

# SCIENTIFIC REPORTS



OPEN

## A whole mitochondria analysis of the Tyrolean Iceman's leather provides insights into the animal sources of Copper Age clothing

Received: 27 January 2016

Accepted: 15 July 2016

Published: 18 August 2016

Niall J. O'Sullivan<sup>1,2</sup>, Matthew D. Teasdale<sup>3</sup>, Valeria Mattiangeli<sup>3</sup>, Frank Maixner<sup>1</sup>, Ron Pinhasi<sup>2</sup>, Daniel G. Bradley<sup>3</sup> & Albert Zink<sup>1</sup>

The attire of the Tyrolean Iceman, a 5,300-year-old natural mummy from the Ötztal Italian Alps, provides a surviving example of ancient manufacturing technologies. Research into his garments has however, been limited by ambiguity surrounding their source species. Here we present a targeted enrichment and sequencing of full mitochondrial genomes sampled from his clothes and quiver, which elucidates the species of production for nine fragments. Results indicate that the majority of the samples originate from domestic ungulate species (cattle, sheep and goat), whose recovered haplogroups are now at high frequency in today's domestic populations. Intriguingly, the hat and quiver samples were produced from wild species, brown bear and roe deer respectively. Combined, these results suggest that Copper Age populations made considered choices of clothing material from both the wild and domestic populations available to them. Moreover, these results show the potential for the recovery of complete mitochondrial genomes from degraded prehistoric artefacts.

The Tyrolean Iceman, a 5,300-year-old (Copper Age) natural mummy, discovered in the Italian Ötztal Alps in 1991 provides a direct archaeological link to prehistoric Europe. Two decades of analysis concerning this individual have provided insights into ancestry, diet, tools, lifestyle, health and attire<sup>1–3</sup>. Despite multiple studies and their relatively good preservation, ambiguity still persists as to the species of origin for the majority of the Iceman's clothes<sup>4,5</sup>. A more complete characterisation of his garments is therefore required, if we are to maximise the information provided by these artefacts and gain further insights into the choice of raw material from which Copper Age populations manufactured their clothes.

Preserved leathers provide rare and valuable information into how ancient populations utilized the secondary products of animal husbandry<sup>6,7</sup>. To date biomolecular research on the Iceman's leather has been confounded by the relatively high decomposition of the material<sup>3</sup>. The structural features of the leather and fur necessary for microscopic identification, such as grain pattern, are either damaged or absent from the clothing<sup>4</sup>. Proteomic and microscopic analyses have managed to some extent to overcome these limitations<sup>4,5</sup>, allowing for a species level resolution of some material, but ambiguity still remains in regard to others. For instance, it has been reported that a peptide mass fingerprinting (PMFs) analysis of the clothes, based on keratins and collagen recovered from the ancient leather, provided insufficient data to differentiate between closely related species such as goat and sheep<sup>6,7</sup>. Analysis of genetic data is therefore critical if we are to distinguish leather sourced from domesticated species from those produced from closely related wild relatives<sup>8</sup>. Furthermore, populations of the same domestic species can be differentiated by genetic markers in the mitogenome (haplogroups) that indicate independent domestications and human-mediated migrations, observing these haplogroups can highlight the genetic origin of the population used in clothing manufacture<sup>9</sup>.

Previous PCR work targeting ancient DNA (aDNA) fragments extracted from leather and furs have reported, in most cases, poor endogenous content<sup>10</sup>. This lack of such aDNA (apart from time dependent decomposition) has been attributed to the deterioration and removal of DNA molecules during the manufacturing process, which

<sup>1</sup>Institute for Mummies and the Iceman, EURAC research, 39100 Bolzano, Italy. <sup>2</sup>School of Archaeology and Earth Institute, University College Dublin, Belfield, Dublin 4, Ireland. <sup>3</sup>Smurfit Institute of Genetics, University of Dublin, Trinity College, Dublin 2, Ireland. Correspondence and requests for materials should be addressed to N.J.O. (email: niall.o-sullivan@ucdconnect.ie)

Sample	Species <sup>a</sup>	Haplo- groups <sup>b</sup>	Reads shotgun (#)	Mapped unique reads shotgun (#)	Reads capture (#)	Mapped unique reads capture (#)	Enrichment fold <sup>c</sup>	Coverage (X)	Contamination (%)
Coat (10304)	Sheep	B1	824,279	14	1,739,134	7,028	238	20.6	0.43
Dark coat (OCD)	Sheep	B1	723,348	6	2,422,032	5,141	256	15.5	1.32
Light coat (OCL2)*	Sheep	B1	533,933	5	594,102	1,197	215	3.6	3.97
Loincloth (OLC)	Sheep	B1	587,776	32	1,287,892	14,879	212	48.3	1.58
Coat (O314)	Goat	A	870,940	2,763	1,993,094	29,521	5	104.6	0.38
Leggings (OL)	Goat	A	962,864	102	4,900,933	30,502	59	110	0.54
Shoelace (OSA)	Cattle	T3	13,234,516	2,169	72,794	7,398	620	22.4	4.76
Quiver (OCA)*	Roe Deer	Central European	4,947,201	194	364,953	822	57	2.0	NA
Hat (OH)*	Brown bear	Western Lineage	NA	NA	21,672	151	NA	0.4	NA

**Table 1. Summary statistics and identification of sequenced libraries of this study.** \*Indicates partially reconstructed mitogenomes. <sup>a</sup>Fastq screen species identification. <sup>b</sup>The haplogroups were identified for each domestic species using the MitoToolPy tool and the observation of previously defined haplotypes for domesticated species. Wild species clades were assigned based on how they clustered in their respective phylogenies (Fig. 2). <sup>c</sup>Enrichment fold roughly estimates the success of targeted enrichment. It is based on a downsampling of shotgun and capture reads to a comparable number. The downsampled number of captured reads is divided by the number of shotgun reads. Two exceptions are for sample OCA and OSA in which the shotgun libraries were more abundant and hence were downsampled rather than the capture libraries.

is thought to involve scraping, exposing to fatty acids and, in some cases, intense heating<sup>11,12</sup>. In addition, the ancient tanning process is believed to introduce contamination from handling by humans or by the addition of organic material from other mammalian species<sup>5</sup>. Finally, the restoration process or the freeze-drying of the leather may have leached away identifying biological markers<sup>11</sup>. Recently, PCR based genetic analyses of hair shafts, recovered from more recent excavations of the Iceman find site have succeeded in further characterising the clothes; providing a species identification and even comparative phylogenies with modern populations<sup>13,14</sup>. These samples are however, unfortunately not directly associated with particular parts of the Iceman's clothing, and as such lack this critical piece of information for the evaluation of material usage by Copper Age individuals.

Conventional PCR based sanger sequencing of aDNA is further restricted by low endogenous DNA content, small fragment sizes and the co-extraction of contaminating molecules that are typical of ancient material<sup>15,16</sup>. The innovations of high throughput sequencing and targeted enrichment of specific DNA markers have overcome many of the obstacles faced by researchers in the study of DNA extracted from such archaeological materials<sup>17</sup>, and in short have led to the identification and characterisation of a wide range of informative genetic markers from ancient human and animal populations<sup>18–28</sup>. The study of ancient animal genetic variation can uncover novel variety or distributions of ancient populations that would not be possible solely from the analysis of modern genetic data, as intensive selective pressure (especially for domestics in the past 200 years<sup>29</sup>) and migrations have shuffled European phylogeography<sup>30</sup>. This has shown that contemporary genetic distributions of European livestock are not necessarily representative of ancient distributions, such as during the Copper Age or Neolithic<sup>31</sup>. In this paper, we report a targeted mitochondrial enrichment of DNA extracted from six articles of the Iceman's personal effects: shoelace, hat, loincloth, coat, leggings, and quiver. We then use these data to provide a discussion of Copper Age leather production and to investigate the relationship between the ancient animals and their current day populations.

## Results

**DNA enrichment, quality and authentication.** The sequencing libraries of the Iceman's clothes were subjected to both a low coverage shotgun sequencing analysis and mtDNA capture (targeted enrichment), using the Illumina MiSeq platform. To assess the quality of the extracted DNA and to measure the success of the mitochondrial capture, the mapped reads from the shotgun and capture libraries were quality controlled and quantified (Table 1 and Supplementary Table 2). Firstly, the libraries were shotgun sequenced (except for the hat sample, which was only sequenced after enrichment) and the presence of species-specific markers was ascertained. However, shotgun depth of coverage was too low for the majority of samples to reconstruct complete mitogenomes or recover enough informative autosomal markers, as most reads in sequenced libraries (often >99%) were exogenous. Hence, deeper shotgun sequencing to recover complete nuclear or mitochondrial genomes was not economically feasible. To overcome this situation, the samples were enriched for mitochondrial templates, which allowed for a four to 620-fold increase in high-quality unique sequences mapping to the mitochondrial reference. Fold-increase was based upon a down-sampling of captured or shotgun libraries to a similar number of reads depending on the depth of sequencing, after PCR-duplicates were removed. With the exception of the hat, quiver and light coat (0.4x–3.6x mean coverage), there was sufficient coverage to reconstruct complete mitogenomes for six of the libraries (15.5x–110x mean coverage). A positive and unambiguous species assignment for all enriched libraries was made regardless of coverage. For the six highly covered libraries a mapDamage<sup>32</sup> analysis (Supplementary Fig. 1) of sequence data showed damage patterns indicative of antiquity. However, this authenticity check is only optimal for relatively high coverage data, as the signal to noise ratio

was observed to decline with increased coverage. This is shown by the unclear damage patterns of the three low covered libraries (Supplementary Fig. 2). Cross-contamination estimates from multiple animal sources were low (<5%) (Table 1). These values were based on the average number of all mismatches at haplogroup defining haplotypes (Supplementary Table 3). Sequence depth of cover was visualized using qualimap software (bamqc)<sup>33</sup> and displayed no disproportionately low coverage that could disrupt consensus calls (Supplementary Figs 11–19) for the six highly covered libraries.

**Phylogeny and haplogroup assignment.** The haplogroups and haplotypes of the mitochondrial genomes sequenced from libraries produced from the leather of domestic species were identified with the aid of MitotoolPy<sup>34</sup> tool and visually inspected using IGVtools<sup>35</sup> (Table 1). All reported mutations are relative to the reference sequences (Supplementary Table 1 and Supplementary Table 3). Phylogenetic reconstructions of completely and partially reconstructed mitochondria were produced for all libraries (Figs 1 and 2, respectively).

The shoelace clustered to haplogroup T3 taurine lineage of cattle (Fig. 1a). The loincloth, dark fur coat and coat (sample 10304) libraries clustered inside sheep haplogroup B1, closer to modern domestics than mouflon (*Ovis orientalis*) (Fig. 1b). The coat sample (O314), and the leggings were identified as goat haplogroup A (Fig. 1c). There was sufficient mitochondrial genetic diversity observed to show that the materials derived from both sheep and goat came from multiple individuals, there were at least four sheep and two goats used in the manufacture. Goat and sheep sequences clustered closely within their respective haplogroups (Figs 1 and 2). Visualisation and contamination estimates of sequences did not show a co-capture of goat nuclear copies of mitochondrial DNA (numt), which have confounded some reference datasets for goat<sup>36</sup>.

The three low coverage libraries provided sufficient markers for clade or haplogroup assignment, but care is needed in their interpretation as SNP verification was constrained by low coverage. The sheep light coat sample exemplifies this problem, as it had a substantially lower coverage than the other identified domestic mitogenomes. This meant that potential sequence errors could not be removed by applying a majority consensus call alone. Hence, the sheep light coat data was mapped at a higher stringency with BWA aln (argument: -n 0.1) to remove confounding sequences. This method removed the confounding data, on inspection with IGVtools these appeared to be mostly coding lesions caused by relatively high degradation, and produced a more conservative consensus sequence. The resulting phylogeny for the light coat had a shorter branch length but was more realistic, which clustered approximately within sheep haplogroup B1 (Fig. 2c). The light coat phylogeny was based on 15,444 bp (base pairs) retrieved sites out of the complete 16,616 bp reference. The hat was identified as brown bear clustering nearest the European Western lineage (Fig. 2a), based on a reduced dataset of 5,363 bp of retrieved sequence out of a complete genome of 16,870 bp. The quiver was roe deer, which clustered closest to the central European clade (Fig. 2b). The roe deer sequence consisted of 7,662 bp, reconstructed compared to the 16,358 bp long reference mitogenome.

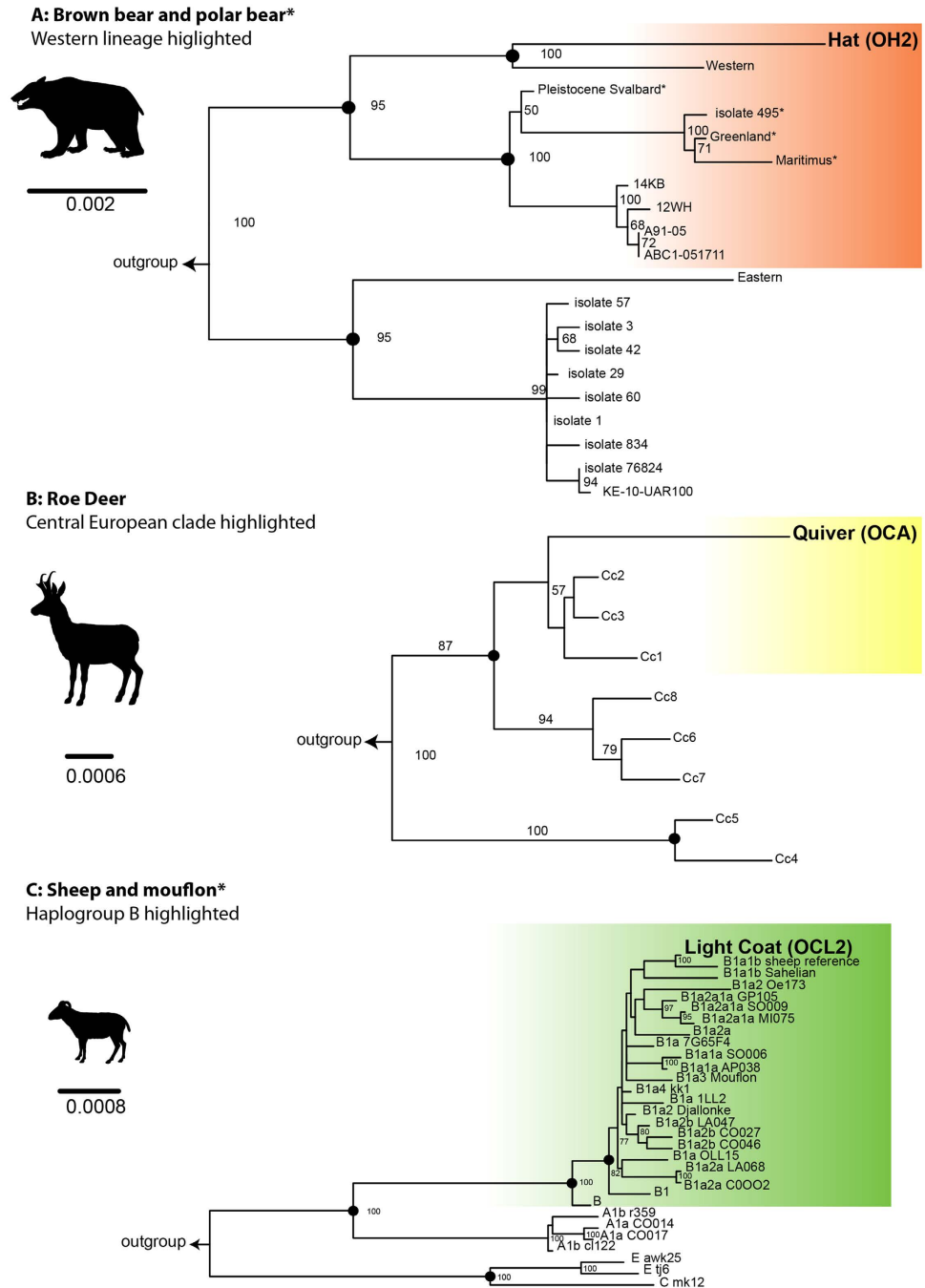
## Discussion

Here we present a comprehensive targeted enrichment and species identification of 5,300-year-old leather from the Iceman's clothes and quiver (Table 1). Of the nine samples taken, all produced libraries with sufficient data to identify the source species of the material. These were cattle, sheep, goat, brown bear and roe deer. Six libraries had sufficient coverage to fully reconstruct the mitogenome and confidently identify mtDNA haplotypes and reconstruct phylogeny (Fig. 1). Three libraries were of reduced coverage, thereby only partially reconstructing the mitogenomes and allowing for an approximated phylogeny (Fig. 2). Nonetheless, the identified cattle, sheep and goat sequences all fall within the range of mitochondrial genetic variability observed today in modern European domestic populations<sup>36–38</sup>. The mitogenomic genetic variability recovered shows that the sheep and goat leather fragments came from multiple hides. The genetic haplotypes of the wild species both roe deer and brown bear are consistent with present day phylogeography in the Alpine region<sup>39,40</sup>.

The Mitochondrial capture analysis allowed for an effective increase in coverage (between a four and 620-fold increase in coverage) and highlighted the effectiveness of NGS assisted with mitochondrial targeted enrichment to retrieve residual endogenous DNA from suboptimal sources such as skin or tanned leather<sup>10,41</sup>. Furthermore, our shotgun analysis is the first NGS quantification of the residual endogenous DNA present in archaeological leather. These data show that the endogenous content was highly variable and considerably lower than that of tissue sampled from the Iceman for genomic reconstruction<sup>3</sup>. This result lends support to the theory that tanning, the restoration processes, or both can remove endogenous markers<sup>10,11</sup> and furthermore that the tanning may have exposed the material to a greater degree of bacterial and chemical degradation<sup>7</sup>. Equally, these results could support the hypothesis that soft tissue has a generally weaker retention of endogenous DNA compared to calcified remains<sup>41</sup>. Interestingly, enriched libraries displayed neither co-enrichment from other mammalian species or cross contamination from other animals in the mapped data (Supplementary Figs 2–10, Table 1). Hence, there is little genetic evidence from these samples to suggest that leathers and furs were treated with fat and oil extracted from other animals during the ancient tanning process<sup>11</sup>. Previous studies have however observed residual plant-based oils and fats in ancient leather indicating their use in leather manufacturing. Such as the identification of plant lipids in Neolithic leathers from the contemporaneous Schnidejoch Pass site in Switzerland<sup>7</sup>.

The Iceman's garments and quiver are from an assemblage of at least five different species of animal. The coat alone was a combination of at least four hides and two species: goat and sheep. This result may indicate a haphazard stitching together of clothing based upon materials that were available to the Iceman, as ancient rudimentary leather is posited to rapidly deteriorate after manufacture<sup>6,12</sup>. However, the leggings were composed of goat leather, which was also used in the manufacture of a 4,500-year-old leggings from Schnidejoch, Switzerland. This result lends support to the idea that Copper Age individuals in the Alpine region selected species for specific attributes when manufacturing clothing<sup>8</sup>. This may also indicate a functional choice of material based on flexibility or insulating potential.





**Figure 2. Species phylogeny based on partially reconstructed mitogenomes.** Aligned sequences had all gap sites removed with Gblocks, except the sheep light coat (OCL2). Asterisks on branch labels of trees (A) and (C) correspond to clade/haplogroup names where inter-subspecies mono-phylogeny occur. Outgroups used to root trees have been removed (Supplementary Table 1). Black circles on nodes indicate that there was more than 90% bootstrap support in neighbour joining, maximum parsimony and maximum likelihood topologies. Coloured boxes highlight associated haplogroups of samples. Animal illustrations provided by Matthew Collins (*pers coms*).

The haplogroups identified in these samples from sheep, goat and cattle all fall within the domestic variability of the source animal in modern Europe, novel SNPs have been recovered that may have been lost from contemporary domestic haplogroups (Supplementary Table 3). In this study, the goat mitogenomes were identified as haplogroup A, yet a goat B1 haplogroup was observed in a study of 4,500-year-old swiss leggings<sup>8</sup>. The ancient distribution of domestic goat haplogroups are unknown, yet B1's modern European distribution is restricted to Greece and Turkey<sup>8,42</sup>. Goat haplogroup B1 is more widespread in western Asia, but is closely related to haplogroup A. The Iceman and his leggings predate the Swiss find by c. 800 years, suggesting an overlap in the distribution of goat haplogroups that is no longer present in Europe. The sheep leather and fur were identified within



haplogroup B1 from the three coat fragments and the loincloth. One sheep (coat 10304) mitogenome has the same distinct mutations at loci 16147 and 16440 reported by Olivieri *et al.*<sup>13</sup> in a PCR analysis of fibres from a later excavation the Iceman's find site. However, other mutations reported by Olivieri *et al.*<sup>13</sup> were not present indicating that they identified sequences from yet another sheep used in the garment assemblage and recovered from the find site. The shoelace mitogenome closely resembled the cattle reference sequence of haplogroup T3, which is the predominant mitochondrial haplotype in European domestic cattle and is observed in high frequency throughout the Neolithic of Europe<sup>43</sup>.

The quiver was identified as roe deer, clustering near the contemporary central European clade. This is the first time this species has been identified as being part of the Iceman's leather attire. Notably, red deer (*Cervus elaphus*) was identified from PCR of hair shafts at the finding site<sup>14</sup> and contents of the Iceman's stomach<sup>44</sup>, and from PMFs of the quiver<sup>5</sup>. The evidence that the Iceman utilised at least two deer species shows the variety of wild materials from which he subsisted. The hat sequence clusters inside the Western lineage of brown bears, suggesting regional continuity of this genetic marker. Previous studies of brown bears report complex phylogeography across its range, typified by overlapping ranges of divergent clades of brown bear and introgression with polar bears (*Ursus maritimus*) due to convoluted gene flow caused by Pleistocene climatic changes<sup>40,45</sup>. The assignment of the hat to brown bear is important given that previous studies could not resolve its taxonomy beyond *Carnivora*<sup>5</sup>, and secondly that it is unambiguously a wild species that was used for clothing manufacture.

Despite the Iceman being established as an agro-pastoralist<sup>2,3</sup>, the hat and quiver provide evidence of hunting and trapping of wild animals. Other studies have proposed parts of the diet, clothing and equipment as belonging to wild species such as wild canid (for the hat, species previously unresolved), red deer and chamois<sup>5,14</sup>. Rollo *et al.*<sup>44</sup> demonstrated that the Iceman consumed red deer and alpine ibex (*Capra ibex*). The choices that Copper Age people made with respect to animals exploitation were likely dependent on availability, necessity, functionality and symbolism<sup>46</sup>. Given the diversity of the analysed material (each analysed fragment comes from a different animal), the lack of any obvious symbolism and the rapid degradation rate of ancient leather<sup>11</sup> it would appear that the Iceman's leathers and furs were first chosen haphazardly for subsistence as rudimentary leather pieces rapidly disintegrated and required replacement with strips of new hide. There is however likely to be a regional variation in the selection of wild and domestic populations throughout the Copper Age and the Iceman provides a window into the alpine region which may have been unlike other locations<sup>46</sup>.

These results are based on uni-parental maternal genetic markers, hence the authors cannot exclude the possibility that the analysed domestic species were sourced from herds that were re-stocked with wild males as occurred between cattle and aurochs populations as is postulated in ancient Britain<sup>31</sup>. Whole genome sequencing of the Iceman's attire in the future may increase our knowledge to the degree of introgression with local populations or breeding selection for functional genes.

## Materials and methods

**Ancient DNA extraction and library preparation.** DNA extractions were based on the protocol of Gamba *et al.*<sup>47</sup> and conducted in a dedicated ancient DNA laboratory in Trinity College Dublin. Barcoded Illumina NGS libraries were prepared following the protocol of Meyer and Kircher<sup>48</sup> with modifications by Gamba *et al.*<sup>47</sup>. The one exception was the hat, which was extracted and library prepared at the clean room facility at the EURAC - Institute for Mummies and the Iceman, Bolzano. The extraction protocol for the hat was adapted from Rohland *et al.*<sup>49</sup>. The hat library was subsequently indexed and captured at Trinity College Dublin.

**Mitochondrial DNA enrichment and Sequencing.** Eight DNA extracts from five articles of the Iceman's clothing and one from the quiver were selected for mtDNA enrichment. Firstly, libraries were shotgun sequenced (excluding the hat) on a Illumina MiSeq to check the quality of the extracted DNA and the percentage endogenous content. Prior to sequencing, the indexed libraries were pooled and enriched for mitochondrial genomes using a custom RNA capture produced from reference domestic mtDNA genomes by MYcorarray (Mycorarray, 5692 Plymouth Road, Ann Arbor, MI 48105) following the manufacturer's instructions. The shotgun and captured libraries were both sequenced on the Illumina MiSeq platform at TrinSeq St James's Hospital Dublin.

**Processing of raw data and mapping.** Universal Illumina adapters were trimmed from the 3' end of raw sequences using Clip&Merge<sup>50</sup> with a minimum read length of 30 base pairs (bp) and 2 bp clipped from both the 5' and 3' end termini. Fastq Screen ([http://www.bioinformatics.babraham.ac.uk/projects/fastq\\_screen](http://www.bioinformatics.babraham.ac.uk/projects/fastq_screen)) was used as a first-pass species identification using a cohort of reference animal mitogenomes (cattle, sheep, goat, deer, bear, human, ibex, pig, chamois and dog) obtained from the refseq database at NCBI<sup>51</sup> (Supplementary Table 4). Fastq Screen displays which reference the sequenced libraries preferentially map to and have the most species specific reads. Reads were then re-mapped to the most similar reference mitogenome with BWA<sup>52</sup> using standard parameters for aDNA<sup>53</sup>. For each mitochondrial reference (Supplementary Table 1), the first 30bp were manually copied and pasted to the end of the reference to compensate for the circular nature of the mitochondrial genome<sup>54</sup>. PCR duplicate (clonal) reads were removed using SAMtools<sup>55</sup> and ancient DNA authenticity was assessed using mapDamage<sup>32</sup> (analysing adapter removed but unclipped data). MapDamage was further used to rescale the base quality of putatively damaged bases i.e. G to A or C to T transitions approaching the 5' and 3' termini. With one exception for sheep whose mtDNA genome contains a repeat expansion<sup>38</sup>, all reads were filtered for a minimum mapping quality of 30 with the SAMtools view<sup>52</sup>. Rough mitochondrial cross-contamination estimates from other animals were made following the methods adapted from Sánchez-Quinto *et al.*<sup>56</sup> and Teasdale *et al.*<sup>20</sup> analysing mismatches at haplogroup defining positions.

**Phylogenetic analysis and haplogroup assignment of samples.** A consensus fasta sequence was generated for each sample using ANGSD (analysis of next generation sequencing data) tool<sup>57</sup>, filtering out base

quality below 30 and making a majority consensus call from overlapping reads. Bam files were visualised with IGV<sup>35</sup> during construction of fasta consensus sequences to investigate if any polymorphisms in the consensus were confounded by low coverage (less than 3 times covered) and to identify INDELS that may have been overlooked during consensus calling.

Reference mitochondrial sequences were acquired from the NCBI database, where possible using a RefSeq sequence. The sheep reference sequence AF010406, used by Dometree<sup>34</sup>, corresponds exactly to the sheep reference NC\_001941 from Hiendleder *et al.*<sup>58</sup>. The consensus sequences were aligned to a cohort of mitogenomes from NCBI (Supplementary Table 1) using MUSCLE<sup>59</sup> with maximum of two iterations, which represent modern diversity in each species and appropriate outgroups to root the phylogenetic trees, and where possible ancient data sequenced for that species (Supplementary Table 1). Neighbour-joining (NJ), maximum-likelihood (ML), and maximum-parsimony (MP) phylogenetic trees were constructed in Seaview<sup>60</sup>. The Jukes-Cantor substitution model, 100 bootstrap iterations and gaps ignored setting were used for neighbour-joining. The generalised time reversible (GTR) and 100 bootstraps settings were used for the maximum-likelihood trees. The MP used a less thorough tree search and 100 bootstraps. Ultimately, the ML trees were chosen and visualised using FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>). However, bootstrap values from NJ and MP above 90 were considered. Loci were excluded from the phylogenetic trees if they were cited in previous publications as confounding analysis or caused polyphylogeny among established haplogroups<sup>34,36,38–40,54</sup>.

Phylogenetic analysis for all species was based on an assemblage of mitogenomes within their respective genera. For sheep (*Ovis aries*) this was based on an assemblage of *Ovis* mitogenomes from Meadows *et al.*<sup>38</sup>, Dometree<sup>34</sup> and Lancioni *et al.*<sup>61</sup>, using the haplogroup nomenclature of the later. For cattle (*Bos taurus*) phylogeny was based on Dometree<sup>34</sup>, Bonfiglio *et al.*<sup>62</sup>, Zhang *et al.*<sup>54</sup> and Achilli *et al.*<sup>37</sup>. Goat (*Capra hircus aegagrus*) phylogenies were based on Dometree<sup>34</sup>, Colli *et al.*<sup>42</sup> and Hassanin *et al.*<sup>36</sup>. Mitogenomes identified as clustering within the major haplogroups of cattle, sheep and goat by Peng *et al.*<sup>34</sup> were added to the phylogenies to increase the resolution. Some of the nomenclatures for the sheep and goat subhaplogroups were updated from the dometree website using the above publications. Haplogroups and novel haplotypes for domestic species were identified by analysing the consensus sequences with MitoToolPy<sup>34</sup> tool against a dataset of established haplotypes for domestic animals<sup>34</sup>. In this study, haplogroup assignments incorporated the above publications for the inclusion of up-to-date haplotypes. Roe deer (*Capreolus capreolus*) sequence data were aligned against an assemblage of Eurasian *Capreolus* variety<sup>39</sup>. Brown bear (*Ursus arctos*) sequence data were aligned with an assemblage of European, Polar, Eurasian and North American variety<sup>40,45</sup>.

For sequenced libraries with insufficient coverage to fully reconstruct mitogenomes, the partial data was tentatively converted to a consensus sequence with gap sites. The sequences were then inspected using IGV (the integrative genome visualizer)<sup>35</sup>. The data were aligned using methods as above, in addition GBlocks<sup>63</sup> was used (except for Sample OCL2, which included gaps) to create a subset of aligned data that excluded all gaps from phylogenetic analysis.

## References

- Rollo, F., Ascii, W., Marota, I. & Sassaroli, S. In *Der Mann im Eis* 91–105 (Springer Vienna, 1995).
- Müller, W., Fricke, H., Halliday, A. N., McCulloch, M. T. & Wartho, J.-A. Origin and migration of the Alpine Iceman. *Science* **302**, 862–866 (2003).
- Keller, A. *et al.* New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. *Nat. Commun.* **3**, 698 (2012).
- Groenman-van Waateringe, W. Analyses of the hides and skins from the Hauslabjoch. *Jahrbuch des Römisch-Germanischen Zentralmuseums Mainz* **39**, 114–128 (1992).
- Hollemeyer, K., Altmeyer, W., Heinze, E. & Pitra, C. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry combined with multidimensional scaling, binary hierarchical cluster tree and selected diagnostic masses improves species identification of Neolithic keratin sequences from furs of the Tyrolean Iceman Oetzi. *Rapid Commun. Mass Spectrom.* **26**, 1735–1745 (2012).
- Brandt, L. Ø. *et al.* Species identification of archaeological skin objects from Danish bogs: comparison between mass spectrometry-based peptide sequencing and microscopy-based methods. *PLoS One* **9**, e106875 (2014).
- Spangenberg, J. E., Ferrer, M., Tschudin, P., Volken, M. & Hafner, A. Microstructural, chemical and isotopic evidence for the origin of late neolithic leather recovered from an ice field in the Swiss Alps. *J. Archaeol. Sci.* **37**, 1851–1865 (2010).
- Schlumbaum, A. *et al.* Ancient DNA, a Neolithic legging from the Swiss Alps and the early history of goat. *J. Archaeol. Sci.* **37**, 1247–1251 (2010).
- Larson, G. *et al.* Current perspectives and the future of domestication studies. *Proc. Natl. Acad. Sci. USA* **111**, 6139–6146 (2014).
- Vuissoz, A. *et al.* The survival of PCR-amplifiable DNA in cow leather. *J. Archaeol. Sci.* **34**, 823–829 (2007).
- Püntener, A. G. & Moss, S. Ötzi, the Iceman and his Leather Clothes. *CHIMIA International Journal for Chemistry* **64**, 315–320 (2010).
- Covington, A. D. Modern tanning chemistry. *Chem. Soc. Rev.* **26**, 111–126 (1997).
- Olivieri, C. *et al.* Phylogenetic position of a Copper Age sheep (*Ovis aries*) mitochondrial DNA. *PLoS One* **7**, e33792 (2012).
- Olivieri, C. *et al.* Positioning the red deer (*Cervus elaphus*) hunted by the Tyrolean Iceman into a mitochondrial DNA phylogeny. *PLoS One* **9**, e100136 (2014).
- Allentoft, M. E. *et al.* The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proc. Biol. Sci.* **279**, 4724–4733 (2012).
- Skoglund, P. *et al.* Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proc. Natl. Acad. Sci. USA* **111**, 2229–2234 (2014).
- Maricic, T., Whitten, M. & Pääbo, S. Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. *PLoS One* **5**, e14004 (2010).
- Immel, A. *et al.* Mitochondrial Genomes of Giant Deers Suggest their Late Survival in Central Europe. *Sci. Rep.* **5**, 10853 (2015).
- Skoglund, P. *et al.* Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science* **336**, 466–469 (2012).
- Teasdale, M. D. *et al.* Paging through history: parchment as a reservoir of ancient DNA for next generation sequencing. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **370**, 20130379 (2015).
- Meyer, M. *et al.* A mitochondrial genome sequence of a hominin from Sima de los Huesos. *Nature* **505**, 403–406, doi: 10.1038/nature12788 (2013).

22. Raghavan, M. *et al.* Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* **505**, 87–91 (2014).
23. Green, R. E. *et al.* A draft sequence of the Neandertal genome. *Science* **328**, 710–722 (2010).
24. Reich, D. *et al.* Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* **468**, 1053–1060 (2010).
25. Orlando, L. *et al.* Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* **499**, 74–78 (2013).
26. Knapp, M., Clarke, A. C., Horsburgh, K. A. & Matisoo-Smith, E. A. Setting the stage - building and working in an ancient DNA laboratory. *Ann. Anat.* **194**, 3–6 (2012).
27. Maixner, F. *et al.* The 5300-year-old *Helicobacter pylori* genome of the Iceman. *Science* **351**, 162–165 (2016).
28. Ermini, L. *et al.* Complete mitochondrial genome sequence of the Tyrolean Iceman. *Curr. Biol.* **18**, 1687–1693 (2008).
29. Kijas, J. W. *et al.* Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biol.* **10**, e1001258 (2012).
30. Larson, G. & Burger, J. A population genetics view of animal domestication. *Trends Genet.* **29**, 197–205 (2013).
31. Park, S. D. E. *et al.* Genome sequencing of the extinct Eurasian wild aurochs, *Bos primigenius*, illuminates the phylogeography and evolution of cattle. *Genome Biol.* **16**, 234 (2015).
32. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684 (2013).
33. García-Alcalde, F. *et al.* Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics* **28**, 2678–2679 (2012).
34. Peng, M.-S. *et al.* DomeTree: a canonical toolkit for mitochondrial DNA analyses in domesticated animals. *Mol. Ecol. Resour.* **15**, 1238–1242 (2015).
35. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
36. Hassanin, A., Bonillo, C., Nguyen, B. X. & Cruaud, C. Comparisons between mitochondrial genomes of domestic goat (*Capra hircus*) reveal the presence of numts and multiple sequencing errors. *Mitochondrial DNA* **21**, 68–76 (2010).
37. Achilli, A. *et al.* The multifaceted origin of taurine cattle reflected by the mitochondrial genome. *PLoS One* **4**, e5753 (2009).
38. Meadows, J. R. S., Hiendleder, S. & Kijas, J. W. Haplogroup relationships between domestic and wild sheep resolved using a mitogenome panel. *Heredity* **106**, 700–706 (2011).
39. Matosiuk, M., Sheremetyeva, I. N., Sheremetyev, I. S., Saveljev, A. P. & Borkowska, A. Evolutionary neutrality of mtDNA introgression: evidence from complete mitogenome analysis in roe deer. *J. Evol. Biol.* **27**, 2483–2494 (2014).
40. Hirata, D. *et al.* Molecular phylogeography of the brown bear (*Ursus arctos*) in Northeastern Asia based on analyses of complete mitochondrial DNA sequences. *Mol. Biol. Evol.* **30**, 1644–1652 (2013).
41. Seguin-Orlando, A. *et al.* Pros and cons of methylation-based enrichment methods for ancient DNA. *Sci. Rep.* **5**, 11826 (2015).
42. Colli, L. *et al.* Whole mitochondrial genomes unveil the impact of domestication on goat matrilineal variability. *BMC Genomics* **16**, 1115 (2015).
43. Scheu, A. *et al.* The genetic prehistory of domesticated cattle from their origin to the spread across Europe. *BMC Genet.* **16**, 54 (2015).
44. Rollo, F., Ubaldi, M., Ermini, L. & Marota, I. Otzi's last meals: DNA analysis of the intestinal content of the Neolithic glacier mummy from the Alps. *Proc. Natl. Acad. Sci. USA* **99**, 12594–12599 (2002).
45. Edwards, C. J. *et al.* Ancient hybridization and an Irish origin for the modern polar bear matriline. *Curr. Biol.* **21**, 1251–1258 (2011).
46. Doppler, T. *et al.* Landscape opening and herding strategies: Carbon isotope analyses of herbivore bone collagen from the Neolithic and Bronze Age lakeshore site of Zurich-Mozartstrasse, Switzerland. *Quat. Int.*, doi: 10.1016/j.quaint.2015.09.007 (2015).
47. Gamba, C. *et al.* Genome flux and stasis in a five millennium transect of European prehistory. *Nat. Commun.* **5**, 5257 (2014).
48. Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* **2010**, db.prot5448 (2010).
49. Rohland, N., Siedel, H. & Hofreiter, M. A rapid column-based ancient DNA extraction method for increased sample throughput. *Mol. Ecol. Resour.* **10**, 677–683 (2010).
50. Peltzer, A. *et al.* EAGER: efficient ancient genome reconstruction. *Genome Biol.* **17**, 60 (2016).
51. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
52. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
53. Schubert, M. *et al.* Improving ancient DNA read mapping against modern reference genomes. *BMC Genomics* **13**, 178 (2012).
54. Zhang, H. *et al.* Morphological and genetic evidence for early Holocene cattle management in northeastern China. *Nat. Commun.* **4**, 2755 (2013).
55. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
56. Sánchez-Quinto, F. *et al.* Genomic Affinities of Two 7,000-Year-Old Iberian Hunter-Gatherers. *Curr. Biol.* **22**, 1494–1499 (2012).
57. Korneliusson, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* **15**, 356 (2014).
58. Hiendleder, S., Mainz, K., Plante, Y. & Lewalski, H. Analysis of mitochondrial DNA indicates that domestic sheep are derived from two different ancestral maternal sources: no evidence for contributions from urial and argali sheep. *J. Hered.* **89**, 113–120 (1998).
59. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
60. Gouy, M., Guindon, S. & Gascuel, O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **27**, 221–224 (2010).
61. Lancioni, H. *et al.* Phylogenetic relationships of three Italian merino-derived sheep breeds evaluated through a complete mitogenome analysis. *PLoS One* **8**, e73712 (2013).
62. Bonfiglio, S. *et al.* The enigmatic origin of bovine mtDNA haplogroup R: sporadic interbreeding or an independent event of *Bos primigenius* domestication in Italy? *PLoS One* **5**, e15760 (2010).
63. Castresana, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17**, 540–552 (2000).

## Acknowledgements

F.M., N.J.O. and A.Z. acknowledge the funding provided by the South Tyrolean grant legge 14. M.D.T, V.M, and the ancient DNA work described were supported by European Research Council Investigator Grant 295729-CodeX to D.G.B. We also acknowledge Science Foundation Ireland (SFI) ERC Support Award 12/ERC/B2227. The authors wish to thank Prof Matthew Collins and his team at York University for providing us with sampling tools and mammalian figures outlines for phylogenetic trees. We are grateful to the Archaeological Museum of Bolzano for assistance in sampling.



### Author Contributions

A.Z., R.P. and D.G.B. supervised this study. A.Z. and F.M. provided samples. N.J.O., V.M. and M.D.T. extracted the DNA and prepared the libraries for sequencing. V.M. performed the sequencing. N.J.O., V.M., M.D.T. and F.M. analysed the genetic data. N.J.O., V.M. and M.D.T. wrote the manuscript with contributions from all authors.

### Additional Information

**Accession codes:** Sequence reads were uploaded to the Sequence Read Archive at the ENA (European Nucleotide Archive), study accession PRJEB13144.

**How to cite this article:** O'Sullivan, N. J. *et al.* 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 (2016).

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016