Exceptional ancient DNA preservation and fibre remains of a Sasanian saltmine sheep mummy in Chehrābād, Iran

Conor Rossi, Gabriela Russ-Popa, Valeria Mattiangeli, Fionnuala Mcdaid, Andrew Hare, Hossein Davoudi, Haeedeh Laleh, Zahra Lorzadeh, Roya Khazaeli, Homa Fathi, et al.

BIOLOGY LETTERS

royalsocietypublishing.org/journal/rsbl

Research

Cite this article: Rossi C *et al*. 2021 Exceptional ancient DNA preservation and fibre remains of a Sasanian saltmine sheep mummy in Chehrābād, Iran. *Biol. Lett.* **17**: 20210222. https://doi.org/10.1098/rsbl.2021.0222

Received: 29 April 2021
Accepted: 21 June 2021

Subject Areas:
evolution

Keywords:
ancient DNA, sheep, mummy

Author for correspondence:
Kevin G. Daly
e-mail: dalyk1@tcd.ie

Population genetics

Exceptional ancient DNA preservation and fibre remains of a Sasanian saltmine sheep mummy in Chehrābād, Iran

Conor Rossi[1], Gabriela Ruß-Popa[2], Valeria Mattiangeli[1], Fionnuala McDaid[1], Andrew J. Hare[1], Hossein Davoudi[3], Haeedeh Laleh[3,4], Zahra Lorzadeh[3], Roya Khazaeli[3], Homa Fathi[3], Matthew D. Teasdale[5], Abolfazl A'ali[6], Thomas Stöllner[7], Marjan Mashkour[3,8] and Kevin G. Daly[1]

[1]Smurfit Institute of Genetics, Trinity College Dublin, Dublin 2, D02 VF25, Ireland
[2]Austrian Academy of Sciences, Austrian Archaeological Institute, Archaeological Sciences, Hollandstraße 11-13, 1020 Vienna, Austria
[3]Central Laboratory, Bioarchaeology Laboratory, University of Tehran, 1417634934 Tehran, Iran
[4]Faculty of Humanities, Department of Archaeology, University of Tehran, 1417935840 Tehran, Iran
[5]McDonald Institute for Archaeological Research, Dept. of Archaeology, University of Cambridge, Cambridge CB2 3ER, UK
[6]Zanjan Cultural Heritage Centre, Archaeological Museum of Zanjan, Emaarate Zolfaghari, Taleghani St., Zanjan, Iran
[7]Research Department, Haus der Archäologien, Ruhr University Bochum, Institute for Archaeological Studies and Deutsches Bergbau-Museum Bochum, Am Bergbaumuseum 31, D-44791 Bochum, Germany
[8]Archéozoologie, Archéobotanique, Sociétés, Pratiques et Environnements (AASPE), Muséum national d'Histoire naturelle, Sorbonne Université, CNRS, CP 56, 55 rue Buffon, 75005 Paris, France

CR, 0000-0003-4561-8878; GR-P, 0000-0001-8492-6238; VM, 0000-0001-9785-1714; AJH, 0000-0001-8595-6965; HD, 0000-0002-5236-1444; HF, 0000-0002-1764-0130; MDT, 0000-0002-7376-9975; AA, 0000-0002-6621-8665; MM, 0000-0003-3630-9459; KGD, 0000-0002-5579-6144

Mummified remains have long attracted interest as a potential source of ancient DNA. However, mummification is a rare process that requires an anhydrous environment to rapidly dehydrate and preserve tissue before complete decomposition occurs. We present the whole-genome sequences (3.94 X) of an approximately 1600-year-old naturally mummified sheep recovered from Chehrābād, a salt mine in northwestern Iran. Comparative analyses of published ancient sequences revealed the remarkable DNA integrity of this mummy. Hallmarks of postmortem damage, fragmentation and hydrolytic deamination are substantially reduced, likely owing to the high salinity of this taphonomic environment. Metagenomic analyses reflect the profound influence of high-salt content on decomposition; its microbial profile is predominated by halophilic archaea and bacteria, possibly contributing to the remarkable preservation of the sample. Applying population genomic analyses, we find clustering of this sheep with Southwest Asian modern breeds, suggesting ancestry continuity. Genotyping of a locus influencing the woolly phenotype showed the presence of an ancestral 'hairy' allele, consistent with hair fibre imaging. This, along with derived alleles associated with the fat-tail phenotype, provides genetic evidence that Sasanian-period Iranians maintained specialized sheep flocks for different uses, with the 'hairy', 'fat-tailed'-genotyped sheep likely kept by the rural community of Chehrābād's miners.

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.5486259.

THE ROYAL SOCIETY PUBLISHING

1. Introduction

In 1993, a remarkably preserved human body dating to approximately 1700 years Before Present (BP) was discovered in the Douzlākh salt mine near Chehrābād village in the Zanjan Province of northwest Iran [1–3]. A total of eight 'Salt Men' have been identified at the mine [4,5], several retaining keratinous tissues such as skin, hair and both endo- and exoparasites, despite dating to the Achaemenid (550–330 BCE, 2500–2280 BP) and Sasanian (224–651 CE, approximately 1700–1300 BP) periods. The mine, also known as Chehrābād, was active in various periods, and its archaeological refilling layers represent an extraction history that ranged from the sixth century BCE to twentieth-century CE. In addition to the 'Salt Men', textiles, leather objects and animal remains have been discovered [6,7], likely preserved by the high salinity and low moisture content of the mine. Isotopic, genetic and lipid analyses have been reported for this material [1], and studies have been carried out to characterize genomic DNA survival [8]. These human and animal remains are examples of natural mummification—the spontaneous desiccation of soft tissue by a dry environment that rapidly dehydrates soft tissue before decay begins [9].

Mummification has been suggested as a mechanism that may sufficiently preserve keratinized tissue for ancient DNA (aDNA) sequencing [9]. The effects of age-related damage in aDNA are well documented and include base misincorporation at strand overhangs, fragmentation and low endogenous content [10]. Both deamination and depurination, associated with postmortem transition error and DNA fragmentation, respectively, require water as a substrate [10]. Ancient DNA from Chehrābād, a highly saline, anhydrous environment, presents an opportunity to investigate potential differences in nucleotide degradation resulting from this unusual taphonomic context.

In this study, we sequenced DNA from the approximately 1600-year-old (Sasanian period) mummified sheep leg 4305, recently discovered in a large mining gallery in the northwestern edge of the Douzlākh saltmine of Chehrābād by Iranian–German researchers during archaeological excavations (figure 1*a*) [2]. The specimen was likely deposited during refilling activities in the fourth–fifth centuries CE after the gallery's reopening in the Early Sasanian period (second–third centuries CE) and following its initial collapse between 405 and 380 BCE. The leg was possibly discarded during food preparation activities, as both sheep and goat were likely used as provisioning for Sasanian-period miners; equines may have been used as beasts of burden [16]. By this time, sheep were an established commodity for their meat and secondary products such as wool fibre, which was widespread by the fourth millennium BCE and showed regional specialization by the third millennium BCE [17].

We find unusual survival patterns of endogenous DNA given its distance from the equator, implying that exceptional preservation of nucleic acid integrity was afforded by the unique salt-rich environment. This enables characterization of the mummy skin metagenome and population genomic profiling of this sheep in the context of modern breeds. We also genotype 51 SNPs within the first intron of the platelet-derived growth factor D (*PDGFD*) that are highly differentiated between fat-tailed and thin-tailed breeds [18]. Finally we genotype the antisense *EIF2S2* retrogene insertion within the 3′ UTR of the *IRF2BP2* gene that influences the woolly phenotype and is derived relative to the ancestral coarse 'hairy' coat [19], in tandem with fibre analysis using scanning electron microscopy (SEM).

2. Material and methods

A sample of the mummified sheep skin (table 1, MUM2) from sheep leg 4305 was directly radiocarbon dated at the ^{14}CHRONO Centre (Queen's University Belfast). OxCal 4.3.2 [20] was used to calibrate its age (95.4% confidence interval) using [21].

Sample preparation, extraction and library preparation were performed in a dedicated aDNA laboratory in the Smurfit Institute of Genetics, Trinity College Dublin according to standard protocols (see electronic supplementary material). Sequencing of MUM2 and two Iranian sheep bone samples (Khor1 and Azer2) of approximately similar ages (table 1) for comparison was performed on Illumina MiSeq (50 bp SE) and HiSeq 2500 platforms (100 bp SE and 100 bp paired-end (PE)).

Sequencing reads were aligned to OviAri3.1 and filtered to produce bam files following standard aDNA sequencing pipelines (electronic supplementary material). Damage patterns were assessed using mapDamage2.0 [22].

Filtered reads not aligned to either sheep or human genomes were taxonomically assigned using the metagenomic classifier Kraken 2 [23]. Microbial sources were estimated using SourceTracker2 [24] with a custom metagenomic database [25–29] (electronic supplementary material). Bacterial species abundances were generated using MIDAS [30].

Mitochondrial sequences were produced using ANGSD [31] and a maximum-likelihood phylogenetic tree was generated using SeaView and PhyML [32–34] with the HKY85 substitution model, selected using jModelTest2 [35,36] and 100 bootstrap repeats.

A SNP dataset of modern breeds [37] was used to investigate genomic affinities (electronic supplementary material). LASER (v. 2.03) principal components analysis (PCA) [38], outgroup f_3 statistics [39], TreeMix [40] and ADMIXTURE [41] analyses were completed (electronic supplementary material, table S5).

We investigated the woolly locus located on chromosome 25 [19]. Two modified OviAri3.1 assemblies were produced, one representing the ancestral 'hairy' phenotype, the other representing the 'woolly' phenotype (electronic supplementary material). Final bam files were visualized using IGV [42]. Hair fibres were examined using scanning electron and light microscopes at USTEM, TU Wien and Austrian Archaeological Institute, respectively. We assessed 51 SNPs in the *PDGFD* gene associated with the derived fat-tail phenotype [18], using the genotype calls of modern fat and thin-tailed breeds to define the derived allele [43,44]. As the average genome coverage was too low for accurate diploid genotype calls, we report base calls for both alleles.

3. Results and discussion

(a) DNA preservation and metagenomics

The Chehrābād mummy sample (MUM2) was directly dated to the fifth–sixth century CE (2 sigma 1621–1481 cal BP, uncalibrated 1600 ± 30 BP; electronic supplementary material, figure S3). This aligns with the Sasanian Empire period of Iran, a time when the mine was in active use [1]. Initial DNA screening indicated high endogenous DNA for MUM2, and also the comparative Iranian sheep samples from relatively close time periods (table 1).

Sequencing of the Chehrābād mummy produced a 3.94 X genome after quality filtering (electronic supplementary material, table S1), in addition to the low coverage comparative genomes (0.0 4 X and 0.07 X). MUM2 differs from the two comparative sheep samples in displaying longer fragment lengths (median 107 bp versus 52 bp and 56 bp; figure 1*b*; collapsed reads-only 90 bp versus 50 bp and 55 bp) and substantially lower rates of deamination (figure 1*d*) (δS, single-strand cytosine deamination probability, mean $\delta S = 0.012$ versus 0.382

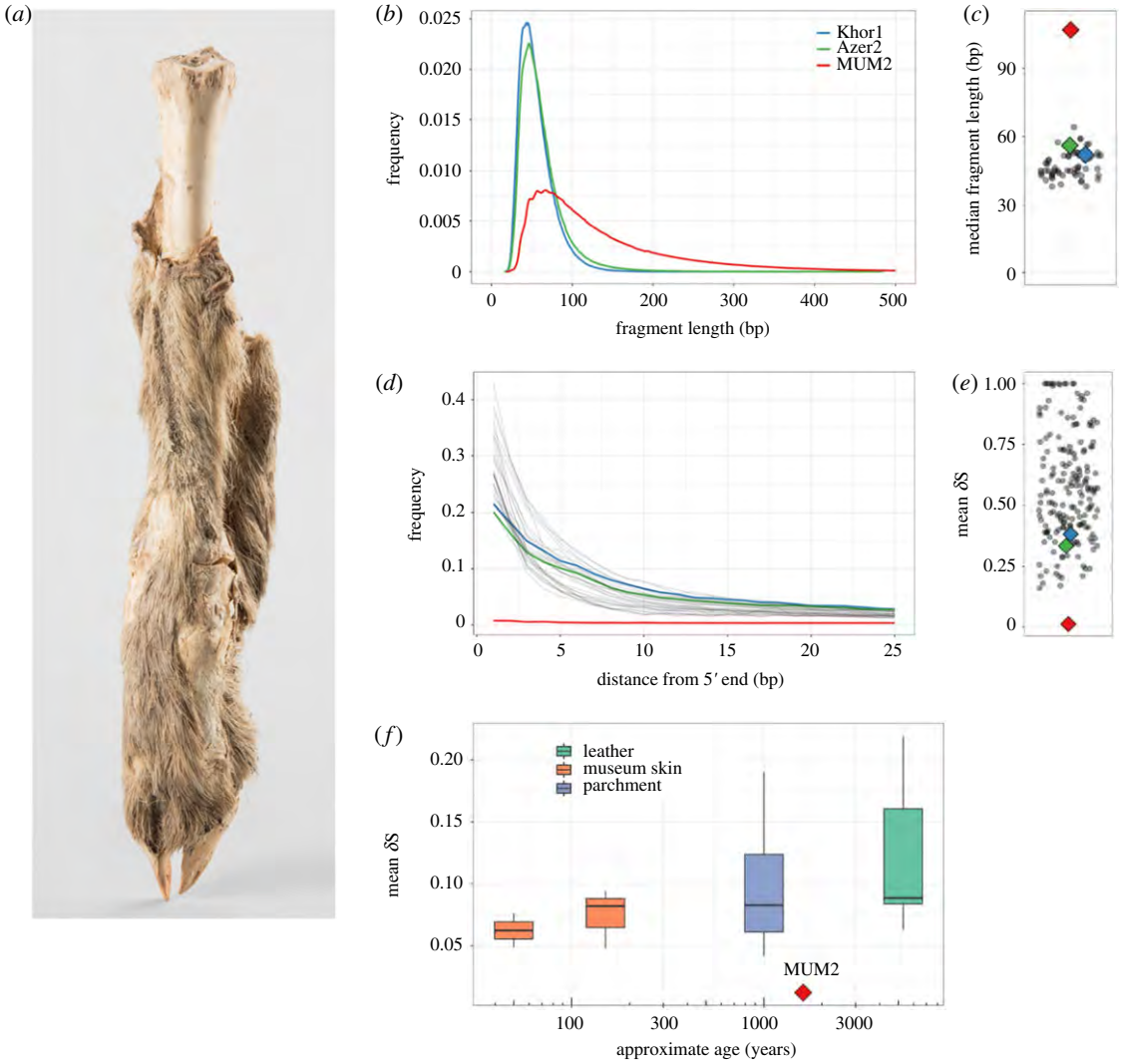


Figure 1. (*a*) Mummified sheep leg (4305) after cleaning. Photography: N. Tehrani. (*b*) Read length distributions of MUM2, Khor1 and Azer2, calculated from PE data. MUM2 shows a reduced rate of fragmentation. The median read length of MUM2 (107 bp) exceeds the median read length of Khor1 and Azer2 (52 bp and 56 bp, respectively). (*c*) Median read lengths of 61 published ancient Ovicaprid samples [11]. The median read length of MUM2 (107 bp; 90 bp among collapsed reads only) exceeds the longest among published ovicaprid genomes (64 bp). (*d*) Deamination patterns of MUM2, Khor1, Azer2 and other ancient ovicaprids for non UDG-treated libraries. Low levels of base misincorporation at the 5′ ends of reads were observed for MUM2 compared to Khor1 and Azer2. (*e*) Mean δS of 182 published ancient bone samples [11,12]. The mean δS of MUM2 (0.012) is unusual in its low levels of deamination. (*f*) Comparison of mean δS of published ancient skins [13–15]. Lower damage rates are recorded compared to all samples, including some of approximately 50 years old.

Table 1. Summary information of samples sequenced in this study.

name	tissue	origin	period	age	endogenous DNA %	coverage (X)
MUM2	mummified skin	Chehrābād, Iran	Sasanian Empire period	399–539 cal CE[a]	31.01	3.94
Khor1	petrous bone	Nishapur, Iran	Sasanian–Islamic periods	600–1200 CE	58.44	0.04
Azer2	petrous bone	Tepe Hasanlu, Iran	Iron Age III	800–600 BCE	31.32	0.07

[a]directly dated.

and 0.334). Contrasting previously published ancient ovicaprid data from Southwest Asia and Europe (electronic supplementary material, table S3), MUM2 falls outside the ranges of both median fragment length and mean δS values (figure 1*c*, *e*), indicating remarkably low fragmentation and deamination of the Chehrābād sheep mummy genomic material given its latitude. Similar length distributions have been reported primarily from high latitude and permafrost environments [45–48]. A low level of thermal fluctuations may also contribute to DNA preservation [12], as comparable fragment lengths have been reported in a human sample from Wezmeh Cave, Iran [49].

Recent models of postmortem DNA fragmentation suggest rate-constant hydrolytic depurination over time [50], or age-independence, driven by environment-dependent biotic and abiotic factors [12]. The depurination rates of MUM2 are similar to the more-fragmented comparative samples (electronic supplementary material, figure S4),

Table 2. Predicted source proportion of metagenomic reads by SourceTracker2.

	MUM2 (species)	Azer2 (species)	Khor1 (species)	MUM2 (genus)	Azer2 (genus)	Khor1 (genus)
tissue decomposers	0.0212	0.1458	0.0386	0.0426	0.3524	0.3727
salt-rich	0.4725	0.0026	0.0036	0.7458	0.0145	0.0151
laboratory reagents	0.0003	0.0002	0.0002	N/A	N/A	N/A
sheep skin	0.0285	0.0850	0.0463	0.0829	0.3294	0.2244
soil	0.0009	0.0007	0.0046	0	0	0
unknown	0.4766	0.7657	0.9067	0.1287	0.3037	0.3878

implying that other processes in the Chehrābād environment underlie the lower fragmentation rates. The highly alkaline, cool and anhydrous conditions may have contributed to the inhibition of cellular nucleases that would otherwise degrade and fragment endogenous DNA [9]. Postmortem DNA deamination via cytosine hydrolysis [51] is thought to be strongly correlated with age [52] and thermal age [12]. The substantially lower rates of deamination observed in MUM2 are likely owing to the scarcity of environmental free water, required for hydrolytic deamination. These results are consistent with Chehrābād providing a taphonomic environment conducive to genome preservation.

DNA preservation may also be influenced by its tissue-of-origin; for example, bone hydroxyapatite rather than keratin fractions is associated with smaller fragment size [53]. As hydrophobic keratinized tissue may provide resistance to environmental water [54], we compared MUM2 to published genomes of ancient skins (figure 1*f*) to determine if tissue providence was solely responsible for DNA preservation. The mean δS of MUM2 falls outside the range of other ancient skin genomes, including twentieth-century CE goat skins [13] and leather recovered from the Tyrolean Iceman [14]. While this does not discount keratinized tissue being specifically enriched with longer DNA fragments, the Chehrābād sheep mummy appears to be singular in its DNA integrity among published skin samples.

Given the distinctive geochemical composition of Chehrābād, we examined if its salt-rich environment was reflected in the metagenomic profile of MUM2. Taxonomic assignment and abundance estimation assigned 57.13% of classified reads to the halophilic Class of Archaea *Halobacteria* (electronic supplementary material, table S2). Similarly, SourceTracker2 predicted that 0.4725–0.7458 of the microbial community originated from a salt-rich environment (table 2; electronic supplementary material, figure S5). A complementary analysis using MIDAS identified 76 unique bacterial species in the mummified sheep (electronic supplementary material, table S4). The most abundant species is the halophilic bacterium *Actinopolyspora halophila 58532*, accounting for approximately 29% of identified reads. This signal of a dominant halophilic microbial community is not replicated in comparison samples or controls (table 2; electronic supplementary material). Rapid colonization by saprophytic microbial communities, with key decomposers being ubiquitous across soil types, is typical for mammalian corpses postmortem [55]. The halophilic metagenome profile observed in the Chehrābād sheep mummy skin indicates that the typical decomposers may be less abundant in this alkaline, salt-rich setting, which may have contributed to soft tissue and molecular preservation.

(b) Population genomics

We investigated how the Chehrābād sheep MUM2 relates to modern populations using a mitochondrial and autosomal variation. A 664 X mitochondrial genome of MUM2 falls within the C haplotype cluster in a maximum-likelihood phylogeny of modern sheep mitochondria (electronic supplementary material, figure S7). This clade is found at its highest frequency in southwest and east Asia [56,57], has been reported in ancient samples from Bronze Age Turkey [58] and is consistent with past and present-day patterns of mitochondrial diversity.

PCA from autosomal variation clusters MUM2 with modern southwest Asian breeds, using both global and Asian reference panels (electronic supplementary material, figure S8). f_3 outgroup statistics show that MUM2 shares the most genetic drift with southwest Asian breeds, particularly those from Iran (figure 2*a*). ADMIXTURE and TreeMix analysis also confirmed the affinity of MUM2 with modern sheep breeds from southwest Asia (electronic supplementary material). Overall, there is genetic continuity between west Iranian sheep populations in Sassanid and modern time periods, although PCA using Ovine SNP50 genotypes of Asian breeds places MUM2 apart from sampled breeds (electronic supplementary material, figure S9), suggesting a degree of genetic flux during the past 1500–1600 years in Iranian sheep. This is consistent with evidence for genetic exchange across Asia prior to the development of modern breeds [59–61].

(c) Fibre genotype and phenotype and fat-tail genotype

The derived 'woolly' coat phenotype is thought to be influenced by an approximately 1.5 kbp insertion of a *EIF2S2* retrogene into the *IRF2BP2* 3′ UTR, recessive to the ancestral allele associated with 'hairy' coat [19]. We exploited the length of the MUM2 DNA fragments to investigate this 'woolly' locus by searching for read pairs that either encompassed or overlapped the insertion breakpoint, indicative of a copy of the 'hairy' allele. No reads were found to overlap the diagnostic insertion breakpoints of the 'woolly' insert, which would indicate a copy of the 'woolly' allele (figure 2*b*). Five reads were found to uniquely map to the 'hairy' allele diagnostic position, with a further two read pairs inferred to overlap this breakpoint (figure 2*c*). We therefore infer this animal to be either homozygous or heterozygous for the dominant 'hairy' allele. In addition, SEM imaging of the mostly unpigmented mummified hair fibres revealed mosaic scales typical of sheep [62], with fine lines on the scale surface (figure 2*d*;

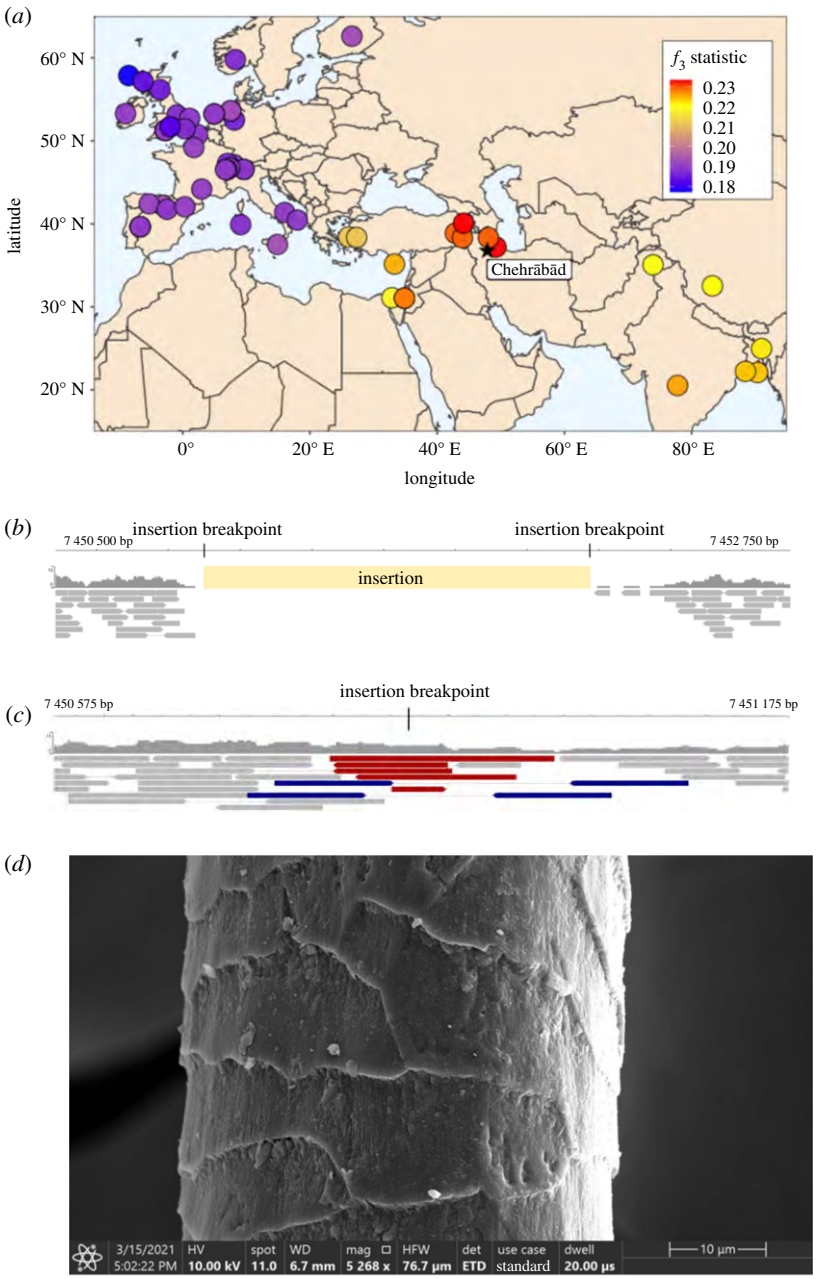


Figure 2. (*a*) Shared genetic drift between MUM2 and modern sheep populations. Higher f_3 values, in red, indicate higher shared drift, relative to the outgroup Asiatic mouflon. Visualization of read coverage of filtered bam files at woolly locus, in assemblies with (*b*) and without (*c*) the insertion. Reads highlighted in red overlap the insertion breakpoint, blue indicates an inferred overlap of the insertion point by the straddling read pair. Highlighted reads map only to one assembly and do not align to the other. (*d*) SEM image of MUM2 hair fibre, displaying the mosaic scales typical of a sheep hair shaft. Image by A. Steiger-Thirsfeld and G. Ruß-Popa.

electronic supplementary material), a characteristic of sheep hair fibres and particularly for mouflon and medium-wool breeds [63]. This may reflect MUM2 coming from a herd maintained for meat or milk production rather than wool, consistent with suggestions that ovicaprids were used as food for workers, and that sections of the mine were used as stables [1].

We also find evidence of a fat-tail associated allele (48/51 SNPs) (electronic supplementary material, table S6) at *PDGFD*, a gene likely controlling tail phenotype [18,44]. This observation, along with MUM2 sharing mitochondrial haplotype C with the majority of modern fat-tailed breeds [61], and the genomic affinity of MUM2 to modern fat-tailed breeds, although based on SNP-chip data, is intriguing. While we cannot determine the MUM2 tail phenotype directly, its genotype is similar to a medium-wool or hairy-coated fat-tail breed [1]. Hairy-coated sheep may have lower mortality rates, have higher birth weights, and be more robust than woolly coated [64], while fat-tail breeds are thought to be better adapted to arid environments [18]. If phenotypically similar to these sheep breeds, the flock represented by MUM2 could have provided a reliable meat and fat source for Chehrābād's miners. The faunal assemblage of the mine and paleoparasitological studies, although not very abundant, support the fact that sheep/goats were the most consumed animals by the miners [6,7,16,65].

Both woolly and fat-tailed sheep are depicted in the Early Bronze Age Mesopotamia but the spread of these phenotypes may have been uncoupled, and occurred via distinct processes [66,67]. Fat-tailed breeds were likely introduced from Southwest to East Asia in a period (700 BCE–1000 CE) broadly coinciding with the age of MUM2 [60]; the observed *PDGFD* genotype supports an ancient origin of this economically important trait. Wider aDNA analysis may elucidate when

woolly and fat-tailed associated genotypes arose and how they may have influenced sheep breed development, which have their origins in fourth millennium BCE Mesopotamia [66,67]. Although the archaeozoological assemblages in the Iranian Plateau from the Antiquity and later Mediaeval periods are still limited, the diversity of the size of sheep bones is already an indication of the diversification of breeds in these periods [65,68]. Our results are consistent with MUM2 deriving from a herd used for meat and/or milk rather than wool production, and reflect sophisticated Sasanian-period husbandry practices and specialized sheep breeding.

Data accessibility. Sequencing reads and mitochondrial sequences are available under European Nucleotide Archive accession PRJEB43881.

The data are provided in the electronic supplementary material [69].

Authors' contributions. C.R., K.G.D. and M.M. designed study; C.R., K.G.D., V.M., A.J.H. and F.M. performed laboratory work; C.R. and M.D.T. performed bioinformatic work; G.R.-P. was responsible for SEM analysis; H.D., H.L., Z.L., R.K., H.F., A.A., T.S. and M.M. worked directly with and provided the archaeological samples. All authors contributed to writing the manuscript, approve this study, and are accountable for all aspects of the work.

Competing interests. We declare we have no competing interests.

Funding. Study supported by ERC Investigator grant no. 295729-CodeX. M.D.T. was additionally supported by ERC investigator grant no. 787282-B2C.

Acknowledgements. We thank Daniel Bradley for his mentorship, discussions and reading of the paper; Andreas Steiger-Thirsfeld for SEM images and N. Tehrani for sample photography.

References

1. Aali A, Stöllner T, Abar A, Rühli F. 2012 The salt men of Iran: the salt mine of Douzlākh. *Chehrābād.* **42**, 61–81.
2. Aali A, Stöllner T. (eds) 2015 *The archaeology of the Salt Miners. Interdisciplinary Research 2010–2014*. Bochum, Germany: Metalla.
3. Stöllner T, Aali A, Bagherpour Kashani N. 2020 *Tod im Salz. Eine archäologische Ermittlung in Persien. Begleitbuch, Katalog und graphic novel*. Bochum, Germany: Nünnerich-Asmus Verlag Media.
4. Öhrström LM, Seiler R, Böni T, Aali A, Stöllner T, Rühli FJ. 2016 Erratum to: Radiological findings in an ancient Iranian salt mummy (Chehrābād ca. 410–350 BC). *Skeletal Radiol.* **45**, 433. (doi:10.1007/s00256-015-2316-0)
5. Nasab H V, Aali A, Kazzazi M, Pollard M, Stöllner T. 2019 Reappraisal of the number of salt mummies identified in Chehrābād Salt Mine, Zanjan, Iran. *Bioarchaeol. Near East* **13**, 23–47.
6. Mashkour M. 2014 Faunal report of the 2010 and 2011 excavation campaigns. In *The archaoelogy of the salt miners* (eds A Aali, T Stöllner), pp. 112–117. Bochum, Germany: Metatta.
7. Mashkour M, Fathi H, Davoudi H. 2020 Interaktionen zwischen Mensch und Tier im Salzbergwerk Douzlākh, Chehrābād. In *Tod im salz. Eine archäologische Ermittlung in Persien. Begleitbuch, Katalog und graphic novel* (eds T Stöllner, A Aali, N Bagherpour Kashani), pp. 219–224. Veröffentlichungen aus dem Deutschen Bergbau-Museum Bochum 246, Bochum/Oppenheim, Germany: Nünnerich-Asmus Verlag Media.
8. Warinner C, Bouwman A. 2015 Ancient DNA investigation of the Chehrābād Mummies: Interim Report. In *The archaeology of the salt miners: interdisciplinary research 2010–2014* (eds A Aali, T Stöllner), pp. 90–96. Bochum, Germany: Deutsches Bergbau-Museum.
9. David R. 2008 *Egyptian mummies and modern science*. Cambridge, UK: Cambridge University Press.
10. Dabney J, Meyer M, Pääbo S. 2013 Ancient DNA damage. *Cold Spring Harb. Perspect. Biol.* **5**, a012567. (doi:10.1101/cshperspect.a012567)
11. Daly KG *et al.* 2018 Ancient goat genomes reveal mosaic domestication in the Fertile Crescent. *Science* **361**, 85–88. (doi:10.1126/science.aas9411)
12. Kistler L, Ware R, Smith O, Collins M, Allaby RG. 2017 A new model for ancient DNA decay based on paleogenomic meta-analysis. *Nucleic Acids Res.* **45**, 6310–6320. (doi:10.1093/nar/gkx361)
13. Cassidy LM, Teasdale MD, Carolan S, Enright R, Werner R, Bradley DG, Finlay EK, Mattiangeli V. 2017 Capturing goats: documenting two hundred years of mitochondrial DNA diversity among goat populations from Britain and Ireland. *Biol. Lett.* **13**, 20160876. (doi:10.1098/rsbl.2016.0876)
14. O'Sullivan NJ, Teasdale MD, Mattiangeli V, Maixner F, Pinhasi R, Bradley DG, Zink A. 2016 A whole mitochondria analysis of the Tyrolean Iceman's leather provides insights into the animal sources of Copper Age clothing. *Sci. Rep.* **6**, 31279. (doi:10.1038/srep31279)
15. Teasdale MD *et al.* 2017 The York Gospels: a 1000-year biological palimpsest. *R. Soc. Open Sci.* **4**, 170988. (doi:10.1098/rsos.170988)
16. Nezamabadi M, Aali A, Stöllner T, Mashkour M, Le Bailly M. 2013 Paleoparasitological analysis of samples from the Chehrabad salt mine (Northwestern Iran). *Int. J. Paleopathol.* **3**, 229–233. (doi:10.1016/j.ijpp.2013.03.003)
17. Arbuckle BS, Hammer EL. 2019 The rise of pastoralism in the Ancient Near East. *J. Archaeol. Res.* **27**, 391–449. (doi:10.1007/s10814-018-9124-8)
18. Dong K *et al.* 2020 Genomic analysis of worldwide sheep breeds reveals *PDGFD* as a major target of fat-tail selection in sheep. *BMC Genomics* **21**, 800. (doi:10.1186/s12864-020-07210-9)
19. Demars J *et al.* 2017 Genome-wide identification of the mutation underlying fleece variation and discriminating Ancestral Hairy Species from Modern Woolly Sheep. *Mol. Biol. Evol.* **34**, 1722–1729. (doi:10.1093/molbev/msx114)
20. Ramsey CB. 2009 Bayesian analysis of radiocarbon dates. *Radiocarbon* **51**, 337–360. (doi:10.1017/S0033822200033865)
21. Reimer PJ *et al.* 2013 IntCal13 and marine13 radiocarbon age calibration curves 0–50,000 Years cal BP. *Radiocarbon* **55**, 1869–1887. (doi:10.2458/azu_js_rc.55.16947)
22. Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L. 2013 mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684. (doi:10.1093/bioinformatics/btt193)
23. Wood DE, Lu J, Langmead B. 2019 Improved metagenomic analysis with Kraken 2. *Genome Biol.* **20**, 257. (doi:10.1186/s13059-019-1891-0)
24. Knights D, Kuczynski J, Charlson ES, Zaneveld J, Mozer MC, Collman RG, Bushman FD, Knight R, Kelley ST. 2011 Bayesian community-wide culture-independent microbial source tracking. *Nat. Methods* **8**, 761–763. (doi:10.1038/nmeth.1650)
25. Emmons AL, Mundorff AZ, Keenan SW, Davoren J, Andronowski J, Carter DO, Debruyn JM. 2020 Characterizing the postmortem human bone microbiome from surface-decomposed remains. *PLoS ONE* **15**, e0218636. (doi:10.1371/journal.pone.0218636)
26. Robinson CK *et al.* 2015 Microbial diversity and the presence of algae in halite endolithic communities are correlated to atmospheric moisture in the hyper-arid zone of the Atacama Desert. *Environ. Microbiol.* **17**, 299–315. (doi:10.1111/1462-2920.12364)
27. Ross AA, Müller KM, Scott Weese J, Neufeld JD. 2018 Comprehensive skin microbiome analysis reveals the uniqueness of human skin and evidence for phylosymbiosis within the class Mammalia. *Proc. Natl Acad. Sci. USA* **115**, E5786–E5795. (doi:10.1073/pnas.1801302115)
28. Moreira-Grez B, Tam K, Cross AT, Yong JWH, Kumaresan D, Nevill P, Farrell M, Whiteley AS. 2019 the bacterial microbiome associated with arid biocrusts and the biogeochemical influence of biocrusts upon the underlying soil. *Front. Microbiol.* **10**, 2143. (doi:10.3389/fmicb.2019.02143)
29. Salter SJ *et al.* 2014 Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* **12**, 1–12. (doi:10.1186/s12915-014-0087-z)
30. Nayfach S, Rodriguez-Mueller B, Garud N, Pollard KS. 2016 An integrated metagenomics pipeline for

strain profiling reveals novel patterns of bacterial transmission and biogeography. *Genome Res.* **26**, 1612–1625. (doi:10.1101/gr.201863.115)

31. Korneliussen TS, Albrechtsen A, Nielsen R. 2014 ANGSD: analysis of next generation sequencing data. *BMC Bioinf.* **15**, 356. (doi:10.1186/s12859-014-0356-4)
32. Edgar RC. 2004 MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797. (doi:10.1093/nar/gkh340)
33. Gouy M, Guindon S, Gascuel O. 2010 SeaView Version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **27**, 221–224. (doi:10.1093/molbev/msp259)
34. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010 New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321. (doi:10.1093/sysbio/syq010)
35. Darriba D, Taboada GL, Doallo R, Posada D. 2012 jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* **9**, 772–772. (doi:10.1038/nmeth.2109)
36. Guindon S, Gascuel O. 2003 A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704. (doi:10.1080/10635150390235520)
37. Kijas JW *et al.* 2012 Genome-wide analysis of the World's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biol.* **10**, e1001258. (doi:10.1371/journal.pbio.1001258)
38. Wang C, Zhan X, Liang L, Abecasis GR, Lin X. 2015 Improved ancestry estimation for both genotyping and sequencing data using projection Procrustes analysis and genotype imputation. *Am. J. Hum. Genet.* **96**, 926–937. (doi:10.1016/j.ajhg.2015.04.018)
39. Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. 2012 Ancient admixture in human history. *Genetics* **192**, 1065–1093. (doi:10.1534/genetics.112.145037)
40. Pickrell JK, Pritchard JK. 2012 Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* **8**, e1002967. (doi:10.1371/journal.pgen.1002967)
41. Alexander DH, Lange K. 2011 Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinf.* **12**, 1–6. (doi:10.1186/1471-2105-12-246)
42. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011 Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26. (doi:10.1038/nbt.1754)
43. Taylor WTT *et al.* 2021 Evidence for early dispersal of domestic sheep into Central Asia. *Nat. Hum. Behav*. (doi:10.1038/s41562-021-01083-y)
44. Li X *et al.* 2020 Whole-genome resequencing of wild and domestic sheep identifies genes associated with morphological and agronomic traits. *Nat. Commun.* **11**, 2815. (doi:10.1038/s41467-020-16485-1)
45. Miller W *et al.* 2012 Polar and brown bear genomes reveal ancient admixture and demographic footprints of past climate change. *Proc. Natl Acad. Sci. USA* **109**, E2382–E2390. (doi:10.1073/pnas.1210506109)
46. Orlando L *et al.* 2013 Recalibrating *Equus* evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* **499**, 74–78. (doi:10.1038/nature12323)
47. Librado P *et al.* 2015 Tracking the origins of Yakutian horses and the genetic basis for their fast adaptation to subarctic environments. *Proc. Natl Acad. Sci. USA* **112**, E6889–E6897. (doi:10.1073/pnas.1513696112)
48. Sikora M *et al.* 2019 The population history of northeastern Siberia since the Pleistocene. *Nature* **570**, 182–188. (doi:10.1038/s41586-019-1279-z)
49. Broushaki F *et al.* 2016 Early Neolithic genomes from the eastern Fertile Crescent. *Science* **353**, 499–503. (doi:10.1126/science.aaf7943)
50. Allentoft ME *et al.* 2012 The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proc. R. Soc. B* **279**, 4724–4733. (doi:10.1098/rspb.2012.1745)
51. Briggs AW, Stenzel U, Meyer M, Krause J, Kircher M, Pääbo S. 2010 Removal of deaminated cytosines and detection of *in vivo* methylation in ancient DNA. *Nucleic Acids Res.* **38**, e87. (doi:10.1093/nar/gkp1163)
52. Sawyer S, Krause J, Guschanski K, Savolainen V, Pääbo S. 2012 Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS ONE* **7**, e34131. (doi:10.1371/journal.pone.0034131)
53. Schwarz C, Debruyne R, Kuch M, Mcnally E, Schwarcz H, Aubrey AD, Bada J, Poinar H. 2009 New insights from old bones: DNA preservation and degradation in permafrost preserved mammoth remains. *Nucleic Acids Res.* **37**, 3215–3229. (doi:10.1093/nar/gkp159)
54. Gilbert MTP *et al.* 2004 Ancient mitochondrial DNA from hair. *Curr. Biol.* **14**, R463–R464. (doi:10.1016/j.cub.2004.06.008)
55. Metcalf JL *et al.* 2016 Microbial community assembly and metabolic function during mammalian corpse decomposition. *Science* **351**, 158–162. (doi:10.1126/science.aad2646)
56. Pereira F, Davis SJM, Pereira L, Mcevoy B, Bradley DG, Amorim A. 2006 Genetic signatures of a Mediterranean influence in Iberian Peninsula sheep husbandry. *Mol. Biol. Evol.* **23**, 1420–1426. (doi:10.1093/molbev/msl007)
57. Chen S-Y, Duan Z-Y, Sha T, Xiangyu J, Wu S-F, Zhang Y-P. 2006 Origin, genetic diversity, and population structure of Chinese domestic sheep. *Gene* **376**, 216–223. (doi:10.1016/j.gene.2006.03.009)
58. Demirci S, Baştanlar EK, Dağtaş ND, Pişkin E, Engin A, Özer F, Yüncü E, Doğan ŞA, Togan İ. 2013 Mitochondrial DNA diversity of modern, ancient and wild sheep (*Ovis gmelinii anatolica*) from Turkey: new insights on the evolutionary history of sheep. *PLoS ONE* **8**, e81952. (doi:10.1371/journal.pone.0081952)
59. Chessa B *et al.* 2009 Revealing the history of sheep domestication using retrovirus integrations. *Science* **324**, 532–536. (doi:10.1126/science.1170587)
60. Zhao Y-X *et al.* 2017 Genomic reconstruction of the history of native sheep reveals the peopling patterns of nomads and the expansion of early pastoralism in East Asia. *Mol. Biol. Evol.* **34**, 2380–2395. (doi:10.1093/molbev/msx181)
61. Lv F-H *et al.* 2015 Mitogenomic meta-analysis identifies two phases of migration in the history of Eastern Eurasian Sheep. *Mol. Biol. Evol.* **32**, 2515–2533. (doi:10.1093/molbev/msv139)
62. Rast-Eicher A. 2016 *Fibres: microscopy of archaeological textiles and furs*. Budapest, Hungary: Archaeolingua Alapítvány.
63. Rast-Eicher A. 2010 Fell- und Lederreste aus den Gräbern 2008. In *Spiez-Einigen, Holleeweg 3, Naturwissenschaftliche Untersuchungen zu den bronzezeitlichen Bestattungen* (eds C Cooper, M Harbeck, M Kühn, A Rast-Eicher, M Schweissing, S Ulrich-Bochsler, P Vandorpe), pp. 175–198. Bern, Switzerland: Service Archéologique du Canton de Berne.
64. Allain D, Pena-Arnaud B, Foulquie D, Bourdillon Y, François D. 2014 Introgression of wool-shedding genes into the Romane breed sheep. In *10. World congress of genetics applied to livestock production*, pp. 17–22. Vancouver, Canada: American Society of Animal Science.
65. Mashkour M. 2013 Animal exploitation during the Iron Age to Achaemenid, Saasanian and early Islamic Periods. In *Persia's imperial power in the late antiquity* (eds EW Sauer, HO Rekavandi, TJ Wilkinson, J Nokan-Den), pp. 548–580. Oxford, UK: Oxbow Books.
66. Pöllath N, Schafberg R, Peters J. 2019 Astragalar morphology: approaching the cultural trajectories of wild and domestic sheep applying geometric morphometrics. *J. Archaeol. Sci.* **23**, 810–821. (doi:10.1016/j.jasrep.2018.12.004)
67. Vila E *et al.* 2021 EVOSHEEP: the makeup of sheep breeds in the ancient Near East. *Antiquity* **95**(379), E2. (doi:10.15184/aqy.2020.247)
68. Khazaeli R. 2013 *The subsistence economy of the Old Nishapur from the formation of the city up to the Mongol Period*. Tehran, Iran: University of Tehran.
69. Rossi C *et al*. 2021 Exceptional ancient DNA preservation and fibre remains of a Sasanian saltmine sheep mummy in Chehrābād, Iran. Figshare.

Supplementary Material: Exceptional ancient DNA preservation and fibre remains of a Sasanian saltmine sheep mummy in Chehrābād, Iran

Conor Rossi[1], Gabriela Ruß-Popa[2], Valeria Mattiangeli[1], Fionnuala McDaid[1], Andrew J. Hare[1], Hossein Davoudi[3], Haeedeh Laleh[3,4], Zahra Lorzadeh[3], Roya Khazaeli[3], Homa Fathi[3], Matthew D. Teasdale[5], Abolfazl A'ali[6], Thomas Stöllner[7], Marjan Mashkour[3,8], Kevin G. Daly[1]

1. Smurfit Institute of Genetics, Trinity College Dublin, Dublin 2, Ireland
2. Austrian Academy of Sciences, Austrian Archaeological Institute, Archaeological Sciences, Hollandstraße 11-13, 1020 Vienna, Austria
3. University of Tehran, Central Laboratory, Bioarchaeology Laboratory, 1417634934 Tehran, Iran
4. University of Tehran, Faculty of Humanities, Department of Archaeology, 1417935840 Tehran, Iran
5. McDonald Institute for Archaeological Research, Dept. of Archaeology, University of Cambridge
6. Zanjan Cultural Heritage Centre, Archaeological Museum of Zanjan, Emaarate Zolfaghari,Taleghani St., Zanjan, Iran
7. Ruhr University Bochum, Institute for Archaeological Studies & Deutsches Bergbau-Museum Bochum, Research Department, Haus der Archäologien, Am Bergbaumuseum 31, D-44791 Bochum
8. Archéozoologie, Archéobotanique, Sociétés, Pratiques et Environnements (AASPE), Muséum national d'Histoire naturelle, Sorbonne Université, CNRS; CP 56, 55 rue Buffon, 75005 Paris, France

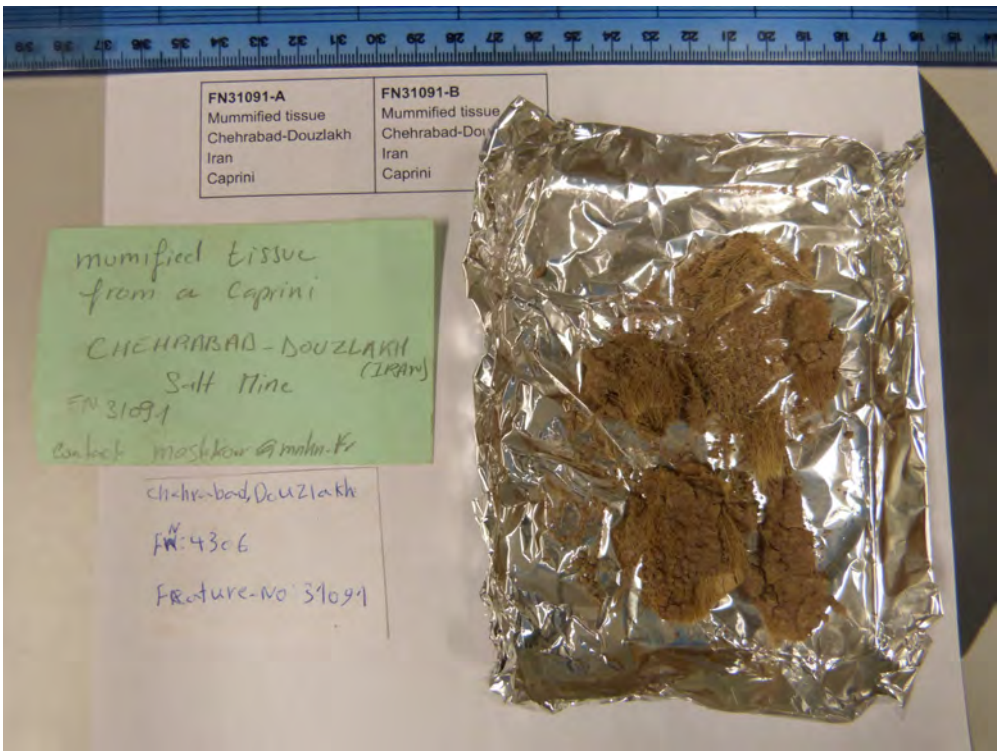
Contextual Information for Archaeological Sites

Chehrābād, Iran

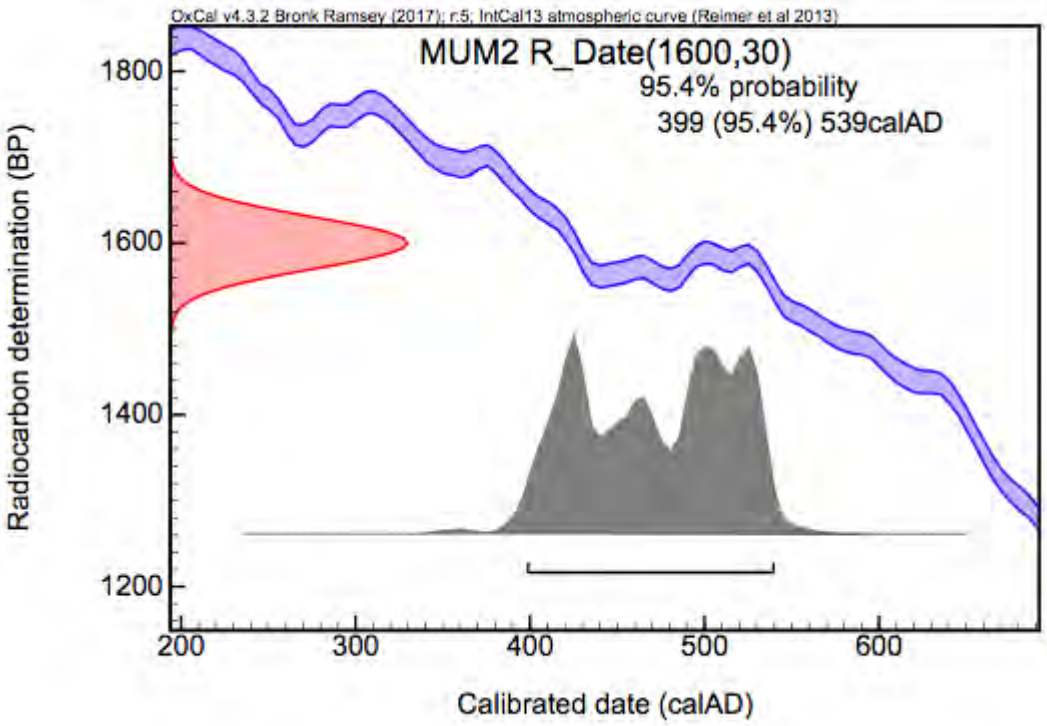
The mummified sheep leg 4305 (Supplementary Figure 1) was discovered during Iranian-German research work in one of the debris layers (feature 31091) in the northwestern edge of the mine during modern archaeological excavations of the Douzlākh saltmine of Chehrābād [1]. At the northwestern edge, part of a large mining gallery was discovered. This part had collapsed between 405 and 380 BCE but was reopened in the Early Sasanian period (2^{nd}-3^{rd} centuries CE). The working area in the northwestern edge was partly refilled most likely between the 4^{th} and 5^{th} century CE, of which the layer 31091 belongs. The dating of the sheep-leg (Table 1, Supplementary Figure 3) is concurrent to the layer dating (reference [1], 43-45, Table 11). It was most likely deposited in the course of refilling activities during a salt extraction phase when mining took place on the northwestern salt rock edge of the mine.

A skin cutting of sheep leg 4305, MUM2 (Supplementary Figure 2), was used for DNA extraction and C_{14} date estimation.

Supplementary Figure 1: Sheep leg (4305) in the archaeological context of layer 31091 (left) and recent condition after cleaning (two angles, centre, right), Photos: DBM/RUB/AMZ, Thomas Stöllner (left), N. Tehrani (centre, right).



Supplementary Figure 2. Photograph of the mummified skin of MUM2. MUM2 was sampled from the sheep leg 4305.



Supplementary Figure 3: OxCal 4.3.2 [2] **calibration of the radiocarbon age of MUM2 using IntCal13** [3]**.**

Tepe Hasanlu, Western Azerbaijan, Iran

Tepe Hasanlu is one of the key sites of northwestern Iran, due to its long-term occupation and
well-defined stratigraphy. It is located in the Solduz valley on the southern shore of Lake Urmia, Western Azerbaijan province of Iran (1043 m ASL, Latitude: 37°0'16.15"N, Longitude: 45°27'31.74"E). Robert H. Dyson Jr. directed ten seasons of excavations at Hasanlu from 1956 to 1977 [4,5]. The site was occupied during ten cultural periods from the Late Neolithic (period X) to the Ilkhanid dynasty (period I) [6]. The most represented periods in the site are the Late Bronze (period V) and Iron Age (period IV-III). The Iron Age II citadel of Hasanlu was destroyed and fired during a battle, probably by the Urartian army, around 800 BC [7,8]. Hasanlu period IIIc and b, are attributed to Iron Age III (Urartian period / 800-600 BCE) and period IIIa allocated to the Achaemenid Empire (550-530 BC). The citadel with a fortification wall, a few massive public buildings like garrison quarters, flimsy structures, stables and other architectural remains had been discovered from period III [9]. Extensive excavations in these periods resulted in the discovery of burned buildings, thousands of artifacts in closed contexts, and a large quantity of human and animal bone remains. Animal bones are among the most abundant materials recovered in the site. Despite the taxonomic diversity of the remains in Hasanlu, domestic sheep, goat, and cattle were the most exploited animals in all the periods and cattle contributed much less to the diet than sheep and goat during the Bronze Age and Iron I to II, while its contribution increased to as much as Caprini during Iron III and Historical periods [10,11]. The sheep sample from Tepe Hasanlu analysed in this study was a petrous bone recovered from deposits of level IIIb (Iron Age III) in campaign 1974, Op. X32, Stratum 5, Lot 26.

Azer2 Azer2 MM TH3

Nishapur, Khorasan, Iran

Nishapur, one of the major urban centers of Khorasan, is located in Northeast Iran on a plain limited by the Binalud heights in the North and the Kashmar heights in the South (1198 m ASL, Latitude: 36°10'16.72"N, Longitude: 58°50'50.82"E). Nishapur as a major social and economic urban area exists in a cultural landscape in which the economic life of rural and urban entities was based on agriculture, horticulture, and animal husbandry. Nishapur had been an important producing center of silk, wool, cotton and textiles and their export throughout the Islamic world and beyond [12]. Archaeological research in Nishapur began by the Metropolitan Museum of Art during the 1930s [13–17]. Since then, surveys and excavations have been realized by national and international missions. Excavations conducted by the Iranian-French team resulted in the discovery of a significant amount of animal remains in the citadel [18,19]. Recently, a new Iranian excavation that has been carried out in the citadel and the urban area also produced substantial bioarchaeological data [20]. 2470 pieces of animal bones from the first collection found in 2005-2006 were studied in the Archaeozoology section in the Bioarchaeology Laboratory of the Central laboratory of the University of Tehran [21]. The remains could be dated from the Sasanian period to the 12th century CE. The results indicate that sheep, goats, and cattle were the most exploited animals for both meat and secondary products. The Nishapur petrous bone sample analyzed in the present study was discovered in the campaign 2006, Tr. 15B located in the north part of the citadel, Square: JXV, 6-12th century CE.

Khor1 Khor1 MM NKD1

Supplementary Methods

DNA extraction

Sample preparation, extraction and library preparation were performed in a dedicated ancient DNA laboratory in the Smurfit Institute of Genetics, Trinity College Dublin. A small section of skin (0.454g) was cut using a disposable blade. Prior to the extraction the sample was washed with 1ml of water (lab grade and UVed) then with 600ul of Ethanol (99%) and then again with 1 ml of water. Each time the supernatant was removed after centrifugation. DNA was extracted using a modified version of the classic phenol/chloroform DNA extraction protocol [22]. A modified extraction buffer was used, where 3mM CaCl2 and 30mM DTT were added to the original extraction buffer.

Library preparation and sequencing

Treatment of ancient DNA with Uracil-DNA glycosylase (UDG) and Endonuclease VIII has been shown to remove base misincorporations caused by deamination in aDNA [23]. For UDG-treated libraries, 5 µl USER® Enzyme (1,000U/ml; Uracil-Specific Excision Reagent, New England BioLabs Inc.) was added to 16.25µl purified DNA and incubated in an Eppendorf ThermoMixer for 3 hours at 37°C, prior to library preparation.

Both UDG-treated and not-treated libraries were constructed based on the protocol of [24] with modifications as reported in [25]. Control tubes (21.25µl H_2O starting material) were carried out for library construction. Indexing Polymerase Chain Reactions (PCRs) were performed using Accuprime *Pfx* Supermix (Invitrogen), primer IS4 (10 µM) and a unique indexing primer (5 µM) as detailed in [24]). The concentration of amplified libraries was quantified with a TapeStation 2200 (Agilent).

Three rounds of sequencing were done: single-end and paired-end shotgun sequencing of USER-treated libraries on Illumina HiSeq 2500 (Macrogen), and single-end shotgun sequencing for non USER-treated libraries on Illumina MiSeq (TrinSeq). The UDG-treated library was sequenced on a HiSeq 2500 Illumina platform (100bp SE and 100bp PE) via a commercial sequencing company (Macrogen, Republic of Korea) while the non-UDG-treated library was sequenced on a MiSeq Illumina platform (50bp SE) at TrinSeq (Trinity College Dublin, Ireland) with Phi X control at 1%.

Raw-read processing

The quality of fastq files was assessed using FastQC [26]. For single-end libraries, adapters were trimmed from reads of raw fastq files using cutadapt 1.9.1 [27] (cutadapt -a AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC -O 1 -m 30). This removed reads shorter than 30bp and trimmed adapter sequence if an overlap of more than 1 bp was found between the read and adapter sequence.

For paired-end libraries, adapters were trimmed using AdapterRemoval v2.1.1 [28] (AdapterRemoval --file1 reads_1.fq --file2 reads_2.fq --basename output_paired --trimns

--trimqualities --minquality 25 --collapse). The --trimns --trimqualities --minquality flags cause stretches of consecutive Ns and/or bases of quality lower than 25 from the 5' and 3' termini to be trimmed. The --collapse flag merges mate pairs that overlap by at least 11 bp into a single read. Base quality scores of collapsed reads are then recalculated using the metadata of the paired reads. The collapsed fastq file and the truncated read pairs were used in downstream analysis.

Read alignment

All trimmed reads of samples were aligned to OviAri3.1 using BWA [29] 0.7.5a with relaxed parameters (-l 1024 -n 0.01 -o 2). These flags disable seeding, allow for more mismatches with the reference and allow 2 gap openings, respectively.

For single-end reads, the resulting sai files were converted to sam file format using BWA samse [29]. Paired-end reads were handled using BWA sampe [29] to collapse sai pairs into one sam file. Sam files were then converted into the binary bam files using SAMtools (v1.7) view [30]. The -F4 flag was used with SAMtools view to discard unmapped reads for all bam files. For paired-end reads, the -f2 flag was also used to remove pairs of reads where one paired read was aligned to a different chromosome. The bam files were then sorted using SAMtools sort and PCR optical duplicates were removed using SAMtools rmdup. Bam files were then filtered for reads of mapping quality less than 30 using SAMtools view. Read groups were added to each bam file using Picard AddOrReplaceReadGroups (https://broadinstitute.github.io/picard/). Finally, these filtered bam files were merged using SAMtools merge. Genome coverage was calculated using GATK DepthOfCoverage [31]. Alignment statistics are summarised in Supplementary Table 1.

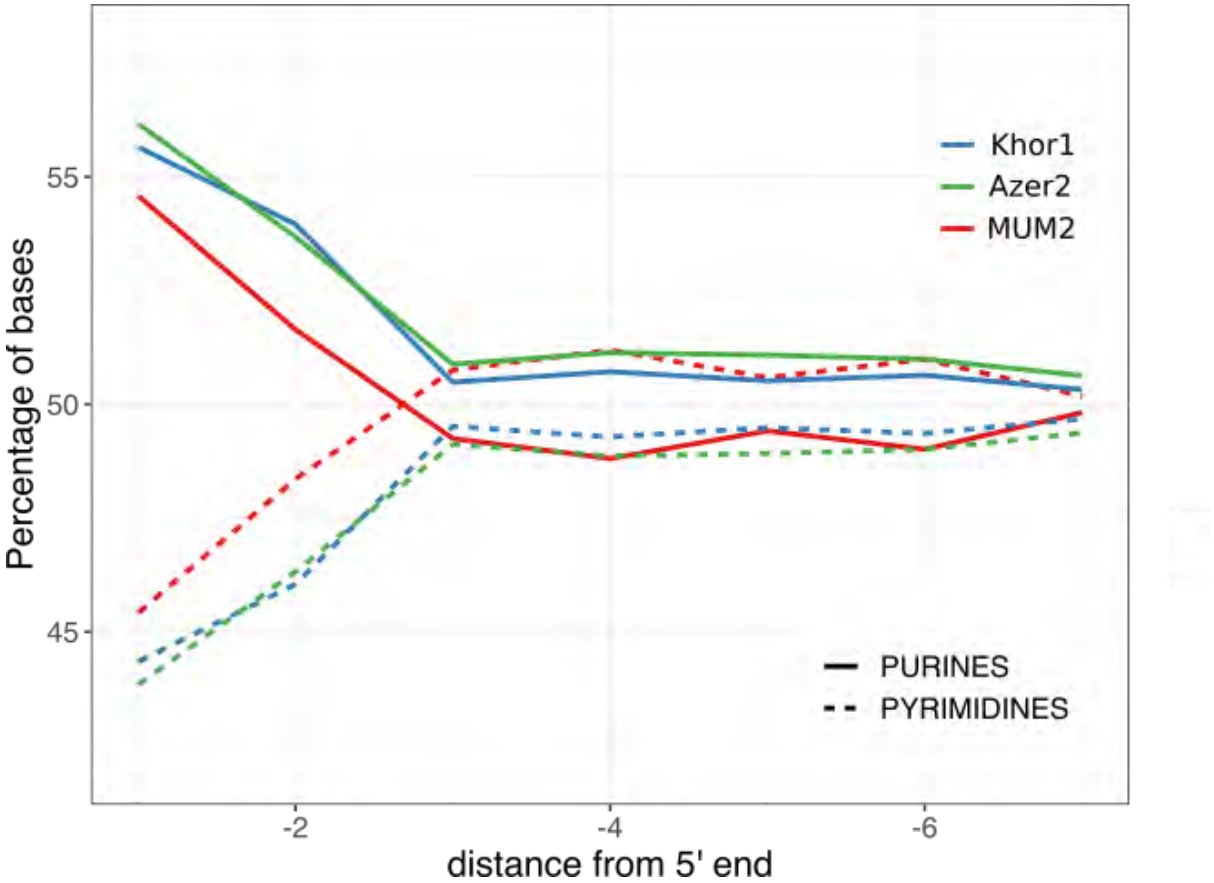
Sample	Library type	Sequencing platform	Raw Reads	Trimmed Reads	Aligned Reads	MQ30, rmdup Aligned Reads	Endog. %
MUM2	Non-USER treated	Illumina MiSeq 63 SE	6,371,514	6,317,966	2,479,430	1,959,646	31.02
MUM2	USER treated	Illumina HiSeq 100 SE	63,389,209	62,636,422	19,172,915	15,576,961	24.87
MUM2	USER treated	Illumina HiSeq 100 PE	598,909,352	348,606,473	109,018,213	86,945,474	24.94
Khor1	Non-USER treated	Illumina MiSeq 65 SE	333,116	319,697	263,477	186,846	58.44
Khor1	USER treated	Illumina HiSeq 100 PE	6,230,996	3,146,933	2,796,988	1,934,890	61.48
Azer2	Non-USER treated	Illumina MiSeq 65 SE	1,430,246	1,379,462	645,402	432,134	31.33
Azer2	USER treated	Illumina HiSeq 100 PE	20,448,674	10,835,702	5,029,767	3,074,445	28.37

Supplementary Table 1: Sequencing and alignment statistics.

Damage patterns

Bam files were assessed using mapDamage2.0 [32]. mapDamage2.0 calculates read length distributions and substitution patterns at the 5' and 3' ends of read fragments of aligned bam files. mapDamage2.0 also outputs a number of damage parameters estimated in a Bayesian manner including δS, the cytosine deamination probability in single strand context. Published sequencing data of other ancient samples were retrieved from NCBI GenBank in order to contextualise the damage patterns of our sample with published data (Supplementary Table 3).

Depurination levels were inferred using frequencies of purines close to the 5' end of fragment strand-breaks. The use of *T4* DNA polymerase in our library preparation removes 3' overhangs, therefore only 5' ends are considered. Only non-UDG libraries were used to calculate these levels as USER excises uracil residues in DNA fragments, creating new strand breaks. Using Qualimap v2.1.3 [33], nucleotide frequencies were calculated on filtered bam files at each position of DNA fragments (Supplementary Figure 4). Following the methods detailed in [34], approximate depurination rates were calculated.



Supplementary Figure 4: Nucleotide frequencies at 5' end of DNA fragments. The depurination rate was 6.3%, 5.6% and 5.3% for MUM2, Khor1, Azer2, respectively.

Metagenomics

For all metagenomic analysis, we first removed reads originating from the host organism and human contamination. To do this, all trimmed reads were aligned to the sheep genome (oviAri3.1) and the human genome (hg38) using methods detailed above. Unmapped reads

were retained using SAMtools view -f 4 and converted to fastq files using SAMtools fastq. The retained reads were then deduplicated using PrinSeq [35]. In total 202,815,624 reads were retained for MUM2 (176,858,319 collapsed/single end reads and 25,957,305 paired end reads). This was repeated for Khor1 and Azer2, resulting in 281,649 and 4,315,523 reads, respectively.

Unmapped, deduplicated reads were searched against the Kraken2's standard database using default settings [36]. This database includes sequences of bacterial, archaeal and viral genomes. Kraken2 classified 11.96% (24,247,293 reads) of the unmapped, deduplicated reads using this database. To calculate Class abundances, microbial abundances were recalculated using Bracken [37] est_abundance.py, with a minimum threshold of 0.1% of total classified reads. 57.13% (13,499,165) of the classified reads belonged to the salt-adapted *Halobacteria* Class which reflects the high salinity environment of Chehrābād.

Class	Total Reads Assigned	% of Assigned Reads	Domain
Halobacteria	13499165	57.13	Archaea
Actinobacteria	3879524	16.42	Bacteria
Clostridia	1243905	5.26	Bacteria
Alphaproteobacteria	914412	3.87	Bacteria
Gammaproteobacteria	823160	3.48	Bacteria
Betaproteobacteria	659427	2.79	Bacteria
Bacilli	295377	1.25	Bacteria
Other and Host	155083	8.54	N/A

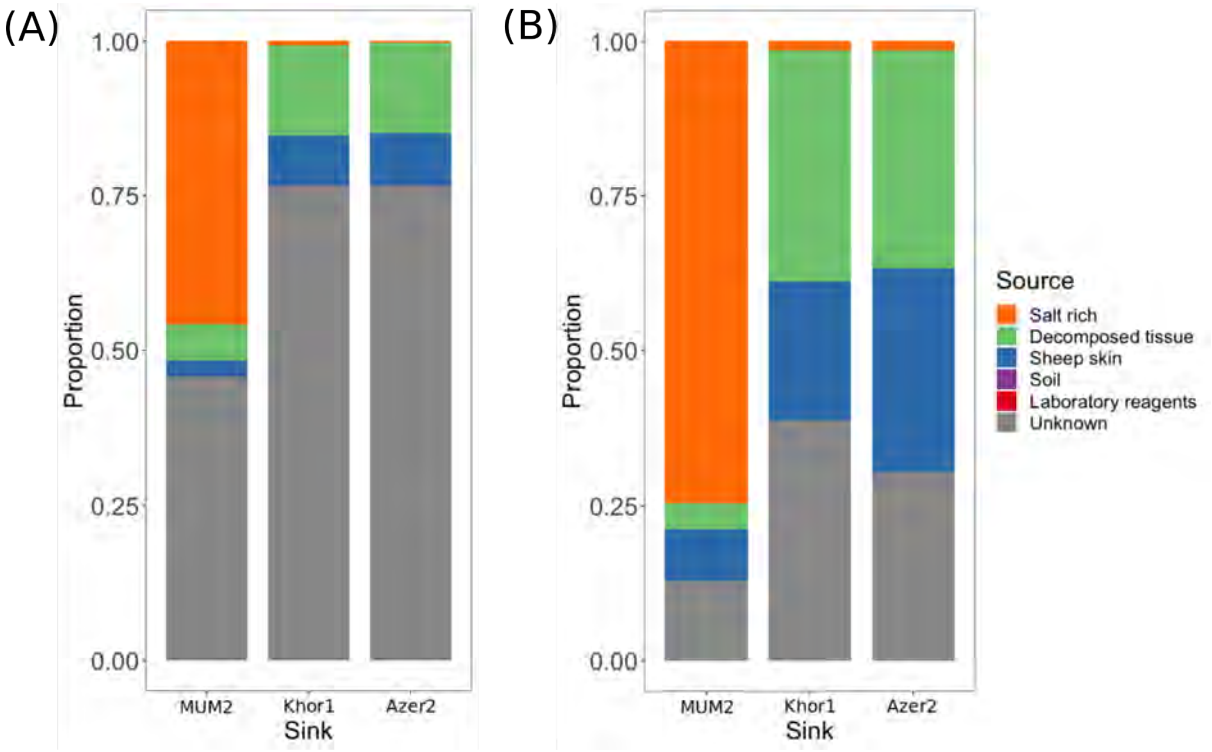
Supplementary Table 2: Taxonomically assigned reads of MUM2 by Kraken2, recalculated by Bracken, at Class level. The salt-adapted Class *Halobacteria* predominates with 57.13% of assigned reads.

Using the same methods, *Halobacteria* were found in very low abundance in our comparative bone samples, Khor1 and Azer2 (0.7% and 0.67% of assigned reads, respectively with a threshold of 0.1%). As a comparison with other ancient skins, metagenomic analysis of the Tyrolean Iceman's preserved clothes [38] and ancient parchment [39] also revealed very low levels of *Halobacteria* present (0-0.02% and 0-0.3% of assigned reads, respectively). The uniquely high levels of halophilic archaea present in MUM2 is suggestive of the large effect the saltmine environment had on the microbial communities.

We used SourceTracker2 to further define the metagenomic composition of MUM2. SourceTracker2 is a Bayesian source-prediction tool that can estimate the proportion of sequencing reads of samples (sinks) originating from selected metagenomes (sources) [40]. DNA sequences of various relevant metagenomes (salt-rich, live sheep skin, soil, tissue decomposers, and laboratory reagents) were downloaded from NCBI (Supplementary Table 4). The sequences were trimmed using AdapterRemoval v2.1.1 and taxonomic assignment for these reference sequences was done by Kraken2, as detailed above. The kraken reports

were merged and converted to a BIOM table using kraken-biom (https://github.com/smdabdoub/kraken-biom) and the human taxid (9606) was removed.

Source prediction of MUM2, Azer2 and Khor1 was done at both Species and Genus level using default settings. We were unable to use the “Laboratory reagent” source at Genus level as there was an insufficient number of taxonomically assigned Genuses (27) for our rarefaction level (1,000). At Species level, SourceTracker2 estimates 0.4725 of the microbial community of MUM2 originates from a salt-rich environment., while only 0.003-0.007 of the comparative bone samples were estimated to be from this environment (Table 2; Supplementary Figure 5A). This estimation of salt rich increases to 0.746 when estimated at Genus level (Table 2; Supplementary Figure 5B).



Supplementary Figure 5: Prediction of metagenomic source proportions by SourceTracker2 at (A) Species level and (B) Genus level.

In order to cross-validate our results, and to investigate individual bacterial species, MIDAS [41] was selected as it has been reported to be the amongst most accurate for aDNA microbial communities [42]. This tool quantifies bacterial species abundance from shotgun metagenomes. MIDAS clusters more than 31,000 bacterial reference genomes into nearly 6,000 species groups based on sequence identity of 30 universal marker genes.

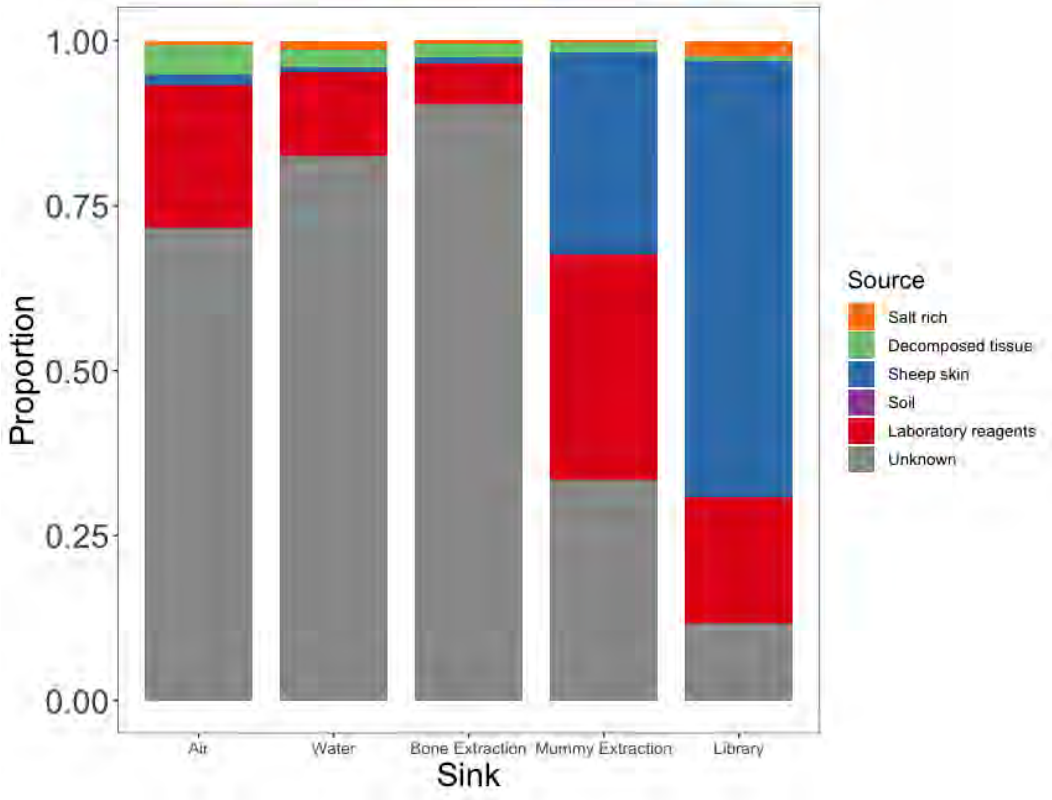
We observe a large proportion of detected microbial species in MUM2 (8 of the 9 most abundant species) as being salt adapted or salt tolerant (Supplementary Table 4). Of these, just one species is identified in ancient sample Azer2. This species, *Streptomyces sp 59965*, is a generalist. Similarly, none of the defined salt adapted bacteria in MUM2 were detected in shotgun data of leather from the Tyrolean Iceman ([38], results not shown). Top MIDAS hits for ancient parchment shotgun data ([39], results not shown) largely comprised of bacteria associated with animal skins (e.g. *Erysipelothrix rhusiopathiae*, *Propionibacterium acnes*,

Staphylococcus equorum, Brevibacterium linens). Some salt-tolerant species (Table 2) were found in low abundance (<2%) of several ancient parchment shotgun data except for *Saccharomonospora glauca 62535* which was found at 4% abundance in sample SAMEA104143134. Overall, these results support the observation that MUM2's microbial profile reflects its unique taphonomic context.

To validate the authenticity of the most abundant bacterium detected, all sequencing data was aligned to GCF_00371785.1, an *Actinopolyspora halophila* reference genome (Actinopolyspora halophila DSM 43834, retrieved from NCBI GenBank). A 0.33X genome was generated (std per bp = 1.13X). Using Qualimap v2.1.3 [33], the coverage of this bam file was visualised. The coverage of the bam file was even, as is recommended as a validation in [43]. The median read length of the aligned reads was 74 bp.

Controls

Controls that were done throughout the sample preparation, extraction and library preparation were also screened for halophilic signatures. The number of reads taxonomically assigned by Kraken2 was low (643 - 4648 assignments) and in all the presence of the *Halobacteria* clade was low (0.009-0.01). Metagenomic source prediction was done at a Species level using the same SourceTracker2 dataset as above (Supplementary Figure 6). The salt-rich environment was not predicted to contribute highly to any of the controls (0.003-0.024). After "Unknown", "Laboratory reagents" were predicted to be the highest contribution to all controls except the Library Control. Here, SourceTracker2 predicted 0.66 of its metagenome to belong to "Sheep skin", suggesting possible skin contamination during library preparation.

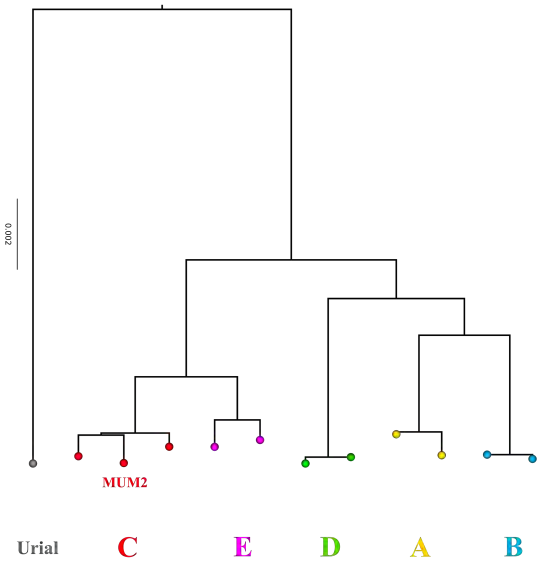


Supplementary Figure 6: Prediction of metagenomic source proportions of controls by SourceTracker2 at Species level.

Mitochondrial tree

All UDG-treated sequencing data was aligned to the domestic sheep mitochondrial reference, AF010406.1 (retrieved from NCBI GenBank) using an identical method as to the nuclear genome alignment. Consensus fasta sequences were generated using ANGSD [44] (angsd -doFasta 2 -doCounts 1 -setMinDepth 3 -minQ 20 -minMapQ 30). With these flags, base calls must have a minimum quality of 20 and a minimum MAQ of 30 to be considered. In the event that there are more than one base at a given site, the most common base is chosen. For a site to be called a minimum depth of 3 is required.

Ten modern sheep whole genome sequences and one urial sheep whole genome sequence published in [45] were retrieved from NCBI GenBank. Multiple sequence alignments were performed with MUSCLE [46]. A maximum-likelihood phylogeny (100 bootstraps) was generated. The HKY85 substitution model was selected for tree-building using jmodeltest2 [47,48]. Alignments and maximum-likelihood phylogeny constructions were implemented in SeaView v. 5.0.1 [46,49,50].



Supplementary Figure 7: Whole mitochondrion maximum-likelihood phylogeny of domestic sheep (100 bootstraps). MUM2 falls within Clade C. All clades of the phylogenetic tree are well-supported by bootstrap values (>0.97).

SNP Calling

The ovine SNP50 HapMap dataset used for the analyses described was provided by the International Sheep Genomics Consortium and obtained from www.sheephapmap.org in agreement with the ISGC Terms of Access. This dataset is genotyped on the Illumina OvineSNP50 Genotyping BeadChip (Illumina), covering >54,000 SNPs evenly spaced over the sheep genome. 1503 genotyped individuals are included in this dataset, representing modern breeds from Africa, the Americas, Europe, SW Asia and Asia [51].The genomic positions of the OvineSNP50 were updated for the OviAri3.1 genome build (https://doi.org/10.6084/m9.figshare.8424935.v2) and SNPs flipped to the forward strand. Using PLINK v1.9 [52], these SNPs were filtered down to 44,223 sites using a MAF filter of ≤ 0.5.

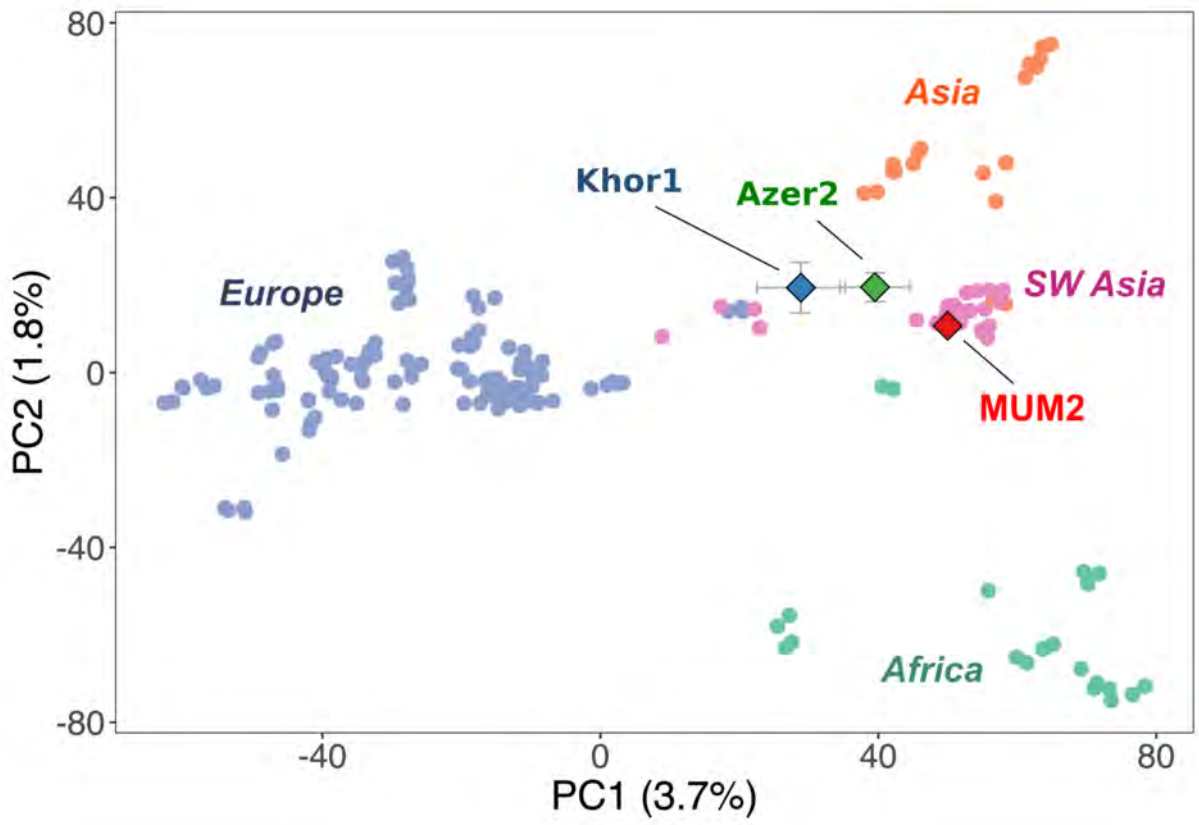
The coverage of our sample was too low (approx. 4X) for accurate diploid genotype calling. To circumvent this issue, the pileup tool in GATK [31] v3.7 was used to report sample base calls for each read at each position of the filtered SNP list (44,223 SNP positions). A minimum base quality of 30 was required to be considered and sites with three or more different bases present were removed. A single base call was then chosen at random from each site and duplicated to create a pseudohaploid homozygous genotype at that position for each individual. In total 43,027 SNPs were called for MUM2.

Using PLINK v1.90, the pseudohaploidised MUM2 genotype calls were merged with the modern sheep breeds dataset. In cases where the allele of MUM2 matched neither of the alleles given in the sheep breed position, the SNP position was flipped using PLINK v1.90 and attempted to be merged again. Any remaining triallelic SNPs were removed. The final merged dataset was used for *f3* statistics, ADMIXTURE and TreeMix analysis.

PCA

Projection PCA using Procrustes analysis was performed using LASER v2.03 [53]. The PCA reference space and projection transformation were constructed using a subsample of our modern breeds dataset. This subsample contained representation from all modern breeds, chosen randomly. Using the OvineSNP50 reference panel, pileup files were generated for MUM2, Khor1 and Azer2 using SAMtools mpileup (-q 30 -Q 20). Each sample was projected onto this PCA space and averaged across ten independent runs. The uncertainty based on the standard deviation * 2 for each sample was plotted on the map (Supplementary Figure 8).

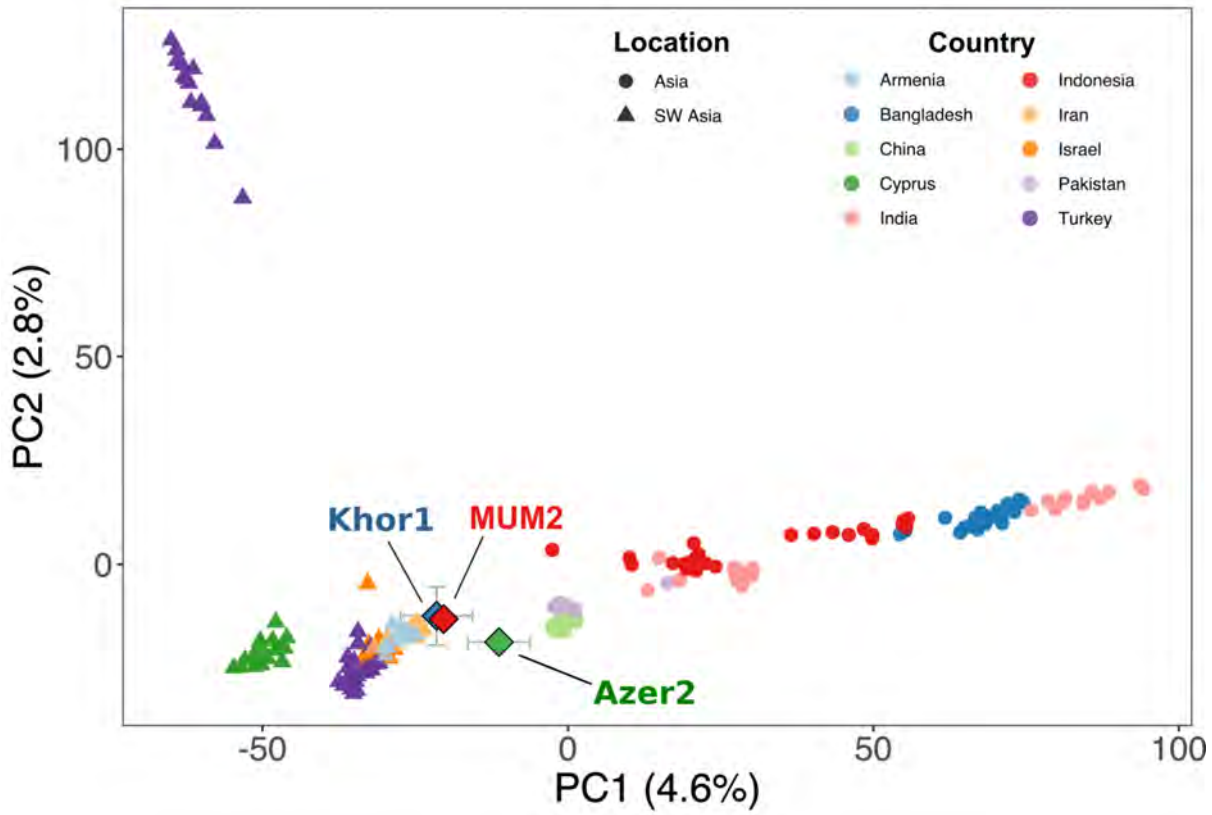
The plot of PC1 vs PC2 shows that PC1 differentiates modern European breeds from breeds of the rest of the world. PC2 differentiates non-European breeds. A cline from African, Asian and South West Asian appears. These breeds do not create discrete clusters, however, and some individuals from different locations can be seen to cluster together in geographically-heterogeneous groups. MUM2 falls within the main cluster of SW Asian breeds. Khor1 and Azer2 both project within the diversity of SW Asian breeds but fall away from this main cluster on PC1. However we note that the uncertainty of PC1 of Khor1 and Azer2 is large, likely due to the lower SNPs called while the coordinates of MUM2 had a low uncertainty.



Supplementary Figure 8: PCA of global sheep affinity and Procrustes projection of MUM2, Khor1 and Azer2. Error bars represent the standard deviation of each PC coordinate between independent runs * 2. PC1 separates breeds of European origin and the rest of the world, while PC2 separates Asian and African breeds into an approximate cline of Central/East Asian, SW Asian, African [51].

Local affinity

To investigate local affinity, a PCA was produced using only breeds of SW Asia and Asia using LASER v2.03. Here, PC1 differentiates Asian and SW Asian breeds. PC2 then separates one SW Asian (Turkish) breed from all other SW Asian breeds. When MUM2, Khor1 were projected onto this reference space using LASER, they clustered with SW Asian breeds rather than those from SE Asia (Supplementary Figure 9). Azer2 falls away from this cluster of SW Asian breeds, however, there is a larger amount of uncertainty again due to lower SNPs called.



Supplementary Figure 9: PCA containing only Asian breeds. Error bars represent the standard deviation of each PC coordinate between independent runs * 2. MUM2 clusters with Southwest Asian breeds. Some amount of genetic flux in local breeds in the past 1600 years is inferred by MUM2 falling slightly away from the main cluster of Southwest Asian breeds. Azer2 falls further away from the main cluster of SW Asian breeds, however, we note increased relative uncertainty of its PC1 position due to the lower number of called SNPs.

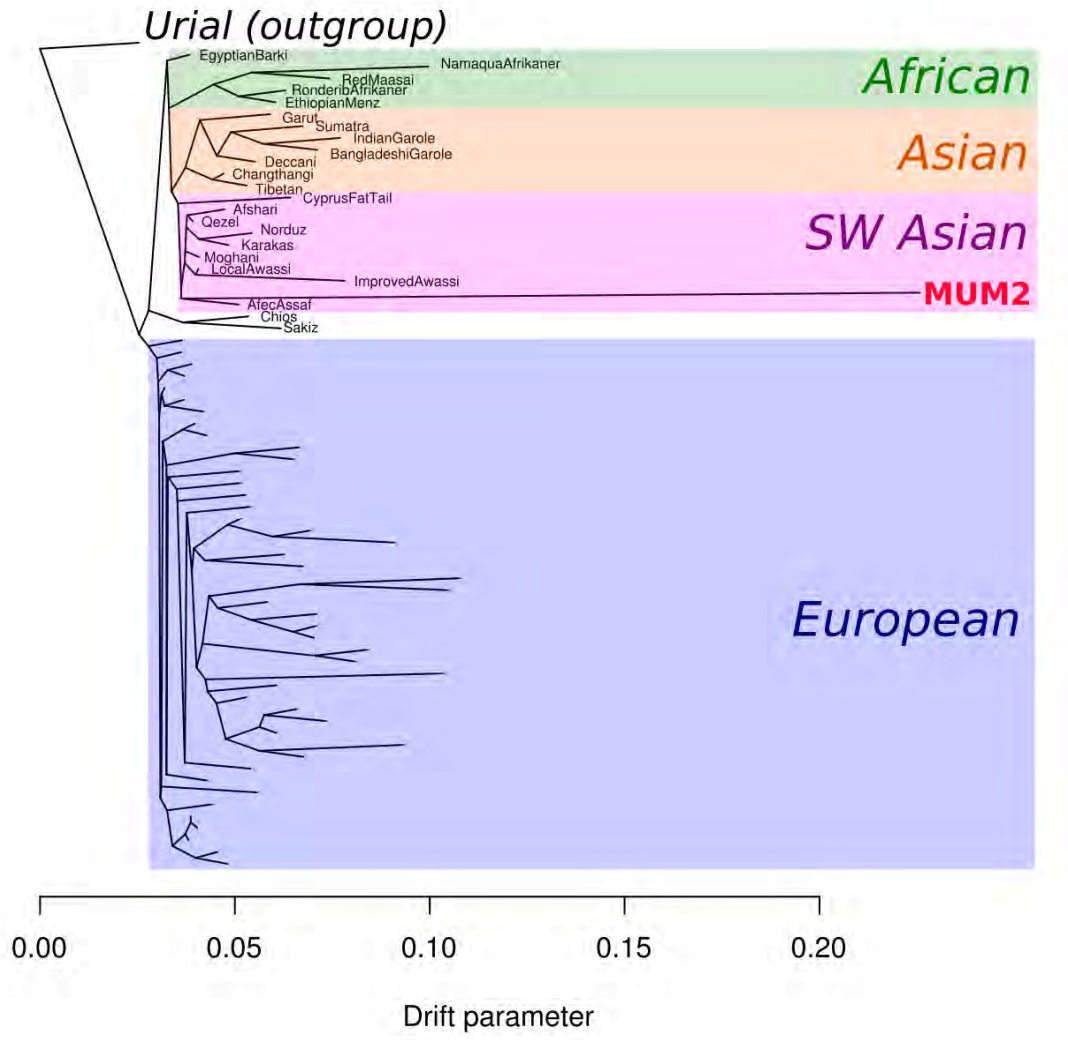
f_3 Statistics

f_3 statistics were computed to quantify the amount of genetic drift shared between MUM2 and modern sheep breeds. Three Asiatic mouflon individuals that were genotyped on the OvineSNP50 Genotyping BeadChip were used as an outgroup. Using PLINK v1.9, these were merged with our full dataset of modern breeds and MUM2. Using PLINK v1.9, a filter of 0 for missingness was applied to this merged dataset. The dataset was then filtered for linkage disequilibrium (LD) between markers, with the command --indep-pairwise 50 5 0.5. Outgroup f_3 values were not estimated for Khor1 and Azer2 due to the lower number of called SNPs.

Using this complete dataset, we calculated the f_3 statistics in the form of f_3(MUM2, modern breed; Asiatic Mouflon) with ADMIXTOOLS [54] (Supplementary Table 5). f_3 statistics were plotted using approximate geographic origins of breeds with R using the sf package [55] (Figure 2A).

TreeMix

The complete dataset used was converted into TreeMix format with the plink2treemix.py utility [56]. TreeMix [56] was used to construct a model of population splits with no migration events. The tree was rooted with Asiatic mouflon and blocks of 500 SNPs were selected (-k 500). No sample size correction was used (-noss). Clades emerge based on the geographic location of the sheep breeds. MUM2 falls in the SW Asian clade.



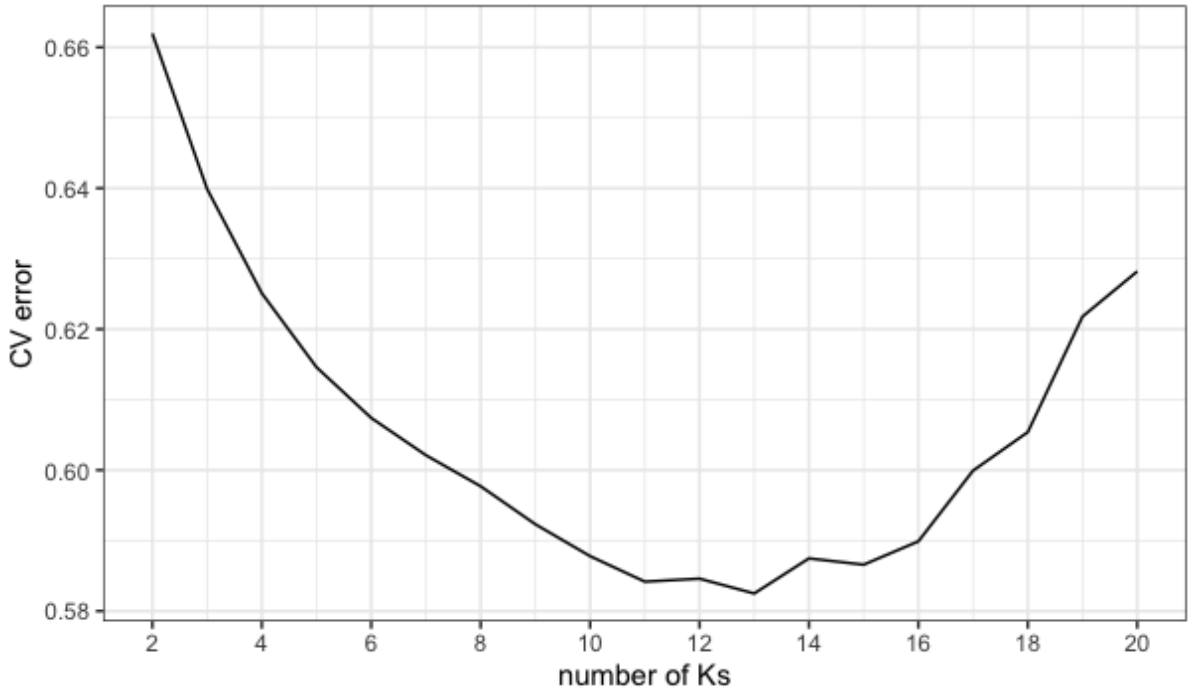
Supplementary Figure 10: TreeMix plot with 0 migration edges. MUM2 falls into the SW Asian breed clade. European breed labels were removed for clarity.

ADMIXTURE

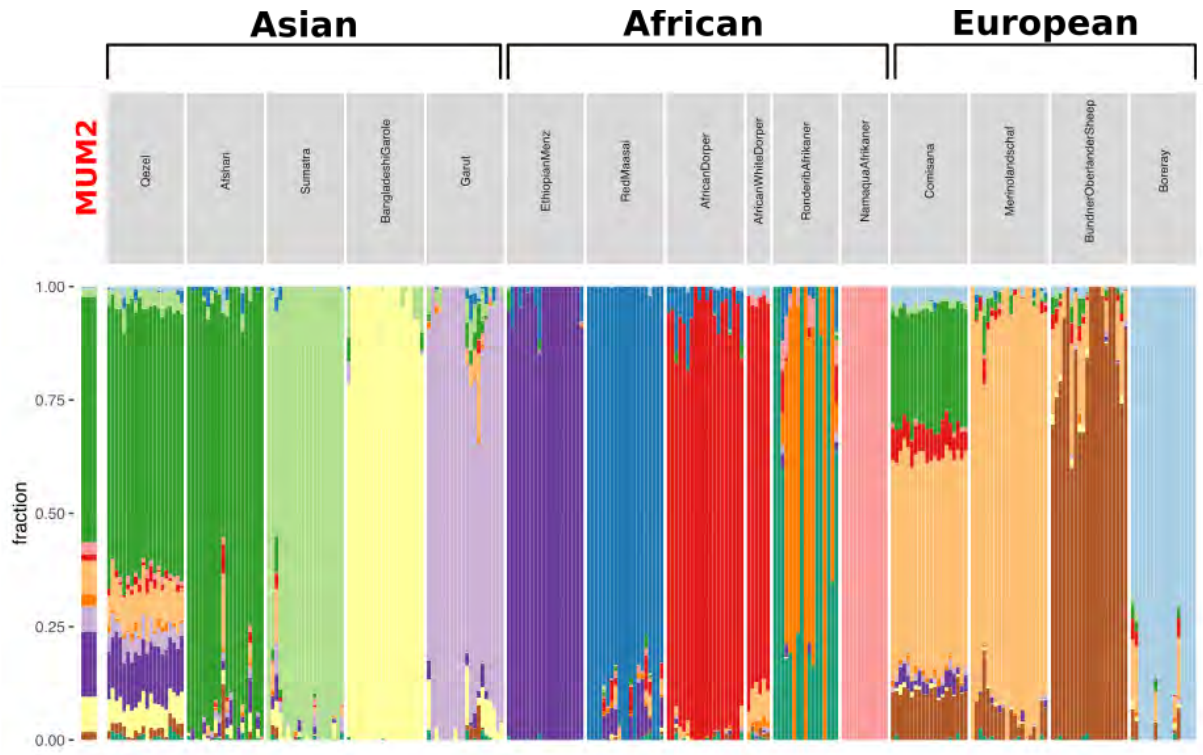
The genotyped breed dataset was downsampled using PLINK v1.9 to include approximately 90 individuals from each Asia, Europe and Africa. Using PLINK v1.9, a filter of 0 for missingness was applied to this merged dataset. The dataset was then filtered for linkage disequilibrium (LD) between markers, with the command --indep-pairwise 50 5 0.5. Finally, transitions were removed using PLINK v1.9 and VCFtools [57]. This fully filtered dataset contained 31,261 sites.

Unsupervised ADMIXTURE was explored with a range of hypothetical ancestral populations, *K*s (2-20). ADMIXTURE runs were replicated three times for each K and set a different

random seed for each. The three replicated runs for each K were averaged, and the cross-validation values which describe the approximate error of each ADMIXTURE run were inspected [58]. The run with the lowest cross-validation error was selected for presentation, *K*=13. In this ADMIXTURE analysis, the predicted ancestral profile of MUM2 most resembles Qezel, a sheep breed local to Iran.



Supplementary Figure 11: Average cross-validation values of unsupervised ADMIXTURE for each *K*, hypothetical ancestral population.



Supplementary Figure 12: ADMIXTURE plot with K=13 ancestral components. MUM2 displays a similar ancestral profile to modern Iranian sheep, represented by the Qezel breed.

Modified Genome Assembly

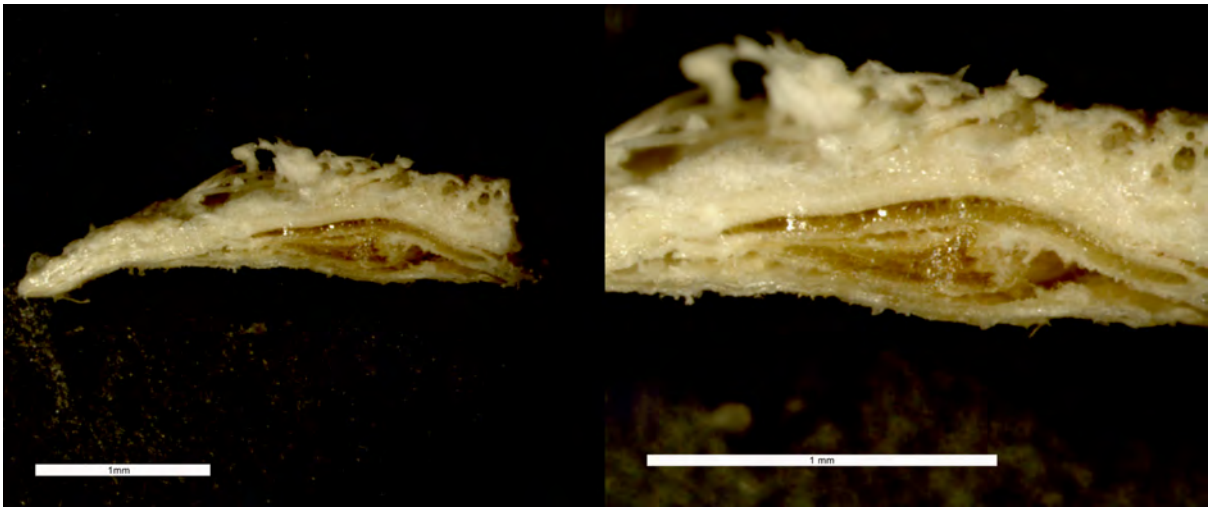
To investigate the woolly locus, two modified OviAri3.1 genome assemblies were produced. To exactly reproduce the locus detailed in [59], the sequence: 'ACTTCTGTAAAATGCAAGATAAATTAAAGTTATTATAACAGTGATTCTTTCAAAAAAATAAAA ACACCATGA' was inserted at position chr25:7,452,231/7,452,232. This modified genome represents the woolly locus. To construct the ancestral locus (i.e. with no asEIFS insertion in chromosome 25), this assembly was edited to replace sequence chr25:7,450,861-7,452,231 with 'TAAAAACACCATGA', as per [59].

All UDG-treated sequencing data was aligned to these modified assemblies in an identical manner as above to produce two bam files. The coverage at the woolly locus was visualised using Integrative Genomics Viewer (IGV) [60], before and after a MAQ filter of 30 was applied. Since there are three copies of EIF2S2 genes annotated on the OviAri3.1 genome assembly (One genuine EIF2S2 gene located on chromosome 13 and two EIF2S2 pseudogenes located on chromosome 7 and chromosome 25), the mapping quality of reads aligned to the insertion are 0 [59].

The 'breakpoints' of the woolly locus were defined at positions chr25:7,450,874/5 and chr25:7,452,217/8 of the modified oviAri3.1 assembly. Any reads which straddled either of these sites were deemed evidence of a woolly allele. The 'breakpoint' of the ancestral locus was chr25:7,450,861/2. Similarly any read that straddled this point was deemed evidence of the hairy/ancestral allele. Any read attributed to one genotype must only align to one modified assembly and not the other.

Microscopy analysis

Images produced by reflected light and transmitted light microscopy reveal a fat layer between the grain and corium, which is characteristic of domestic sheep skin [61,62] (Supplementary Figure 13). The fibres appear mostly unpigmented, with sporadic dark fibres visible (Supplementary Figure 14).

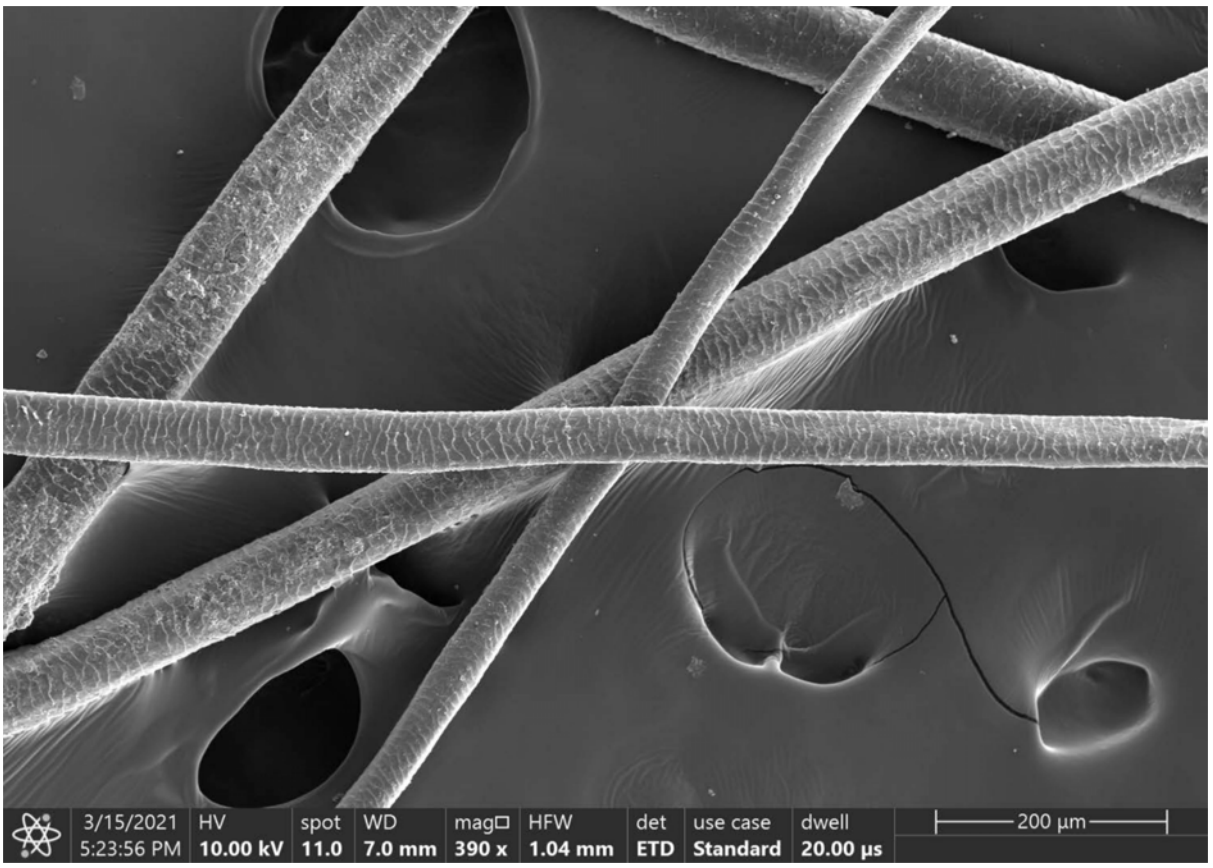


Supplementary Figure 13: Cross section of the sheep skin. Fat is deposited between the corium layer and the grain. Images by G. Ruß-Popa.

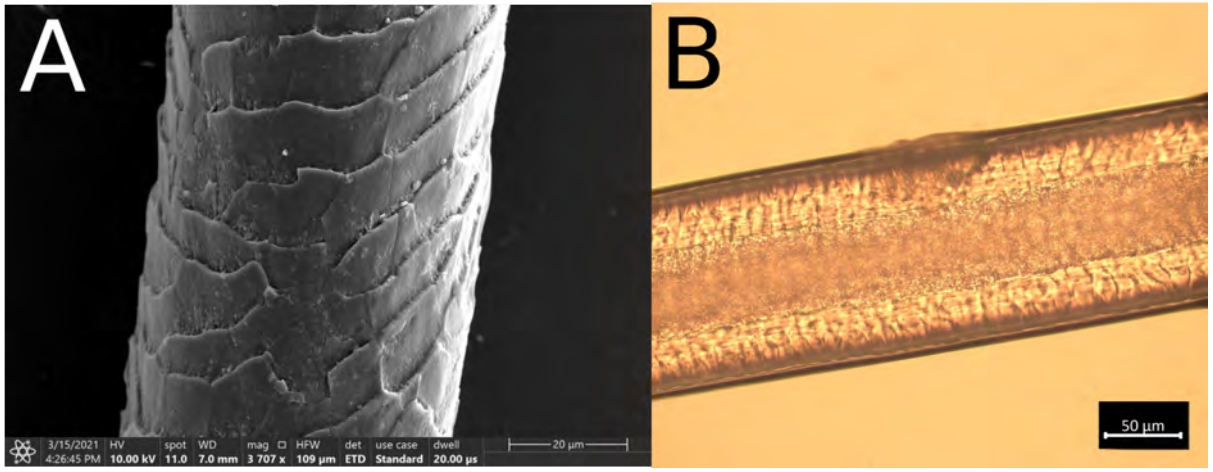
Supplementary Figure 14: Microscopy of fibres reveal a mostly unpigmented phenotype. Image by G. Ruß-Popa.

The leg of the sheep is dominated by coarse fibres (primary fibres) (Supplementary Figure 15). Scanning Electron Microscope (SEM) imagery of these hair fibres showed typical structure of sheep hair fibres, with a mosaic type scale and the fine lines on the scale's surface (Supplementary Figure 16A). This is characteristic of sheep hair fibres, and particularly for mouflon and medium-wool breeds. Light microscopy of these coarse fibres revealed an amorphous medulla (Supplementary Figure 16B). Thin unmedullated fibres (underwool) were also detected (Supplementary Figure 17).

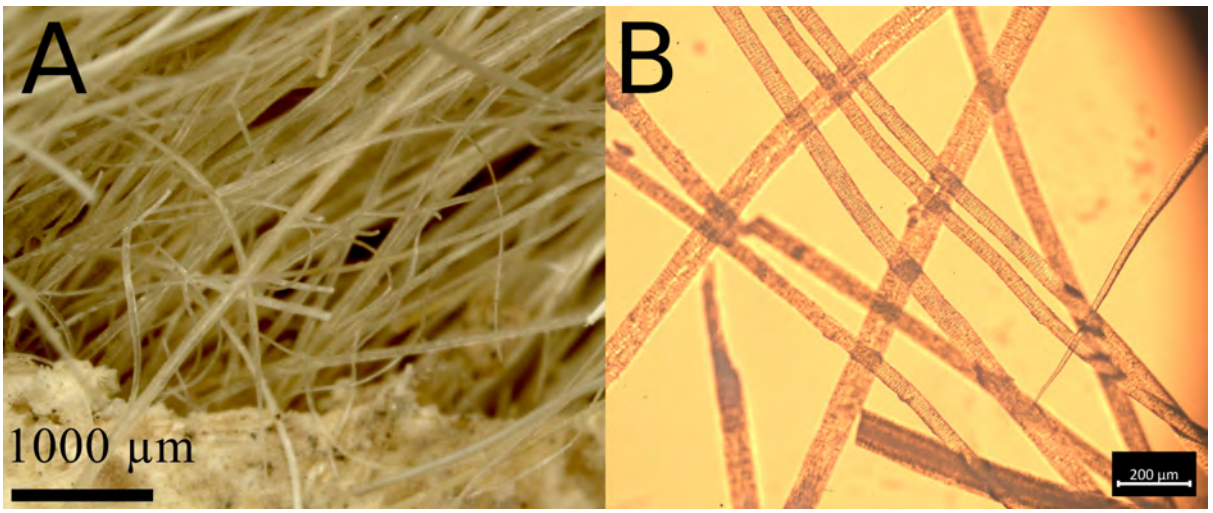
Overall, these results are consistent with MUM2 having a hairy/medium-wool phenotype, although results are not conclusive as the fibre types of the lower leg do not reflect the fleece type from the sheep body. Therefore, comparative analysis like wool measurements [63] of different sheep types are planned, in which samples from all parts of the body, including the lower leg, are to be examined.



Supplementary Figure 15: SEM image of different mummy hair fibres. Image by A. Steiger-Thirsfeld and G. Ruß-Popa.



Supplementary Figure 16A: SEM image of MUM2 hair fibre, displaying the typical mosaic scales of a sheep hair shaft and details of the fine lines on the scale surface. Image by A. Steiger-Thirsfeld and G. Ruß-Popa. **16B Light microscopy of the coarse hair fibre displaying the amorphous medulla.** Image by G. Ruß-Popa.



Supplementary Figure 17A: Microscopy displaying the thin fibres in front. Image by G. Ruß-Popa. **17B: Light microscopy revealing the unmedullated underwool.** Image by G. Ruß-Popa.

SEM images were generated by G. Ruß-Popa and Andreas Steiger-Thirsfeld at USTEM (Universitäre Service Einrichtung für Transmissions-Elektronenmikroskopie), TU Vienna. FEI Quanta 250 FEGSEM with AFM for Structural investigations (EBSD) was used for SEM imagery. Microscopy was done by G. Ruß-Popa at the Austrian Academy of Sciences, Austrian Archaeological Institute, Archaeological Sciences, Hollandstraße 11-13, 1020 Vienna, Austria. Zeiss SteREO Discovery.V20 (Magnification: 1–150 X) was used for Stereomicroscope (reflected light) analysis. Zeiss Axio Scope.A1 Pol. (Magnification: 1 X – 2.5 X – 5 X – 10 X – 20 X – 40 X) was used for the polarized light microscope analysis for the thin section. Axiocam 305 color was the microscope camera.

Genotyping the derived fat-tail allele

The pileup tool in GATK [31] v3.7 was used to report sample base calls at 51 SNP sites defined that are highly differentiated between fat-tailed and thin-tailed breeds, and thought to underlie the fat-tail phenotype [64]. The derived fat-tail alleles were determined using genotypes of the 29 diverse breeds of fat and thin-tailed breeds. These breeds were genotyped as part of the International Sheep Genomics Consortium (ISGC). Genotypic data were retrieved from [65], with methods detailed therein. The average genome coverage of MUM2 (~4X) was too low for accurate diploid genotype calling at these sites. Instead, we report a simple presence/absence of alleles based on base calls (Supplementary Table 6).

14 sites contained both the ancestral and derived allele base calls. Of these sites, 11 are susceptible to misdiagnosis due to deamination-induced base transitions. However, we also note the very low postmortem deamination levels of MUM2. One site (ss1139270994,ss1215832245,ss1198779144) contained one base call of a third allele (G) but was otherwise homozygous for the derived fat-tail allele (T) with four base calls.

Supplementary References

1. Aali A, Stöllner T. 2015 *The Archaeology of the Salt Miners: Interdisciplinary Research 2010-2014*. Bochum: Deutsches Bergbau-Museum.

2. Ramsey CB. 2009 Bayesian Analysis of Radiocarbon Dates. *Radiocarbon*. **51**, 337–360. (doi:10.1017/s0033822200033865)

3. Reimer PJ *et al.* 2013 IntCal13 and Marine13 Radiocarbon Age Calibration Curves 0–50,000 Years cal BP. *Radiocarbon* **55**, 1869–1887.

4. Dyson RH. 1989 The Iron Age Architecture at Hasanlu: An Essay. *Philadelphia* **31**, 107.

5. Muscarella OW. 2006 The Excavation of Hasanlu: An Archaeological Evaluation. *Bulletin of the American Schools of Oriental Research*. **342**, 69–94. (doi:10.1086/basor25066953)

6. Danti MD. 2013 The Late Bronze and Early Iron Age in Northwestern Iran. In *The Oxford Handbook of Ancient Iran* (ed DT Potts), Oxford University Press.

7. Dyson RH, Muscarella OW. 1989 Constructing the Chronology and Historical Implications of Hasanlu IV. *Iran* **27**, 1–27.

8. Danti MD. 2013 Hasanlu V: the Late Bronze and Iron I periods (Hasanlu Excavation Reports 3). *Philadelphia: University of Pennsylvania Museum*

9. Kroll S. 2013 Hasanlu Period III--annotations and corrections. *Iranica Antiqua* **48**, 175–195.

10. Davoudi H. 2017 Subsistence Economy of Iron Age Societies in North-Western Iran Based on Archaeozoological Studies-the Case of Tepe Hasanlu. PhD, Tarbiat Modares University, Iran.

11. Davoudi H, Mashkour M. 2019 Subsistence Economy in Northwestern Iran during Bronze and Iron Ages through Archaeozoological Researches at Tepe Hasanlu. In *Proceedings of the international conference on The Iron Age in Western Iran and Neibouring Regions*, pp. 484–520. RICHT, National Museum of Iran and Kurdistan Province ICHHTO.

12. Lorzadeh Z, Laleh H. 2018 Agricultural Landscape of Nishapur and the Hinterland of a Metropole. *Journal of Islamic Archaeology* **5**, 3–16.

13. Upton JM, Wilkinson CK. 1936 The Persian Expedition 1934-1935: Excavations at Nishapur. *The Metropolitan Museum of Art Bulletin*. **31**, 171. (doi:10.2307/3256246)

14. Dimand MS, Wilkinson CK. 1937 The Iranian Expedition, 1936: The Excavations at Nishapur. *The Metropolitan Museum of Art Bulletin*. **32**, 1. (doi:10.2307/3255486)

15. Dimand MS, Hauser W, Upton JM, Wilkinson CK. 1938 The Iranian Expedition, 1937: The Museum's Excavations at Nishapur. *The Metropolitan Museum of Art*

Bulletin. **33**, 1. (doi:10.2307/3256230)

16. 1942 The Museum's Excavations at Nīshāpūr by Walter Hauser and Charles K. Wilkinson. *The Metropolitan Museum of Art Bulletin* **37**, 83–119.

17. King GRD. 1989 Charles K. Wilkinson, Nishapur: Some Early Islamic Buildings and Their Decoration (New York: The Metropolitan Museum of Art, 1987). Pp. 328. *International Journal of Middle East Studies*. **21**, 283–285. (doi:10.1017/s0020743800032475)

18. Labbaf Khaniki R, Kervran M. 2005-2006 Nishapur 2005-2006, Mission archéologique irano-française. *Unpublished Report, Tehran, ICCHTO*

19. Labbaf Khaniki R. 2006 Archaeological Excavations in Nishapur's Citadel 2005-2006. *Unpublished Report, 2 Vols., Tehran, ICHTO*

20. Laleh H, Lorzadeh Z. 2019 Archaeological Research in the Nishapur Urban Historic Landscape (First Season). In *Proceedings of the 17th Annual Symposium on the Iranian Archaeology 2018-2019* (eds R Shirazi, S Hourshid), pp. 1087–1096. Tehran: Research Center for Cultural Heritage and Tourism Publications.

21. Khazaeli R. 2013 The Subsistence Economy of the Old Nishapur from the Formation of the City up to the Mongol Period. MA, University of Tehran.

22. Sambrook J, Fritsch EF, Maniatis T. 1989 *Molecular Cloning: A Laboratory Manual*.

23. Briggs AW, Stenzel U, Meyer M, Krause J, Kircher M, Pääbo S. 2010 Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. *Nucleic Acids Res.* **38**, e87.

24. Meyer M, Kircher M. 2010 Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* **2010**, db.prot5448.

25. Gamba C *et al.* 2014 Genome flux and stasis in a five millennium transect of European prehistory. *Nat. Commun.* **5**, 1–9.

26. Andrews S. 2010 FastQC. *https://www.bioinformatics.babraham.ac.uk/projects/fastqc/*.

27. Martin M. 2011 Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12.

28. Schubert M, Lindgreen S, Orlando L. 2016 AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes* **9**, 1–7.

29. Li H, Durbin R. 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760.

30. Li H *et al.* 2009 The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079.

31. McKenna A *et al.* 2010 The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303.

32. Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L. 2013 mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684.

33. Okonechnikov K, Conesa A, García-Alcalde F. 2016 Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* **32**, 292–294.

34. Sawyer S, Krause J, Guschanski K, Savolainen V, Pääbo S. 2012 Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS One* **7**, e34131.

35. Schmieder R, Edwards R. 2011 Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**, 863–864.

36. Wood DE, Lu J, Langmead B. 2019 Improved metagenomic analysis with Kraken 2. *Genome Biol.* **20**, 257.

37. Lu J, Breitwieser FP, Thielen P, Salzberg SL. 2017 Bracken: estimating species abundance in metagenomics data. *PeerJ Comput. Sci.* **3**, e104.

38. O'Sullivan NJ, Teasdale MD, Mattiangeli V, Maixner F, Pinhasi R, Bradley DG, Zink A. 2016 A whole mitochondria analysis of the Tyrolean Iceman's leather provides insights into the animal sources of Copper Age clothing. *Sci. Rep.* **6**, 31279.

39. Teasdale MD *et al.* 2017 The York Gospels: a 1000-year biological palimpsest. *R Soc Open Sci* **4**, 170988.

40. Knights D, Kuczynski J, Charlson ES, Zaneveld J, Mozer MC, Collman RG, Bushman FD, Knight R, Kelley ST. 2011 Bayesian community-wide culture-independent microbial source tracking. *Nat. Methods* **8**, 761–763.

41. Nayfach S, Rodriguez-Mueller B, Garud N, Pollard KS. 2016 An integrated metagenomics pipeline for strain profiling reveals novel patterns of bacterial transmission and biogeography. *Genome Res.* **26**, 1612–1625.

42. Velsko IM, Frantz LAF, Herbig A, Larson G, Warinner C. 2018 Selection of Appropriate Metagenome Taxonomic Classifiers for Ancient Microbiome Research. *mSystems* **3**. (doi:10.1128/mSystems.00080-18)

43. Warinner C, Herbig A, Mann A, Fellows Yates JA, Weiß CL, Burbano HA, Orlando L, Krause J. 2017 A Robust Framework for Microbial Archaeology. *Annu. Rev. Genomics Hum. Genet.* **18**, 321–356.

44. Korneliussen TS, Albrechtsen A, Nielsen R. 2014 ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* **15**, 356.

45. Meadows JRS, Hiendleder S, Kijas JW. 2011 Haplogroup relationships between

domestic and wild sheep resolved using a mitogenome panel. *Heredity* **106**, 700–706.

46. Edgar RC. 2004 MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797.

47. Darriba D, Taboada GL, Doallo R, Posada D. 2012 jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* **9**, 772–772.

48. Guindon S, Gascuel O. 2003 A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. *Syst. Biol.* **52**, 696–704.

49. Gouy M, Guindon S, Gascuel O. 2010 SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. *Mol. Biol. Evol.* **27**, 221–224.

50. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010 New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321.

51. Kijas JW *et al.* 2012 Genome-Wide Analysis of the World's Sheep Breeds Reveals High Levels of Historic Mixture and Strong Recent Selection. *PLoS Biol.* **10**, e1001258.

52. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. 2015 Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**. (doi:10.1186/s13742-015-0047-8)

53. Chaolong Wang Xiaowei Zhan Liming Liang Gonçalo R. Abecasis Xihong Lin. 2015 Improved Ancestry Estimation for both Genotyping and Sequencing Data using Projection Procrustes Analysis and Genotype Imputation. *Am. J. Hum. Genet.* **96**, 926–937.

54. Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. 2012 Ancient admixture in human history. *Genetics* **192**, 1065–1093.

55. Pebesma E. 2018 Simple Features for R: Standardized Support for Spatial Vector Data. *R J.* **10**, 439–446.

56. Pickrell JK, Pritchard JK. 2012 Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data. *PLoS Genet.* **8**, e1002967.

57. Danecek P *et al.* 2011 The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158.

58. Alexander DH, Lange K. 2011 Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics* **12**, 1–6.

59. Demars J *et al.* 2017 Genome-Wide Identification of the Mutation Underlying Fleece Variation and Discriminating Ancestral Hairy Species from Modern Woolly Sheep. *Mol. Biol. Evol.* **34**, 1722–1729.

60. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011 Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26.

61. Michel A. 2014 Skin deep: an outline of the structure of different skins and how it influences behaviour in use. A practitioner's guide. In *Why Leather?: The Material and Cultural Dimensions of Leather* (eds S Harris, AJ Veldmeijer), pp. 23–40. Leiden: Sidestone Press.

62. Ruß-Popa G. 2018 Der Gebrauch von Schaffell in der mitteleuropäischen urgeschichtlichen Bekleidung. *Annalen des Naturhistorischen Museums in Wien, Serie A* **120**, 157–176.

63. Rast-Eicher A, Jørgensen LB. 2013 Sheep wool in Bronze Age and Iron Age Europe. *Journal of Archaeological Science*. **40**, 1224–1241. (doi:10.1016/j.jas.2012.09.030)

64. Dong K *et al.* 2020 Genomic analysis of worldwide sheep breeds reveals PDGFD as a major target of fat-tail selection in sheep. *BMC Genomics* **21**, 800.

65. Taylor WTT *et al.* 2021 Evidence for early dispersal of domestic sheep into Central Asia. *Nat Hum Behav* (doi:10.1038/s41562-021-01083-y)