxa and the absence of protein sequences available in the existing databases for mid-sized wild bovids, we simultaneously conducted proteomics on modern museum collection specimens of an Ethiopian sheep and a Namibian goat, as long as springbok and impala, two species which were present in the vicinity of Leopard Cave at the relevant time periods (Table S2).

Identifying taxonomic markers. Proteomics analysis of the twentieth-century museum collection specimens of springbok and impala allowed us to reconstruct alpha 1 and alpha 2 chains of type I collagen (COL1A1 and COL1A2) using the PEAKS algorithm, which generates database-assisted de novo sequences (see Methods section). The generated type I collagen sequences of springbok and impala were aligned to available bovid reference sequences. The peptides observed in our dataset cover 97% and 67% of the reference mature COL1A1 sequence (i.e. the secreted protein, which only represents the helical regions of collagen) and 91% and 52% of the reference mature COL1A2 sequence for springbok and impala, respectively (Fig. S1 and Fig. S2). Sequence alignments point out several amino acid substitutions between the two species, but also with cattle (*Bos taurus*), sheep, goat and an Asiatic wild caprine, the chiru (*Pantholops hodgsonii*). Several peptide taxonomic markers were thus identified: one on the alpha 1 chain and four on the alpha 2 chain of type I collagen (Table 1).

The sequences are written below, with specific amino acids in boldface and modified amino acids underlined (post-translational modifications). The springbok and the impala present one shared derived peptide on COL1A2, GPPGESGAAG**P**AG**P**IGSR (DP#2; Fig. S3a, Fig. S6). The springbok presents two uniquely derived peptides on COL1A2, that discriminates the species from the others bovids: GEAG**A**AG**P**AG**P**AGSR (DP#3; Fig. S3b, Fig. S6) and GEN**G**PV**G**PT**G**P**A**GAAG**P**SG**P**NG**P**PG**P**AGSR (DP#4; Fig. S3c, Fig. S6). These three peptide markers allow the discrimination of springbok and impala from all other bovid species. In addition, our analyses indicate that the two peptide markers traditionally used for distinguishing sheep from goat on COL1A1¹⁶ and COL1A2^{57,69,70} are shared by the antelope species: the COL1A1 peptide PGEV**GG**PPGPPG**PAGEK** (919–934, DP#1; Table S5) is shared between goat, springbok, and impala, and the COL1A2 peptide GPSG**E**PGTAG**P**PG**T**PG**P**Q**G**LL**G**AP**G**FL**G**LP**G**SR (845–877, DP#5; Table S5) is shared between springbok and sheep (Table 1). For impala, we could not specify the sequence of DP#5, since both the sequences of sheep and goat were proposed as PSM (Table S5). The sequence PGEV**GG**PPGPPG**PAGEK** was detected in almost every goat, impala and springbok sample, but always with at least one missed cleavage, contrary to the sheep sequence AGEV**GG**PPGPPG**PAGEK** (Table S5). This is expected given the disfavoured proteolytic cleavage at the N-terminus of proline residues⁷¹.

Taxonomic identification of archaeological samples and direct dating. The collagen content of the archaeological samples, measured using infrared spectroscopy (specifically, Fourier transform infrared spectroscopy in Attenuated Total Reflectance mode [ATR FT-IR]), varied between 11.2% and 22.1% weight (wt) of collagen preserved (Table S4). Compared with the modern bone value of 22%wt of collagen, the remains from Leopard Cave thus appear well to very well preserved; these values are certainly above the threshold of ~3%wt of collagen determined for radiocarbon and proteomics analysis^{16,72}.

The protein extracts were analysed using liquid chromatography coupled to tandem mass spectrometry (a method known as ultra-high-performance liquid chromatography-tandem mass spectrometry [UHPLC-MS/MS]), and the data generated were examined for collagen correspondence through our local collagen sequence database, containing the de novo sequences of springbok and impala described above, and available collagen references of bovids (see Methods section for all accession numbers used). Protein extractions succeeded on each analysed remain selected for this study. Confidently identified peptides were assigned to alpha 1 and alpha 2 chains of type I collagen with sequence coverages of, on average, 70% of mature collagen for the 28 samples. Different algorithms were used, and raw data were manually assessed by extracting ion chromatograms and verifying fragment spectra of each taxonomic peptide reported for sheep and goat, as well as for the springbok and impala ones previously reported (Table S5).

Identifications were secured for all 28 of the remains by the presence of all the taxonomic markers of the corresponding species (Table 2). The three remains from layers E and F (LC_113, LC_149, and LC_176, respectively) present the COL1A1 peptide marker of goat and the COL1A2 marker of sheep, as well as the three new COL1A2 markers: the one shared between springbok and impala and the two others specific of springbok. The presence of the specific markers on the COL1A2 of springbok allows us to confidently identify the remains to this species (Table 2). One of the remains from layer D (LC_134) present a combination of COL1A1 and COL1A2 markers specific of the impala. The 24 other remains presented COL1A1 and COL1A2 peptide markers of sheep (Table 2, Table S5).

Two of the remains from layer D (LC_131 and LC_135) identified as sheep were selected for radiocarbon dating. These provided ¹⁴C ages of 870 ± 30 BP (789–680 cal. BP, 2 σ) and 960 ± 30 BP (919–760 cal. BP, 2 σ) (Fig. 3). These newly generated direct dating results make sense regarding the position of the dated samples within the stratigraphy (Fig. 2 and Table S3).

Discussion

The proteomics analyses we performed allow us to propose de novo collagen sequences for two wild antelopes, the impala and the springbok, which were not yet available in the international databases. We identified three new markers on the alpha 2 chain of type I collagen that will be useful to identify archaeological remains of closely related mid-sized bovid species. We noticed that the markers of both of the type I collagen chains that are commonly used to discriminate between sheep and goat are shared by springbok and impala. In the case of the impala, we detected in our samples the COL1A2 marker DP#5 of both sheep/springbok and goat, suggesting that there might be an intra-specific variation of the peptide in this species. Since those peptides are widely used for species identification of archaeological remains^{16,57,69,70}, we highly recommend the use of both of the type I collagen chain markers to assess attribution of bones and teeth to domestic caprines, especially in cases where the presence of wild antelopes within the archaeological assemblage is a possibility or already attested. All identified type I collagen peptide markers differ by one amino acid only between the different species, highlighting their phylogenetic proximity and the need to use multiple markers when working on closely related animal remains.

Based on the specific markers identified for springbok and impala and those previously reported in the literature for sheep and goat, we were able to disprove the original species attributions of the Leopard Cave remains dated to ca. 2300–2100 cal. BP. Our results suggest that domestic caprines were not present during this time range, and that they first appear in the cave sequence 1500 years later. We propose that environmental conditions could be one of the factors to explain this late presence of domesticates at Leopard Cave. Indeed, several authors have argued that the medieval warm period (occurring before the Little Ice Age) may have favoured increased precipitation in southern Africa between 1000 and 700 BP^{75–77}. Wetter conditions may have created watering holes and improved grazing land availability and, thus, better environmental conditions for the management of livestock in the Namib region, particularly within each of the three inselbergs of Brandberg, Spitzkoppe, and Erongo. Remains from the LSA sites of Geduld⁴¹, Charè^{45,46}, Mirabib⁷⁸, Falls Rock, and Snake Rock⁴⁰ may represent the earliest archaeological evidence of caprine introduction to what is now Namibia. However, none of these remains have been directly dated or molecularly analysed, raising questions about their antiquity and species determination. As our study illustrates, there is a need to be cautious under these circumstances. Thus, if collagen preservation of remains morphologically identified as domesticates allows it, and bearing in mind recent recommendations on the sampling of faunal remains⁷⁹, we recommend the use of palaeoproteomics coupled with direct radiocarbon dating to document the early presence of caprines in Namibia and, more broadly, in southern Africa. This combination of methods will also help to accurately frame the chronology of the different waves of their introduction.

The archaeological evidence relating to pastoralism is thought to back up genetic and linguistic findings (specifically, the frequency of the lactase persistence allele among populations of 'east Africans' and the linguistic links between the Sandawe and KhoeKwadi languages among Khoekhoe populations)²⁴. The low number of remains of domesticated caprines in LSA sites has been considered additional support for the hypothesis of the diffusion by gradual infiltration. In this model, the process of domestic caprines diffusion through Africa would have left no direct archaeological evidence of early herding practices. This 'invisibility' may also be due to the fact that populations exploited these domesticated animals mainly for their secondary products, such as milk or wool, or even for their carrying capacity, rather than their meat. Thus, the initial number of bones deposited and their





























| Sample code | Layer | Morphological taxonomic attribution | Palaeoproteomics species identification | Conventional ¹⁴ C age | # of identified discriminating peptide that allowed identification |
|-------------|-------|-------------------------------------|---|----------------------------------|--|
| LC_113 | F | Caprine (sheep/goat) |  <i>A. marsupialis</i> | 2190 ± 40 BP | #2 - #3 - #4 |
| LC_114 | D | Bovid size 2/3 |  <i>O. aries</i> | | #1 - #5 |
| LC_115 | D | Bovid size 2/3 |  <i>O. aries</i> | | #1 - #5 |
| LC_116 | D | Bovid size 2/3 |  <i>O. aries</i> | | #1 - #5 |
| LC_117 | D | Bovid size 2/3 |  <i>O. aries</i> | | #1 - #5 |
| LC_118 | D | Bovid size 2/3 |  <i>O. aries</i> | | #1 - #5 |
| LC_119 | D | Bovid size 2/3 |  <i>O. aries</i> | | #1 - #5 |
| LC_120 | D | Bovid size 2/3 |  <i>O. aries</i> | | #1 - #5 |
| LC_121 | D | Bovid size 2/3 |  <i>O. aries</i> | | #1 - #5 |
| LC_122 | D | Bovid size 2/3 |  <i>O. aries</i> | | #1 - #5 |
| LC_124 | D | Bovid size 2/3 |  <i>O. aries</i> | | #1 - #5 |
| LC_125 | D | Caprine (sheep/goat) |  <i>O. aries</i> | | #1 - #5 |
| LC_126 | D | Caprine (sheep/goat) |  <i>O. aries</i> | | #1 - #5 |
| LC_127 | D | Caprine (sheep/goat) |  <i>O. aries</i> | | #1 - #5 |
| LC_128 | D | Caprine (sheep/goat) |  <i>O. aries</i> | | #1 - #5 |
| LC_129 | D | Caprine (sheep/goat) |  <i>O. aries</i> | | #1 - #5 |
| LC_130 | D | Caprine (sheep/goat) |  <i>O. aries</i> | | #1 - #5 |
| LC_131 | D | <i>Ovis aries</i> |  <i>O. aries</i> | 871 ± 30 BP | #1 - #5 |
| LC_132 | D | <i>Antidorcas/Aepyceros</i> |  <i>O. aries</i> | | #1 - #5 |
| LC_133 | D | <i>Antidorcas/Aepyceros</i> |  <i>O. aries</i> | | #1 - #5 |
| LC_134 | D | <i>Antidorcas/Aepyceros</i> |  <i>A. melampus</i> | | #2 - #3 - #4 |
| LC_135 | D | Caprine (sheep/goat) |  <i>O. aries</i> | 959 ± 30 BP | #1 - #5 |
| LC_136 | D | <i>Antidorcas/Aepyceros</i> |  <i>O. aries</i> | | #1 - #5 |
| LC_137 | D | <i>Antidorcas/Aepyceros</i> |  <i>O. aries</i> | | #1 - #5 |
| LC_138 | D | Bovid size 2/3 |  <i>O. aries</i> | | #1 - #5 |
| LC_139 | D | Bovid size 2/3 |  <i>O. aries</i> | | #1 - #5 |
| LC_149 | F | Caprine (sheep/goat) |  <i>A. marsupialis</i> | 2270 ± 40 BP | #2 - #3 - #4 |
| LC_176 | E | Caprine (sheep/goat) |  <i>A. marsupialis</i> | 2430 ± 50 BP | #2 - #3 - #4 |

Table 2. Species identifications of the 28 remains from Leopard Cave. Each sample is identified by its lab code, the layer it was recovered from, the morphological identification, the molecular identification, and, if applicable, the corresponding radiocarbon date. The dates in boldface blue represent the direct dating for this study; the ones in black come from ref.¹⁵. The last column lists the discriminating peptide number present in the corresponding sample used for the species identification.

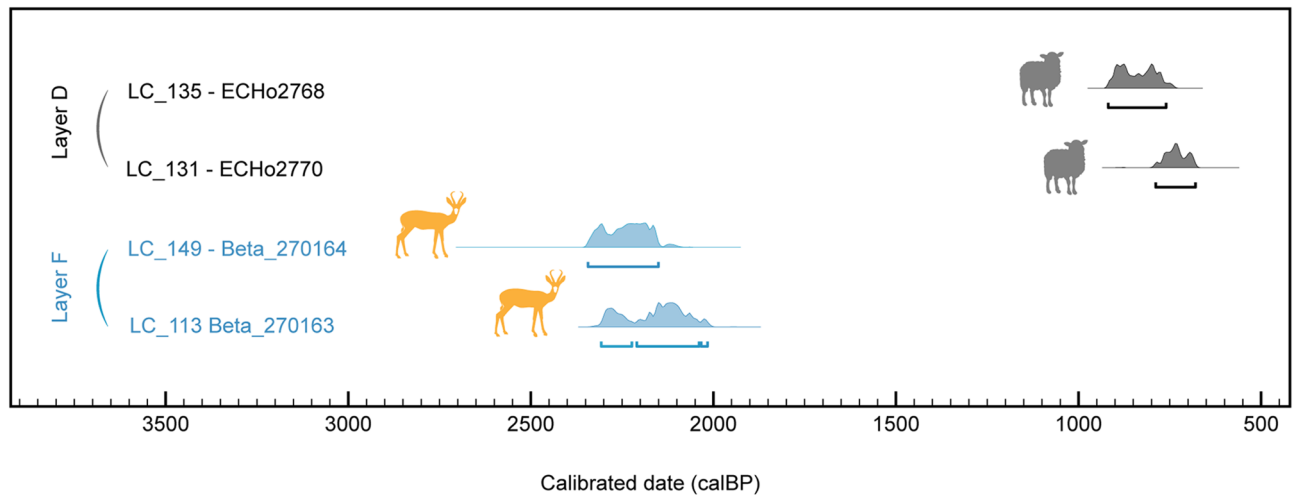


Figure 3. Leopard Cave mid-sized bovid remains direct radiocarbon dates. Direct dates previously reported for what are now known, based on proteomics, to be two springbok remains are shown in blue (LC_113, LC_149). New radiocarbon dates on remains identified as sheep by palaeoproteomics are shown in grey (LC_131, LC_135). Radiocarbon dates were calibrated at 2σ (95.4%) using OxCal 4.3⁷³ and the SHCal13⁷⁴ calibration curve, for the southern hemisphere.

degree of preservation, given variable taphonomic conditions, may influence the presence and detectability of domesticates in faunal assemblages. Hence, traces of herding practices may not be visible archaeologically, even though the animals being herded were eventually consumed by humans. Here again, the use of molecular methods would be appropriate, as the characterisation of pottery residues^{80–82} or dental calculus^{83,84} can reveal the exploitation of secondary products.

In addition, the identification of the first introduced sheep or goat is of particular importance to understanding the ecological, social, and subsistence behaviours of the first (semi-) pastoral societies. Because sheep and goat differ in their ability to adapt to climate change and harsh conditions⁸⁵, people may have had a preference for one over the other when they first started exploiting domesticated caprines. Palaeoproteomics can thus play a leading role in the long-standing debate concerning the introduction of domestic sheep and/or goat throughout the subcontinent. This study particularly highlights the need to apply multiple lines of evidence when evaluating the modalities of the spread of pastoralism in southern Africa (and other parts of the continent), an application that is possible in the new era of biomolecular archaeology.

Methods

We extracted and characterised ancient proteins from all the remains from the site morphologically attributed to caprines, *Antidorcas/Aepyceros* or mid-sized bovid, following recent recommendations on sampling and processing^{79,80}. Having assessed their organic phase preservation using ATR FT-IR spectroscopy, we analysed ancient proteins from 25 bones and 3 dental remains (Table S1). In parallel, and based on the possibility of the presence of wild bovid species, we conducted proteomics analyses on molar roots of modern specimens of two wild antelopes (*Artiodactyla*, Bovidae), *Antidorcas marsupialis* (springbok, Antilopini) and *Aepyceros melampus* (impala, Aepycerotini) from the comparative anatomy collections of the Muséum national d'Histoire naturelle (Paris, France) (Table S2). One sheep and one goat remain from Namibia were also sampled for reference. We then radiocarbon dated two of the archaeological remains molecularly identified as domesticates, using accelerator mass spectrometry (AMS), in order to ascertain their antiquity.

Archaeological remains from Leopard Cave. Leopard Cave is a rock shelter of approximately 50 square metres located in the locality of Omandumba West Farm, in the Erongo region of Namibia¹⁵. Archaeological excavations (2007–2019) have uncovered a long stratigraphic sequence, from the end of the Pleistocene to the Holocene (Fig. 1). The individual layers of the stratigraphic sequence have been previously published using numbers. Recent excavations (2017–2018) have allowed the refinement of the stratigraphy into three different complexes, and layers have now been redocumented by means of letter tags accordingly. Within a faunal assemblage largely dominated by wild game, only two dental remains in layer F (previously layer 6), one mandible in layer E (previously layer 5) and 25 remains in anatomical connection from an upper layer (layer D, previously layer 4) were identified morphologically as caprine or mid-size bovid (Table S1). They were all found in squares N6 and N7.

Collagen preservation estimation. Collagen preservation was estimated prior to any palaeoproteomic analyses, in order to avoid unnecessary destruction of archaeological remains. Less than a milligram of bone or tooth powder was drilled and analysed using ATR FT-IR, following the method described for the radiocarbon sample selection⁷², and the exact same parameters were used to assess archaeological collagen preservation¹⁶.

Protein extraction and treatment. All archaeological remains were sampled using a Dremel drill with a diamond bur that was rinsed with 96% EtOH between each sampling to avoid cross-contamination. A mean of 50 mg of bone or tooth powder was collected and chemically prepared following a modified version of the protocol previously published by our team for archaeological remains from African contexts¹⁶. In brief, samples (both modern and archaeological) were decalcified 8–15 days in a Tris–EDTA buffer (0.05 M and 0.5 M, final pH 7.4, Sigma Aldrich, St. Louis, Missouri, United States). Solutions were replaced every day and kept at 4 °C until further preparation (hereafter, decalcification solution, D). Once they were totally decalcified, pellets were rinsed with Milli-Q water until they were pH neutral. The remaining proteins were then resuspended in ammonium bicarbonate (ABC, 50 mM pH 8, Sigma Aldrich) for 3 h at 65 °C. After centrifugation, supernatants were collected (hereafter extract solutions, E) and placed into clean microtubes. D underwent buffer exchange from Tris–EDTA to ABC using 5 mL Vivaspin (3 kDa, Amicon Ultra, Merck Millipore, Burlington, Massachusetts, United States). Both D and E recovered proteins were reduced with dithiothreitol (DTT, final concentration of 10 mM, 20 min, 56 °C, 350 rpm), alkylated with iodoacetamide (10 mM, 30 min, in the dark, at room temperature) and subjected to trypsin hydrolysis (0.1 µg/µL, Trypsin Gold, Promega, Madison, Wisconsin, United States). Finally, D peptides were recovered, and solutions were desalted by solid phase extraction (SPE) using C8 cartridges (Sep-Pak C8 Plus Short 400 mg Sorbent, Waters, Milford, Massachusetts, United States). Protein concentrations were measured on a NanoVue instrument (GE HealthCare, Chicago, Illinois, United States), yielding amounts ranging between 0.54 and 1.48 µg/µL for archaeological samples and between 0.62 and 1.95 µg/µL for modern samples. Peptides in solutions were kept at –20 °C until further analysis.

Mass spectrometry analyses. The digested samples were analysed by ultra-high-performance-liquid-chromatography coupled to tandem mass spectrometry (UHPLC-MS/MS) following ref. ¹⁶. Briefly, we conducted the UHPLC separation on an Ultimate 3000-RSLC (Thermo Fisher Scientific, Waltham, Massachusetts, United States) using a RSLC Polar Advantage II Acclaim column (2.2 µm, 120 Å, 2.1 × 100 mm, Thermo Fisher Scientific). We used a mobile phase gradient composed of A: H₂O + FA 0.1% and B: ACN + FA 0.08%, with a flow rate of 300 µL/min. The Electrospray-Quadrupole-Time of Flight (ESI-Q-TOF) mass spectrometer (Maxis II ETD, Bruker Daltonics, Billerica, Massachusetts, United States) was used in positive mode on the *m/z* range 200–2200 in data-dependent auto-MS/MS mode. MS/MS spectra were generated using collision-induced dissociation, with ion selection set to *m/z* 300–2200 and charge states set to 1⁺ to 5⁺. Data were automatically calibrated with sodium formate injected at the beginning of each analysis. An ABC blank was performed every five analyses to avoid cross-sample contamination.

The LC–MS/MS data were converted to .mgf format using DataAnalysis software (version 4.4, Bruker Daltonics). Sequences of the two wild antelopes species were reconstructed using the database-assisted de novo tool of PEAKS Studio software⁸⁶ (version 7.5, BioInformatics Solutions, Waterloo, Canada). The search was conducted on a database combining the COL1A1 (*B. taurus* P02453, *B. indicus* A0A4W2FAL4, *B. mutus* L8IV51, *C. hircus* A0A452FHU9, *O. aries* W5P481, and *P. hodgsonii* XM_005964647.1) and COL1A2 (*B. taurus* P02465, *B. indicus* A0A4W2FTM9, *B. mutus* L8HQF7, *C. hircus* A0A452G3V6, *O. aries* W5NTT7, and *P. hodgsonii* XM_005985683.1) sequences of bovids. The entire PEAKS workflow was set to ‘trypsin’ enzymatic digestion, allowing 2 missed cleavages and including the following modifications: fixed carbamidomethylation (C) and variable oxidation (M), deamidation (NQ), hydroxylation (P) and phosphorylation (ST). Error tolerance was set to 10 ppm for precursors and 0.02 Da for fragment ions. Both SPIDER and PEAKS DB tools searches were used to propose the new COL1A1 and COL1A2 sequences. Each amino acid substitution proposed by the software was verified manually by looking at the MS/MS spectra, using the reference sequence that allowed the best coverage percentage. The de novo scores were set to 1% FDR, score of 50% and protein hit ≥ 20. The combined Mascot protein results were generated with ProteinScape Version 4.0.2 (Bruker Daltonics), using the Bovidae collagen database used in PEAKS, complemented by the impala and springbok sequences generated de novo. Once confidently identified, the COL1A1 and COL1A2 were aligned using Geneious⁸⁷ software (version 11.0.3, Auckland, New Zealand; substitution matrix Blosum62, gap open penalty 12) to corresponding protein sequences of *B. taurus* (NM_001034039.2 and NM_174520.2), *B. indicus* (XM_019981159.1 and XM_019959690.1), *B. mutus* (XM_005890325.2 and XM_005909695), *C. hircus* (XM_018064893.1, XM_018064894.1 and XM_0056778936.3), *O. aries* (XM_015098715.1 and XM_004007726.3), *O. aries musimon* (XM_012127623.2, XM_012127624.2 and XM_012146567.2) and *P. hodgsonii* (XM_005964647.1 and XM_005985683.1).

Archaeological samples digests data were searched against a restricted protein database using Mascot software (version 2.4.1, www.matrixscience.com, London, UK⁸⁸). The database was assembled using the National Center for Biotechnology Information (NCBI) collagen sequences (last updated December 9, 2018; 88530 sequences for 78817228 residues, all entries). The same set of parameters used for the PEAKS search was used for the Mascot search. Since the results did not allow to propose clear species identification, they were analysed using ProteinScape software (Bruker, Mannheim, Germany) by implementing the new COL1A1 and COL1A2 sequences of springbok and impala previously de novo reconstructed to the bovid species type I collagen sequence databases. Note that we cannot rule out Leu/Ile substitution in the marker sequences as standard MS/MS cannot differentiate between these two isomeric amino acids.

Radiocarbon dating. We immersed coarse bone powder (0.3–0.7 mm) in 1 M HCl for 20 min under continuous stirring. We then separated the acid-insoluble residues from the solution by centrifugation and rinsed them with Milli-Q water, using vortex. We then immersed them in 0.1 M NaOH for 30 min under stirring, with a change of solution after 10 min. After a new rinsing with Milli-Q water using centrifugation, we immersed the alkali-insoluble residues in 0.01 M HCl and performed solubilisation at 95 °C overnight. We filtered the final solution on MF-Millipore membranes (mixed cellulose ester membranes of 5.0 µm pore size from Fisher

Scientific, Illkirch, France) before freeze-drying them and collecting them for analysis. We performed sample graphitization using the automated AGE 3 graphitization unit and AMS measurements using the compact AMS ECHOMICADAS at the Laboratoire des Sciences du Climat et de l'Environnement (LSCE, CEA, CNRS, UVSQ).

Data availability. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE⁸⁹ partner repository, with the dataset identifier PXD017519 and 10.6019/PXD017519. Each file name in ProteomeXchange corresponds to the nomenclature used for the sample's code, both in this paper and in the associated supplementary data. Sequences of type I collagen for springbok and impala have been deposited on the UniProt database.

Received: 6 January 2020; Accepted: 3 August 2020

Published online: 02 September 2020

References

1. Shanahan, T. M. *et al.* The time-transgressive termination of the African Humid Period. *Nat. Geosci* **8**, 140–144 (2015).
2. Gasse, F. Hydrological changes in the African tropics since the Last Glacial Maximum. *Quatern. Sci. Rev.* **19**, 189–211 (2000).
3. Kuper, R. & Kröpelin, S. Climate-controlled holocene occupation in the Sahara: Motor of Africa's Evolution. *Science* **313**, 803–807 (2006).
4. Marshall, F. & Hildebrand, E. Cattle before crops: the beginnings of food production in Africa. *J. World Prehistory* **16**, 99–143 (2002).
5. MacHugh, D. E., Larson, G. & Orlando, L. Taming the past: ancient DNA and the study of animal domestication. *Annu. Rev. Anim. Biosci.* **5**, 329–351 (2017).
6. Vigne, J.-D. *et al.* Etat des connaissances archéozoologiques sur les débuts de l'élevage du mouton dans l'ancien monde. *Ethnozoot-ecnie* **91**, 11–19 (2012).
7. Churcher, C. S., Kleindienst, M. R., Wiseman, M. F. & McDonald, M. M. The Quaternary faunas of Dakhleh Oasis, Western Desert of Egypt. In *The Oasis Papers 2: Proceedings of the Second International Conference of the Dakhleh Oasis Project 1–24* (Oxbow Books, 2008).
8. Klein, R. G. & Scott, K. Re-analysis of faunal assemblages from the Haua Fteah and other late quaternary archaeological sites in Cyrenaican Libya. *J. Archaeol. Sci.* **13**, 515–542 (1986).
9. Vermeersch, P. M., Van Peer, P., Moeyersons, J. & Van Neer, W. Sodmein Cave Site, Red Sea Mountains (Egypt). *Sahara* **6**, 31–40 (1994).
10. Gautier, A. The Early to Late Neolithic Archeofaunas from Nabta and Bir Kiseiba. In *Holocene Settlement of the Egyptian Sahara: Volume 1: The Archaeology of Nabta Playa* (eds Wendorf, F. & Schild, R.) 609–635 (Springer US, New York, 2001). https://doi.org/10.1007/978-1-4615-0653-9_23.
11. Close, A. E. Sinai, Sahara, Sahel: The introduction of domestic caprines to Africa. *Jennerstrasse* **8**, 459–469 (2002).
12. Kindermann, K. *et al.* Palaeoenvironment and Holocene land use of Djara, Western Desert of Egypt. *Quatern. Sci. Rev.* **25**, 1619–1637 (2006).
13. Lesur, J. *Et la gazelle devint chèvre: pré-histoires africaines d'hommes et d'animaux*. In *Collections « Sites et cités d'Afrique »* (eds Bon, F. & Fauvelle, F.-X.) (Presses Universitaires du Midi / Publications Scientifiques du Muséum, 2017).
14. Sadr, K. A short history of early herding in southern Africa. In *Pastoralism in Africa: Past, Present and Future* (eds Bollig, M., Schnegg, M. & Wotzka, H.-P.) 171–197 (Berghahn Books, New York and Oxford, 2013).
15. Pleurdeau, D. *et al.* "Of Sheep and Men": Earliest Direct Evidence of Caprine Domestication in Southern Africa at Leopard Cave (Erongo, Namibia). *PLoS ONE* **7**, e40340 (2012).
16. Le Meillour, L. *et al.* Identification of degraded bone and tooth splinters from arid environments using palaeoproteomics. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **11**, 472–482 (2018).
17. Horsburgh, K. A. & Moreno-Mayar, J. V. Molecular identification of sheep at Blydefontein Rock Shelter, South Africa. *S. Afr. Human.* **27**, 65–80 (2015).
18. Lander, F. & Russell, T. The archaeological evidence for the appearance of pastoralism and farming in southern Africa. *PLoS ONE* **13**, e0198941 (2018).
19. Sadr, K. Livestock first reached Southern Africa in two separate events. *PLoS ONE* **10**, e0134215 (2015).
20. Sadr, K. *Radiocarbon Dates, Stone Tools and the Origin of Herding on the West Coast of South Africa* (Africa Magna Verlag, Frankfurt am Main, 2014).
21. Smith, A. B., Sadr, K., Gribble, J. & Yates, R. Excavations in the South-Western Cape, South Africa, and the archaeological identity of prehistoric hunter-gatherers within the last 2000 years. *S. Afr. Archaeol. Bull.* **46**, 71–91 (1991).
22. Semino, O., Santachiara-Benerecetti, A. S., Falaschi, F., Cavalli-Sforza, L. L. & Underhill, P. A. Ethiopians and khoisan share the deepest clades of the human Y-chromosome phylogeny. *Am. J. Hum. Genet.* **70**, 265–268 (2002).
23. Schlebusch, C. M. & Jakobsson, M. Tales of human migration, admixture, and selection in Africa. *Annu. Rev. Genomics Hum. Genet.* **19**, 405–428 (2018).
24. Vicente, M., Jakobsson, M., Ebbesen, P. & Schlebusch, C. M. Genetic affinities among Southern Africa hunter-gatherers and the impact of admixing farmer and herder populations. *Mol. Biol. Evol.* **36**, 1849–1861 (2019).
25. Montinaro, F. & Capelli, C. The evolutionary history of Southern Africa. *Curr. Opin. Genet. Dev.* **53**, 157–164 (2018).
26. Heine, B. & König, C. What can linguistics tell us about early Khoekhoe history?. *S. Afr. Human.* **20**, 235–248 (2008).
27. Haacke, W. H. G. Linguistic hypotheses on the origin of Namibian Khoekhoe speakers. *S. Afr. Human.* **20**, 163–177 (2008).
28. Guldemann, T. A linguist's view : Khoe-Kwadi speakers as the earliest food-producers of southern Africa. <https://www.ingentaconnect.com/content/sabinet/nmsasah/2008/00000020/00000001/art00007> (2008).
29. Guillemard, I. Equating language, genes and subsistence? Discussion on the appearance of herding in southern Africa. *Azania Archaeol. Res. Afr.* **55**, 97–120 (2020).
30. Schlebusch, C. M., Lombard, M. & Soodyall, H. MtDNA control region variation affirms diversity and deep sub-structure in populations from southern Africa. *BMC Evol. Biol.* **13**, 56 (2013).
31. Breton, G. *et al.* Lactase persistence alleles reveal partial East African Ancestry of Southern African Khoe Pastoralists. *Curr. Biol.* **24**, 852–858 (2014).
32. Ranciaro, A. *et al.* Genetic origins of lactase persistence and the spread of pastoralism in Africa. *Am. J. Hum. Genet.* **94**, 496–510 (2014).
33. Ehret, C. The first spread of food production to southern Africa. In *The Archaeological and Linguistic Reconstruction of African History* (eds Ehret, C. & Posnansky, M.) 158–182 (University of California Press, Berkeley, 1982).

34. Blench, R. Was there an interchange between Cushitic pastoralists and Khoisan speakers in the prehistory of Southern Africa and how can this be detected? Presented at Königswinter (2007).
35. Muigai, A. W. T. & Hanotte, O. The origin of African sheep: archaeological and genetic perspectives. *Afr. Archaeol. Rev.* **30**, 39–50 (2013).
36. Soma, P., Kotze, A., Grobler, J. P. & van Wyk, J. B. South African sheep breeds: population genetic structure and conservation implications. *Small Ruminant Res.* **103**, 112–119 (2012).
37. Prendergast, M. E. Hunters and herders at the periphery: The spread of herding in eastern Africa. In *People and Animals in Holocene Africa: Recent Advances in Archaeozoology*. **2**, 43 (eds Jousse, H. & Lesur, J.) (2011).
38. Sadr, K. The first herders at the Cape of Good Hope. *Afr. Archaeol. Rev.* **15**, 101–132 (1998).
39. Smith, A. B. Early Herders in Southern Africa: a synthesis. In *Animals and People: Archaeozoological Papers in Honour of Ina Plug* (eds Badenhorst, S., Mitchell, P. & Driver, J.) 94–103 (Archaeopress, 2008).
40. Kinahan, J. The origins and spread of pastoralism in Southern Africa. *Oxf. Res. Encycl. Afr. Hist.* <https://doi.org/10.1093/acref/ore/9780190277734.013.678> (2019).
41. Smith, A. B. & Jacobson, L. Excavations at Geduld and the appearance of early domestic stock in Namibia. *S. Afr. Archaeol. Bull.* **50**, 3–14 (1995).
42. Sandelowsky, B. Mirabib—an archaeological study in the Namib. *Modoqua* **10**, 221–283 (1977).
43. Wadley, L. Big Elephant Shelter and its role in the Holocene prehistory of central South West Africa. *Cimbebasia* **3**, 2–75 (1979).
44. Albrecht, M. *et al.* Oruwanje 95/1: a late Holocene stratigraphy in northwestern Namibia. *Cimbebasia* **17**, 1–22 (2001).
45. Sandelowsky, B. *Archaeologically yours, Beatrice Sandelowsky: A personal journey into the prehistory of Southern Africa, in particular Namibia* (Namibia Scientific Society, Windhoek, 2004).
46. Imalwa, E. *Exploitation of animals during Holocene in Namibia. Examples of three archaeological from Erongo mountains and central Namib desert* (Muséum national d'Histoire naturelle, Paris, 2009).
47. Mauraan, G. *et al.* First in situ pXRF analyses of rock paintings in Erongo, Namibia: results, current limits, and prospects. *Archaeol. Anthropol. Sci.* <https://doi.org/10.1007/s12520-019-00787-7> (2019).
48. Scott, K. & Plug, I. Osteomorphology and osteometry versus aDNA in taxonomic identification of fragmentary sheep and sheep/goat bones from archaeological deposits: Blydefontein Shelter, Karoo. *South Africa. S. Afr. Human.* **21**, 61–79 (2016).
49. Horsburgh, K. A., Orton, J. & Klein, R. G. Beware the Springbok in sheep's clothing: how secure are the faunal identifications upon which we build our models?. *Afr. Archaeol. Rev.* **33**, 353–361 (2016).
50. Horsburgh, K. A., Moreno-Mayar, J. V. & Klein, R. G. Counting and miscounting sheep: genetic evidence for pervasive misclassification of wild fauna as domestic stock. *S. Afr. Human.* **30**, 53–69 (2017).
51. Horsburgh, K. A. A reply to Plug 2017: science requires self-correction. *Azania Archaeol. Res. Afr.* **53**, 114–118 (2018).
52. Plug, I. Reply to Horsburgh *et al.* 2016: 'Revisiting the Kalahari debate in the highlands'. *Azania Archaeol. Res. Afr.* **53**, 98–113 (2018).
53. Scott, K. & Plug, I. Beware of the sheep in eland's clothing: A critique of 'springbok in sheep's clothing', by Horsburgh *et al.* *S. Afr. Archaeol. Bull.* **74**, 25 (2019).
54. Salvagno, L. & Albarella, U. A morphometric system to distinguish sheep and goat postcranial bones. *PLoS ONE* **12**, e0178543 (2017).
55. Kinahan, J. Archaeological evidence of domestic sheep in the Namib desert during the first millennium AD. *J. Afr. Archaeol.* **14**, 7–17 (2016).
56. Demarchi, B. *et al.* Protein sequences bound to mineral surfaces persist into deep time. *eLife Sci.* **5**, e17092 (2016).
57. Prendergast, M. E., Janzen, A., Buckley, M. & Grillo, K. M. Sorting the sheep from the goats in the Pastoral Neolithic: morphological and biomolecular approaches at Luxmanda, Tanzania. *Archaeol. Anthropol. Sci.* <https://doi.org/10.1007/s12520-018-0737-0> (2018).
58. Buckley, M. Zooarchaeology by mass spectrometry (ZooMS) collagen fingerprinting for the species identification of archaeological bone fragments. In *Zooarchaeology in Practice* (eds Giovas, C. & Lefebvre, M.) 227–247 (Springer, Cham, 2018). https://doi.org/10.1007/978-3-319-64763-0_12.
59. Brown, S. *et al.* Identification of a new hominin bone from Denisova Cave, Siberia using collagen fingerprinting and mitochondrial DNA analysis. *Sci. Rep.* **6**, 23559 (2016).
60. Hendy, J. *et al.* A guide to ancient protein studies. *Nat. Ecol. Evol.* **2**, 791 (2018).
61. Chen, F. *et al.* A late Middle Pleistocene Denisovan mandible from the Tibetan Plateau. *Nature* <https://doi.org/10.1038/s41586-019-1139-x> (2019).
62. Welker, F. *et al.* Middle Pleistocene protein sequences from the rhinoceros genus *Stephanorhinus* and the phylogeny of extant and extinct Middle/Late Pleistocene Rhinocerotidae. *PeerJ* **5**, e3033 (2017).
63. Welker, F. *et al.* Ancient proteins resolve the evolutionary history of Darwin's South American ungulates. *Nature* **522**, 81–84 (2015).
64. Welker, F. *et al.* Enamel proteome shows that Gigantopithecus was an early diverging pongine. *Nature* <https://doi.org/10.1038/s41586-019-1728-8> (2019).
65. Cappellini, E. *et al.* Early Pleistocene enamel proteome from Dmanisi resolves *Stephanorhinus* phylogeny. *Nature* **574**, 103–107 (2019).
66. Parker, G. J. *et al.* Sex estimation using sexually dimorphic amelogenin protein fragments in human enamel. *J. Archaeol. Sci.* <https://doi.org/10.1016/j.jas.2018.08.011> (2018).
67. Lugli, F. *et al.* Enamel peptides reveal the sex of the Late Antique 'Lovers of Modena'. *Sci. Rep.* **9**, 1–8 (2019).
68. Hendy, J. *et al.* The challenge of identifying tuberculosis proteins in archaeological tissues. *J. Archaeol. Sci.* **66**, 146–153 (2016).
69. Buckley, M. *et al.* Distinguishing between archaeological sheep and goat bones using a single collagen peptide. *J. Archaeol. Sci.* **37**, 13–20 (2010).
70. Buckley, M. & Kansa, S. W. Collagen fingerprinting of archaeological bone and teeth remains from Domuztepe, South Eastern Turkey. *Archaeol. Anthropol. Sci.* **3**, 271–280 (2011).
71. Gershon, P. D. Cleaved and missed sites for trypsin, Lys-C, and Lys-N can be predicted with high confidence on the basis of sequence context. *J. Proteome Res.* **13**, 702–709 (2014).
72. Lebon, M., Reiche, I., Gallet, X., Bellot-Gurlet, L. & Zazzo, A. Rapid quantification of bone collagen content by ATR-FTIR spectroscopy. *Radiocarbon* **10**, 1–15 (2016).
73. Bronk Ramsey, C. *OxCal Program, Version 4.3.* (Oxford Radiocarbon Accelerator Unit: University of Oxford, 2017).
74. Hogg, A. G. *et al.* SHCal13 Southern Hemisphere Calibration, 0–50,000 Years cal BP. *Radiocarbon* **55**, 1889–1903 (2013).
75. Tyson, P. D. & Lindesay, J. A. The climate of the last 2000 years in southern Africa. *Holocene* **2**, 271–278 (1992).
76. Tyson, P., Karlen, W., Holmgren, K. & Heiss, G. The Little Ice Age and medieval warming in South Africa. *S. Afr. J. Sci.* **96**, 121–126 (2000).
77. Heine, K. Little Ice Age climatic fluctuations in the Namib Desert, Namibia, and adjacent areas: Evidence of exceptionally large floods from slack water deposits and desert soil sequences. In *Paleoecology of Quaternary Drylands* (eds Smytkatz-Kloss, W. & Felix-Henningsen, P.) 137–165 (Springer, Berlin, 2004). https://doi.org/10.1007/978-3-540-44930-0_9.
78. Sandelowsky, B. H. Archaeological investigations at Mirabib Hill rock shelter. *Goodwin Ser.* **2**, 65–72 (1974).
79. Pálsdóttir, A. H., Bläuer, A., Rannamäe, E., Boessenkool, S. & Hallsson, J. H. Not a limitless resource: ethics and guidelines for destructive sampling of archaeofaunal remains. *R. Soc. Open Sci.* **6**, 191059 (2019).

80. Hendy, J. *et al.* Ancient proteins from ceramic vessels at Çatalhöyük West reveal the hidden cuisine of early farmers. *Nat. Commun.* **9**, 4064 (2018).
81. Colonese, A. C. *et al.* New criteria for the molecular identification of cereal grains associated with archaeological artefacts. *Sci. Rep.* **7**, 6633 (2017).
82. Greco, E. *et al.* Proteomic analyses on an ancient Egyptian cheese and biomolecular evidence of brucellosis. *Anal. Chem.* **90**, 9673–9676 (2018).
83. Warinner, C. *et al.* Direct evidence of milk consumption from ancient human dental calculus. *Sci. Rep.* **4**, 7104 (2014).
84. Buckley, S., Usai, D., Jakob, T., Radini, A. & Hardy, K. Dental calculus reveals unique insights into food items, cooking and plant processing in prehistoric central sudan. *PLoS ONE* **9**, e100808 (2014).
85. Balasse, M. & Tresset, A. Environmental constraints on the reproductive activity of domestic sheep and cattle: what latitude for the herder?. *Anthropozoologica* **42**(2), 71–88 (2007).
86. Zhang, J. *et al.* PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. *Mol. Cell Proteom.* **11**, M111.010587 (2012).
87. Kearse, M. *et al.* Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649 (2012).
88. Perkins, D. N., Pappin, D. J., Creasy, D. M. & Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567 (1999).
89. Perez-Riverol, Y. *et al.* The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucl. Acids Res.* **47**, D442–D450 (2019).

Acknowledgements

This work was funded by the Agence Nationale de la Recherche under the LabEx ANR-10-LABX-0003-BCD, in the program 'Investissement d'avenir' ANR-11-IDEX-0004-02. Molecular experiments were performed thanks to an Action Thématique Muséum grant under projects ProtéArch (2016) and TopArch (2017) and the 'Project Exploratoire Premier Soutien' (PEPS) D2CAP from the CNRS. This work is part of LLM's PhD thesis, supported by a grant from the DIM MAP of the Région Île-de-France (project 'Introduction et diffusion du petit bétail en Afrique : approches moléculaires' [IBAAM]) and the Muséum national d'Histoire naturelle. The fieldwork was co-funded by the French ministry of foreign and European affairs (Ministère des Affaires étrangères et du Développement international), the National Museum of Namibia and the National Heritage Council of Namibia. The authors are grateful to the Rust family, owners of Omandumba West Farm, Namibia, for their support; to the staff of the comparative anatomy collections of the Muséum national d'Histoire naturelle (Paris, France) for providing access to modern specimens of springbok and impala; and to F. Thil and N. Tisnerat-Laborde (Laboratoire des Sciences du Climat et de l'Environnement, Saclay, France) for graphitisation, accelerator mass spectrometry measurements and data reduction of the collagen samples. We are grateful to the Plateforme de Spectrométrie de Masse Bio-Organique, the Laboratoire de Datation par le Radiocarbone, and the Plateforme de Spectrométrie Infrarouge (especially X. Gallet) of the MNHN, and to the Laboratoire des Sciences du Climat et de l'Environnement for access to the instruments. LLM, FD and DP are also grateful to PIL and TW for their help on early versions of this manuscript. The authors would like to thank S. Needs-Howarth for English editing of the manuscript.

Author contributions

Conceived and designed the experiments: L.L.M., S.Z., A.Z., S.C., M.L., D.P., J.L. Contributed to field work: L.L.M., F.D., E.I., M.L., A.N., D.P., J.L. Performed ATR-FTIR experiments: L.L.M., M.L. Performed proteomics experiments: L.L.M., S.Z. Performed radiocarbon experiments: O.T., A.Z. Analysed data: L.L.M., S.Z., A.Z., M.L., J.L. Contributed reagents, materials, analysis tools: S.Z., A.Z., S.C., F.D., M.L., D.P., J.L. L.L.M., S.Z., A.Z., D.P. and J.L. wrote the early draft of this manuscript with input from all the authors. All authors wrote parts of the paper and agree to be accountable for the content of the work.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-71374-3>.

Correspondence and requests for materials should be addressed to L.L.M. or J.L.

Reprints and permissions information is available at www.nature.com/reprints.

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



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

© The Author(s) 2020