CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Capturing goats: documenting two hundred years of mitochondrial DNA diversity among goat populations from Britain and Ireland
Authors	Cassidy, Lara M.;Teasdale, Matthew D.;Carolan, Sean;Enright, Ruth;Werner, Raymond;Bradley, Daniel G.;Finlay, Emma K.;Mattiangeli, Valeria
Publication date	2017
Original Citation	Cassidy, L. M., Teasdale, M. D., Carolan, S., Enright, R., Werner, R., Bradley, D. G., Finlay, E. K. and Mattiangeli, V. (2017) 'Capturing goats: documenting two hundred years of mitochondrial DNA diversity among goat populations from Britain and Ireland', Biology Letters, 13(3), 20160876(6pp). doi: 10.1098/rsbl.2016.0876
Type of publication	Article (peer-reviewed)
Link to publisher's version	http://rsbl.royalsocietypublishing.org/content/13/3/20160876 - 10.1098/rsbl.2016.0876
Rights	© 2017, the Authors. Published by the Royal Society. All rights reserved. This is a post-peer-review, pre-copyedit version of an article published in Biology Letters. The definitive publisher-authenticated version [Cassidy, L. M., Teasdale, M. D., Carolan, S., Enright, R., Werner, R., Bradley, D. G., Finlay, E. K. and Mattiangeli, V. (2017) 'Capturing goats: documenting two hundred years of mitochondrial DNA diversity among goat populations from Britain and Ireland', Biology Letters, 13(3), 20160876(6pp). doi: 10.1098/rsbl.2016.0876] is available online at: https://dx.doi.org/10.1098/rsbl.2016.0876 - http://royalsocietypublishing.org/licence
Download date	2026-11-06 20:26:36
Item downloaded from	https://hdl.handle.net/10468/6523



University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Capturing goats: Documenting two hundred years of mitochondrial DNA diversity among goat populations from Britain and Ireland

Lara M. Cassidy[1], Matthew D. Teasdale[1], Sean Carolan[2], Ruth Enright[3], Raymond Werner[2], Daniel G. Bradley[1], Emma K. Finlay[4], Valeria Mattiangeli[1].

1 Smurfit Institute of Genetics, Trinity College Dublin, Ireland.

2 Old Irish Goat Society, Mulranny, Ireland.

3 School of Biological, Earth and Environmental Sciences, University College Cork, Ireland.

4 School of Biotechnology, Dublin City University, Dublin, Ireland

Abstract

Domestic goats (*Capra hircus*) play a key role in agriculture, being especially prised in regions of marginal pasture. However, the advent of industrialized breeding has seen a dramatic reduction in genetic diversity within this species, while rising extinction rates among feral herds have further depleted the reservoir of variation available. Here, we present the first survey of whole mitochondrial genomic variation among the modern and historical goat populations of Britain and Ireland using a combination of mtDNA enrichment and high throughput sequencing (HTS). Fifteen taxidermy samples, representing the indigenous 'Old Goat' populations of the islands, were sequenced alongside five modern Irish dairy goats and four feral samples from endangered populations in western Ireland. Phylogenetic and network analyses reveal distinct groupings dominated by historical British and Irish samples within European mitochondrial variation, demonstrating a degree of maternal genetic structure between the goats of insular and continental Europe. Several of the the Irish modern feral samples also fall within these historical clades, suggesting continuity between these dwindling populations and the ancestral 'Old Goats' of Ireland and Britain.

Background

The domestic goat (*Capra hircus*) is an agricultural species of global importance, prized in regions of marginal pasture, where goats provide a reliable source of meat and dairy products [1,2]. Unfortunately, high extinction rates among wild ibex and feral goat populations are beginning to deplete the bank of genetic variation available to breeders. This reduction in genetic diversity is further confounded by modern intensive selection schemes which seek to create improved domestic breeds [3,4]. Suurveys of the genetic variation in surviving feral goat populations may help to identify diverse and irreplaceable groups that are in need of conservation efforts. The indigenous goat populations of Britain and Ireland, introduced to the islands during the Neolithic [5], are hypothesised to be one such candidate. These landrace animals, collectively described as 'Old Goats', were predominantly used in agriculture up until the late 19th century when they began to be replaced by continental breeds [6]. Today they survive only in small feral herds, whose existence, in recent decades, has been threatened by habitat loss, culling and the ongoing impact of continental introgression.

Museums and private natural history collections house an immense bank of taxidermy specimens that represent an extensive, although relatively untapped, source of genetic information [7–9]. In recent years however, these collections have been exploited in a number of landmark studies, which have documented recent temporal changes in diversity and population structure in a range of species [7,10–12], by harnessing the power of high throughput sequencing (HTS). These genetic resources are likely to be especially important in the case of domestic animals, which have seen dramatic drops in effective population sizes and extensive transportation over the past two centuries owing to the advent of industrialized breeding [3,13]. Indeed these collections may now be the only remaining record of diversity that has since been lost to history.

Here, we present a genetic survey of 15 'Old Goat' historical samples, derived from taxidermy specimens, spanning roughly 200 years in Britain and Ireland, as well as 9 modern samples from dairy and feral

populations in Ireland. These are analysed using whole mitochondrial genomes generated through RNA capture and HTS. Mitochondrial DNA (mtDNA) has been the main genetic marker utilized in previous surveys of genetic diversity in goats (14–16) and allows for the placement of individuals from this study into the wider phylogeographical context.

Methods

DNA extraction was performed using standard protocols. Illumina indexed sequencing libraries were prepared, pooled in equimolar amounts and subjected to target-enrichment using an Agilent custom-made mtDNA in solution (RNA) capture (Supplementary methods). Sequencing was performed on the Illumina MiSeq platform (50 bp single-end). NGS reads were aligned and filtered following standard pipelines (Supplementary methods). Fasta consensus sequences [17] were called using UGENE [18] and deposited in GenBank (KY564246-KY564269) and Dryad (doi:10.5061/dryad.g9v24).

Full mitochondrial genomes for 53 European and Turkish goats [16] were used alongside 23 whole mitochondrial sequences of sufficient coverage from this study (>85% of sequence covered at ≥ 3X) to construct a maximum-likelihood phylogeny. The HKY85 [19] substitution model was selected for tree-building using jmodeltest2 [20,21], and performed with 1000 bootstrap repeats.

HV1 control regions (482bp) of 132 Western and Central European domestic goats [14] were retrieved from GenBank (Extended Data Table 2). These samples were collected from indigenous herds belonging to remote villages and thus make an ideal comparative dataset. Notably this included HV1 sequences from 4 additional feral Irish goats, 2 English Kielder goats and 4 Welsh Bagot goats. These data, alongside 22 HV1 control regions from the current study, were used in the creation of a maximum-likelihood phylogeny, constructed using the Kimura 2-parameter model of substitution (1000 bootstrap repeats), as has been standard in previous studies of goat D-loop variation [14,15]. Additionally,

a median-joining [22] haplotype network of the same dataset was created using PopART (http://popart.otago.ac.nz).

Multiple sequence alignments were performed with Muscle [23]. Alignments and maximum-likelihood phylogeny constructions were implemented in seaview v 4.5.4 [24].

Results

In total 24 whole mtDNA genomes were generated at an average coverage of 40X (Table 1, Supplementary Fig 1). As expected, modern samples yielded higher numbers of total sequenced reads compared to the majority of historical samples. However, historical samples in general were seen to posses a higher capture specificity (Supplementary Fig 2). This resulted in a higher average coverage recovered for historical samples (44X) rather than modern samples (33X), albeit with a greater variance (4X-97X) than moderns (12X-51X). These data reiterates recent findings suggesting taxidermy samples are an excellent choice for study in population genetic experiments [10], and highlights the effectiveness of the mitochondrial RNA capture method for these types of specimen.

The whole mitochondrion phylogeny (Fig 1) showed the vast majority of historical samples from this study to fall into two groupings at high bootstrap probability (81 and 87 respectively), both of which are composed solely of samples from Britain and Ireland. Notably, all modern Irish dairy goats are seen to place outside of these groupings. Furthermore, two modern feral goats from the Burren, Co. Clare (M-Irish3 and M-Irish4), both of whom showed morphological signs of Swiss introgression (e.g chin tassels, large ears, conformation) [25], also fall away from the historical clades, highlighting the impact mass importation of continental breeds [6] has had on genetic variation in both commercial and feral Irish goat herds. Importantly from a conservation perspective, two modern goats (M-Irish1 and M-Irish2), belonging to a feral herd in Mulranny, Co. Mayo, are seen clustering with high bootstrap support

alongside a clade of three historical 'Old Scottish' goats (H-Scot1, H-Scot2 and H-Scot8), two of which originate from an extinct population in the Isle of Skye that existed before 1900.

Due to the limited sample size of the whole mitochondrion tree, a HV1 control region phylogeny was also constructed to investigate whether the above clades remained consistent within a larger western and central European dataset (Fig 2.A). Bootstrap values (not shown) were uniformly low across the haplogroup A branches of the phylogeny, reflecting a lack of diversity between sequences. This is illustrated by the median-joining network in Fig 2.B, where many haplotypic groupings are defined by only one or two mutations, emphasizing the need for whole mitochondrion sequences in studies of goat maternal diversity. Despite this lack of bootstrap support, the HV1 control region phylogeny, taken together with the network, support the conclusions drawn from the whole mitochondrion tree. Two peripheral groupings, dominated by historical British and Irish samples are seen apart from the remainder of European Haplogroup A variation, with branching patterns similar to those seen in the whole mitochondrion phylogeny. This includes the placement of the Mulranny goats with those from the Isle of Skye, alongside a third previously published Irish feral sample [14]. Interestingly, this network branch is also seen to contain a Norwegian Landrace and native Polish goat.

Overall, these results suggest some degree of mitochondrial geographic structure between the 'Old Goats' of insular Europe and their continental counterparts, an expected characteristic of island populations. Asides from the branchings described above and a few other possible exceptions, most notably a cluster containing only Welsh, English and Icelandic goats alongside M-Improv1, the HV1 tree and network both show decidedly low levels of geographic structure. This has been previously observed among goat populations, postulated to be the outcome of extensive movement and trade of goats in their history [15].

Discussion

This to our knowledge is the first study in livestock genetics to apply HTS to taxidermy specimens and it further demonstrates the high rates of efficiency of in-solution hybrid capture as a target enrichment method for skin samples [26]. This cost efficient approach can give access to the wealth of genetic information housed in taxidermic collections, an invaluable resource when studying the population history of domesticates. Extensive transportation and intensive selection pressure over the past two centuries has likely obscured more ancient patterns of genetic variation within these species and dramatically reduced their levels of genetic diversity. The study of museum specimens also allows for the identification of present day, unique populations who have managed to retain some of this past diversity, many of which are being pushed to the brink of extinction today. We demonstrate this potential here, by drawing a link between extinct 'Old Goat' populations living on the Isle of Skye over 100 years ago to endangered herds living in Mulranny, Co. Mayo today. Our results also support documented morphological observations suggesting that, in comparison to the Mulranny population, feral herds in the Burren have been subject to a larger degree of introgression. Overall, these findings identify Mulranny goats as some of the last modern representatives of the 'Old Irish' type of goat, once ubiquitous throughout the island, and highlight them for much needed conservation efforts.

Acknowledgments

We thank The Manx Museum & National Trust, Douglas; The Natural History Museum, London; The National Museum of Scotland, Edinburgh; The National Museum Wales, Cardiff; The Moran family and Westport house, Co. Mayo for providing the taxidermy samples. We are grateful to Nigel Monaghan, National Museum of Ireland - Natural History, Dublin, for sampling assistance. We also thank Claudia Marl and Christian Phol, Barba Goat Farm for providing dairy goat samples and The Old Irish Goat

society, Mulranny, for the modern feral samples. This research was funded by ERC Investigator Grant 295729-CodeX to D.G.B. LMC is funded by the Irish Research Council (GOIPG/2013/1219).

References

1. Nicoloso, L. et al. 2015 Genetic diversity of Italian goat breeds assessed with a medium-density SNP chip. Genet. Sel. Evol. 47, 62.
2. Dong, Y. et al. 2013 Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (Capra hircus). Nat. Biotechnol. 31, 135–141.
3. Taberlet, P., Valentini, A., Rezaei, H. R., Naderi, S., Pompanon, F., Negrini, R. & Ajmone-Marsan, P. 2008 Are cattle, sheep, and goats endangered species? Mol. Ecol. 17, 275–284.
4. Rosa García, R., Celaya, R., García, U. & Osoro, K. 2012 Goat grazing, its interactions with other herbivores and biodiversity conservation issues. Small Rumin. Res. 107, 49–64.
5. Yalden, D. 2010 The History of British Mammals. A&C Black.
6. Porter, V. & Tebbit, J. 1996 Goats of the world. Farming Press.
7. Wandeler, P., Hoeck, P. E. A. & Keller, L. F. 2007 Back to the future: museum specimens in population genetics. Trends Ecol. Evol. 22, 634–642.
8. Kemp, C. 2015 Museums: The endangered dead. Nature 518, 292–294.
9. Martínková, N. & Searle, J. B. 2006 Amplification success rate of DNA from museum skin collections: a case study of stoats from 18 museums. Mol. Ecol. Notes 6, 1014–1017.
10. Ávila-Arcos, M. C. et al. 2013 One hundred twenty years of koala retrovirus evolution determined from museum skins. Mol. Biol. Evol. 30, 299–304.
11. Miller, W. et al. 2009 The mitochondrial genome sequence of the Tasmanian tiger (Thylacinus cynocephalus). Genome Res. 19, 213–220.

12. Bi, K., Linderoth, T., Vanderpool, D., Good, J. M., Nielsen, R. & Moritz, C. 2013 Unlocking the vault: next-generation museum population genomics. Mol. Ecol. (doi:10.1111/mec.12516)
13. Goddard, M. E. & Hayes, B. J. 2009 Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nat. Rev. Genet. 10, 381–391.
14. Luikart, G., Gielly, L., Excoffier, L., Vigne, J. D., Bouvet, J. & Taberlet, P. 2001 Multiple maternal origins and weak phylogeographic structure in domestic goats. Proc. Natl. Acad. Sci. U. S. A. 98, 5927–5932.
15. Naderi, S. et al. 2007 Large-scale mitochondrial DNA analysis of the domestic goat reveals six haplogroups with high diversity. PLoS One 2, e1012.
16. Colli, L. et al. 2015 Whole mitochondrial genomes unveil the impact of domestication on goat matrilineal variability. BMC Genomics 16, 1115.
17. Cassidy, L. M., Teasdale, M. D., Carolan, S., Enright, R., Werner, R., Bradley, D. G., Finlay, E. K., Mattiangeli, V. 2017 Capturing goats: Documenting two hundred years of mitochondrial DNA diversity among goat populations from Britain and Ireland. Dryad Digital Repository. (http://dx.doi.org/10.5061/dryad.g9v24)
18. Okonechnikov, K., Golosova, O., Fursov, M. & UGENE team 2012 Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics 28, 1166–1167.
19. Hasegawa, M., Kishino, H. & Yano, T. 1985 Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22, 160–174.
20. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. 2012 jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772.
21. Guindon, S. & Gascuel, O. 2003 A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.
22. Bandelt, H. J., Forster, P. & Röhl, A. 1999 Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48.
23. Edgar, R. C. 2004 MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.

24. 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.
25. R. Werner, "A management strategy for maintaining a feral goat population on the burren" (www.agrobiodiversity.net, 2010), (available at http://www.agrobiodiversity.net/topic_network/feral/Best_Practise/management-plan-for-the-burren-feral-goat.1.pdf).
26. O'Sullivan, N. J., Teasdale, M. D., Mattiangeli, V., Maixner, F., Pinhasi, R., Bradley, D. G. & 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.

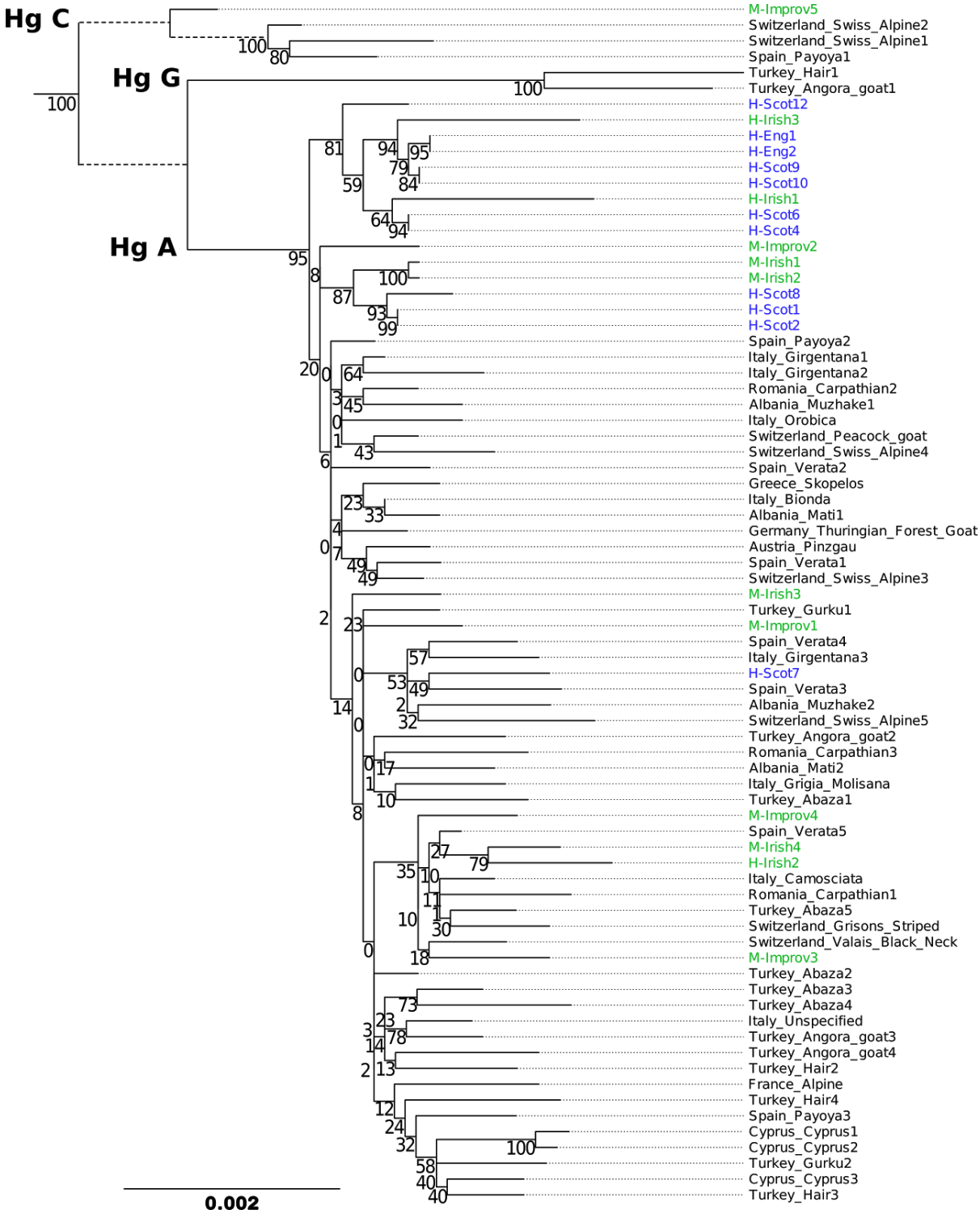


Figure 1. Whole mitochondrion ML phylogeny of 76 domestic goats from Europe and Turkey. Goats from the current study are coloured in green or blue depending on their geographical origin - Ireland and Britain respectively. Bootstrap values and haplogroup (Hg) divisions are displayed along branches. Dashed lines represent branches scaled to a fifth of their true length.

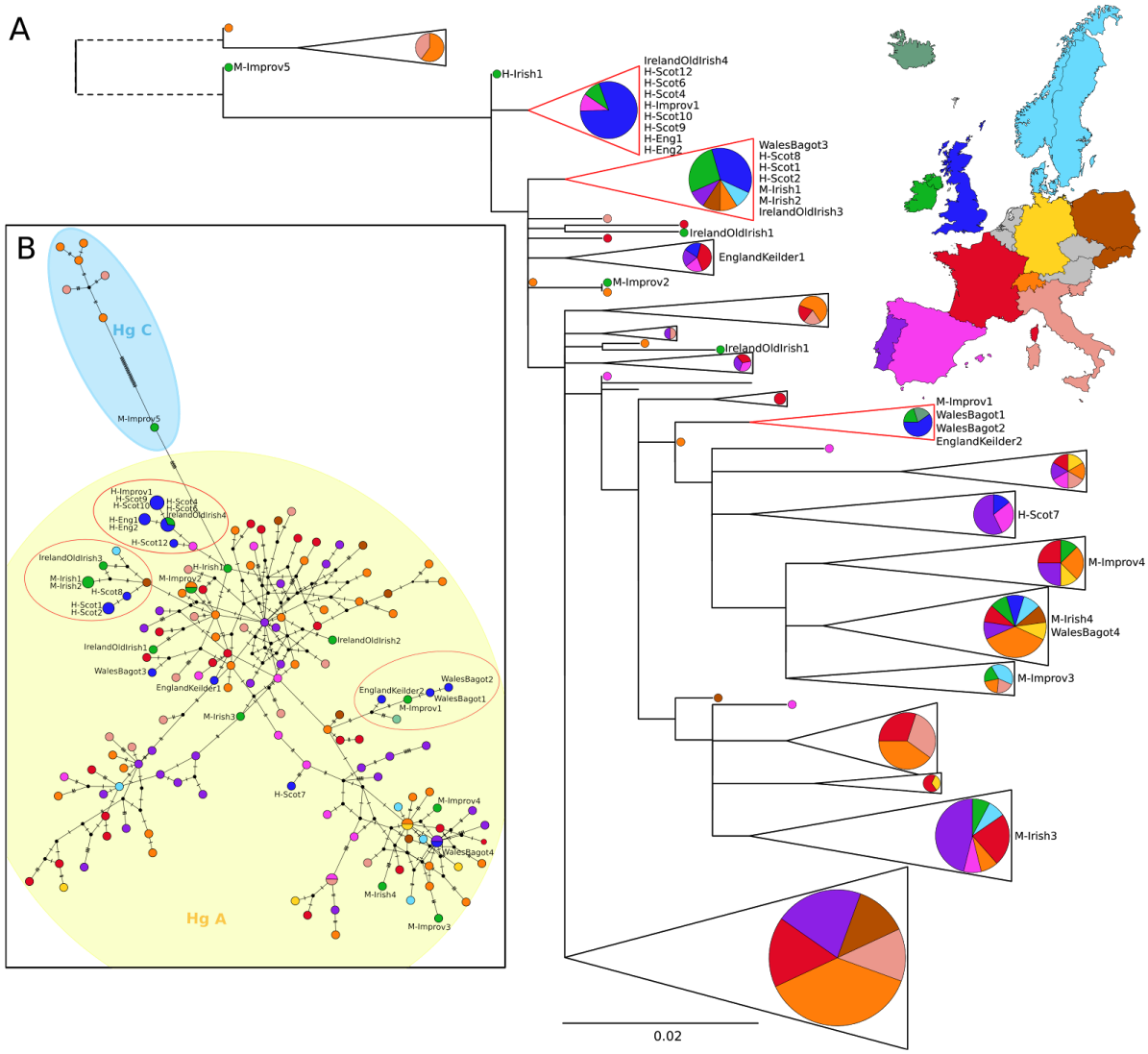


Figure 2. HV1 control region diversity in 154 goats from indigenous populations of Western and Central Europe. Individuals from Ireland and Britain are labelled. Groupings containing a majority of insular samples are highlighted in red. A) shows a HV1 control region maximum-likelihood phylogeny. Dashed lines represent branches scaled to a fifth of their true length. The composition of each collapsed subclade is represented with a pie chart, sized according to number of samples within the clade, and coloured according to the geographical origins of these samples. B) shows a median-joining network following the same colour key. Branch lengths are not proportional to genetic distance. Mutations between nodes are represented through hatch marks.

Map legend	Sample ID	Breed	Type	Origin	Date	Total Reads Sequenced	Aligned mtDNA reads (%)	Mitoch-ondrial Coverag e (X)	Haplogroup
1	M-Irish1	Old Irish	Feral	Mulranny, Ireland	Modern	950744	0.41	11.51	A2c1
2	M-Irish2	Old Irish	Feral	Mulranny, Ireland	Modern	2350466	0.59	40.97	A2c1
3	M-Irish3	Old Irish (w/ Foreign Introgression)	Feral	The Burren, Ireland	Modern	1162452	1.34	46.05	A2c
4	M-Irish4	Old Irish (w/ Foreign Introgression)	Feral	The Burren, Ireland	Modern	1178058	0.52	18.22	A2c1
5	M-Improv1	British Saanen-Toggenburg Dairy	Domestic	Westmeath, Ireland	Modern	508674	2.40	36.28	A2c
6	M-Improv2	British Saanen-Toggenburg Dairy	Domestic	Westmeath, Ireland	Modern	725279	2.30	49.55	A2c1
7	M-Improv3	British Saanen Dairy	Domestic	Westmeath, Ireland	Modern	864103	1.99	51.09	A2c1
8	M-Improv4	British Saanen Dairy	Domestic	Westmeath, Ireland	Modern	730645	1.37	29.80	A2c1
9	M-Improv5	British Alpine Dairy	Domestic	Westmeath, Ireland	Modern	1872889	0.32	17.47	C
-	H-Irish1	Old Irish	Feral	Ireland	Unknown	85113	4.19	8.84	A2c
-	H-Irish2	Old Irish	Feral	Ireland	Unknown	162974	2.54	10.00	A2c1
10	H-Irish3	Old Irish	Feral	Achill Island, Ireland	Unknown	133854	2.32	7.58	A2c1
11	H-Scot1	Old Scottish	Feral	Isle of Skye, Scotland	<1895	320262	8.99	84.35	A2c1
12	H-Scot2	Old Scottish	Feral	Isle of Skye, Scotland	<1898	177917	9.56	49.32	A2c1

Table 1. Sequencing results for 24 modern and historical goats from Britain and Ireland. Aligned mtDNA reads denotes percentage of total reads aligning to the goat mitochondrial genome after filtering for mapping quality and removal of PCR duplicates. Sample IDs are prefixed with an M or H, referring to whether the specimen is modern or historical. Map legend refers to Supplementary Figure 1.

Supplementary Methods

Capturing goats: Documenting two hundred years of mitochondrial DNA diversity among goat populations from Britain and Ireland

Lara M. Cassidy[1], Matthew D. Teasdale[1], Sean Carolan[2], Ruth Enright[3], Raymond Werner[2], Daniel G. Bradley[1], Emma K. Finlay[4], Valeria Mattiangeli[1].

1 Smurfit Institute of Genetics, Trinity College Dublin, Ireland.

2 Old Irish Goat Society, Mulranny, Ireland.

3 School of Biological, Earth and Environmental Sciences, University College Cork, Ireland.

4 School of Biotechnology, Dublin City University, Dublin, Ireland

DNA Extraction

A total of 24 samples were used in this study; 9 modern and 15 historical samples, representing both 'Old Goat' populations and improved dairy breeds (figure S1). Prior to extraction, 1mm3 of hair and dry skin from each historical sample, typically taken from the neck of a mounted head, was washed in 500µl of ddH2O, 300µl of 99% ethanol and then again in 500µl of ddH2O. For each wash samples were centrifuged at 13,500rpm for one minute, discarding the supernatant every time.

Modern specimens consisted of hair and follicles for all samples but M-Irish1 and M-Irish2, for which fresh skin was obtained. For extraction, approximately 0.03g of each modern sample was digested in 450µl TNE buffer (0.5M Tris–HCl, pH 7.4, 0.1M-NaCl, 0.001M-EDTA, 0.5% SDS) and proteinase-K solution (1 mg per sample) at 56°C for 16h. For the washed historical samples, a modified TNE buffer was used, adding 0.003M CaCl2 and 0.05M DTT to the original TNE buffer. Following the 16h incubation, an extra 0.1mg of proteinase-K solution was added to all samples, which were then further incubated for two hours.

A standard phenol-chloroform extraction was then performed (1). The DNA pellet in each tube was re-suspended in 20µl of EBT buffer (10 mM Tris-Cl, pH 8.0 and 0.05% Tween 20). Historical samples in which no DNA was detected (quantified by Qubit®) after the extraction, were excluded from the study at this point and are not presented in Table 1. The modern samples were sonicated (by Bioruptor® for 30 cycles; 30 sec ON/30 sec OFF) to obtain DNA fragments of suitable length for Illumina sequencing (<600bp) and a size selection was carried out using gel electrophoresis. Fragments of approximately 100bp were selected and purified using Qiagen gel purification protocol (QIAquick Gel extraction kit).

Negative controls were carried through the procedure for both modern and historical samples.

Library preparation and mtDNA capture

Indexed single-end sequencing libraries were prepared following the Meyer and Kircher (2010) Illumina sequencing library preparation protocol (2) with modifications as reported in (3). All DNA purification

steps were performed using the MinElute PCR Purification kit protocol. A negative control of 20µl ddH2O was carried through the entire procedure. Indexing and library adapter primers were obtained from Metabion. The library amplification reaction was performed for all samples and controls using Accuprime Pfx Supermix (Life Technology), under the following thermal cycling conditions: 5 min at 95°C; 10-15 cycles of 15 sec at 95°C, 30 sec at 60°C and 30 sec at 68°C; and a final extension step of 5 min at 68°C.

All indexed libraries were then pooled in equimolar amounts (determined by Bioanalyzer). A custom mtDNA capture system for livestock mtDNAs (SureSelect Target Enrichment System, Agilent) was used for target-enrichment of the pooled libraries, following the recommended protocol. The capture design included the complete mtDNA genome of goat (NC_005044), along with 10 d-loop sequences from the major goat haplogroups as well as a large number of diverse domesticate mtDNAs. A control containing 4µl of ddH2O was carried through the in-solution mtDNA capture procedure.

Sequencing and Data Processing

Given the degraded nature and short sequence lengths of historical DNA, a single end approach was adopted for sequencing. The multiplex mtDNA captured from the pooled libraries was sequenced using an Illumina MiSeq platform (50bp single-end). A 1% PhiX positive control (a viral genome with a balanced nucleotide representation) was added prior to the sequencing.

Raw reads from each sample were demultiplexed according to the index and each adapter sequence was trimmed using cutadapt v1.3 (4). Sequence reads were then aligned to the domestic goat (Capra hircus aegagrus) mitochondrial genome (16,443bp) retrieved from NCBI GenBank (NC_005044) using BWA (5) with the parameters -n 0.01, -o 2 and -l 1000. Relatively low coverage was observed over the D-loop region in all samples aligned. High sequence diversity in this region can affect read mapping. Therefore, we replaced the D-Loop in the reference sequence, belonging to haplogroup B (6), with the D-loop sequence of an Irish goat belonging to haplogroup A (EF618086) (7) and repeated the alignment. This modification resulted in substantially increased sample coverage in this region, suggesting that any limiting

step in terms of recovery of diverse sequences appears to be bioinformatic, rather than capture related, again highlighting the efficiency of the in-solution mtDNA capture approach.

Read groups were added using Picard Tools v1.101 (http://broadinstitute.github.io/picard/) and reads were sorted using SAMtools (8). Reads with a mapping quality below 30 were discarded and duplicate reads (clonal PCR amplified reads) were removed, again through the use of SAMtools.

Fasta consensus sequences were called for all individuals using the Unipro UGENE genome analysis suite (9). Haplogroup membership was determined with MitoToolPy (10). The total depth of sequencing coverage for each mitochondrial genome was assessed using GATK (11).

The previously generated trimmed reads for each sample and reference modern goat samples (ERR219543, ERR470101) were aligned using BWA to the complete goat genome CHIR_1.0 with the mtDNA genome replaced by the spliced genome from this study. Duplicate reads were removed. The percentage of mtDNA to autosomal aligned reads generated by this analysis was calculated, with results displayed in figure S2. The two reference samples used are part of data released, prior to publication, from the NextGen Consortium. The use of this data in the current study is in keeping with the Fort Lauderdale principles and these samples were not included in any phylogenetic analysis.

Patterns of postmortem DNA damage were examined for this study's samples using mapDamage 2.0 (12). An ancient Irish bone sample, Rathlin2 (2024-1741 cal BC) (13), was also included in the analysis for comparative purposes. Results are shown in figure S3. Taxidermic samples showed inflated levels of C to T and G to A misincorporations at the 5' and 3' of reads, relative to the modern specimens. However, as expected, these frequencies were dwarfed in comparison to those observed for the ancient bone sample.

References

1. Sambrook J, Fritsch EF, Maniatis T. 1989 *Molecular cloning.* New York: Greene Publishing Associates and John Wiley & Sons.
2. Meyer M, Kircher M. 2010 Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing. *Cold Spring Harb Protoc.* 2010(6):db.prot5448

3. Gamba C, Jones ER, Teasdale MD, McLaughlin RL, Gonzalez-Fortes G, Mattiangeli V, et al. 2014 Genome flux and stasis in a five millennium transect of European prehistory. *Nat. Commun.* 5:5257.
4. Martin M. 2011 Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal.* 17(1):10–2.
5. Li H, Durbin R. 2009 Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics.* 25(14):1754–60.
6. Hassanin A, Bonillo C, Nguyen BX, Cruaud C. 2010 Comparisons between mitochondrial genomes of domestic goat (Capra hircus) reveal the presence of numts and multiple sequencing errors. *Mitochondrial DNA.* 21(3-4):68–76.
7. Naderi S, Rezaei H-R, Taberlet P, Zundel S, Rafat S-A, Naghash H-R, et al. 2007 Large-scale mitochondrial DNA analysis of the domestic goat reveals six haplogroups with high diversity. *PLoS One.* 2(10):e1012.
8. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, et al. 2009 The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics.* 25, 2078-9.
9. Okonechnikov K, Golosova O, Fursov M & UGENE team. 2012 Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics.* 28, 1166–1167.
10. Peng M-S, Fan L, Shi N-N, Ning T, Yao Y-G, Murphy RW, et al. 2015 DomeTree: a canonical toolkit for mitochondrial DNA analyses in domesticated animals. *Mol Ecol Resour.* 15(5):1238–42.
11. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. 2010 The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20(9):1297–303.
12. Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L. 2013 mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics.* 9(13):1682–1684
13. Cassidy LM, Martiniano R, Murphy EM, Teasdale MD, Mallory J, et al. 2016 Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. *Proc Natl Acad Sci USA* 113(2):368–373.

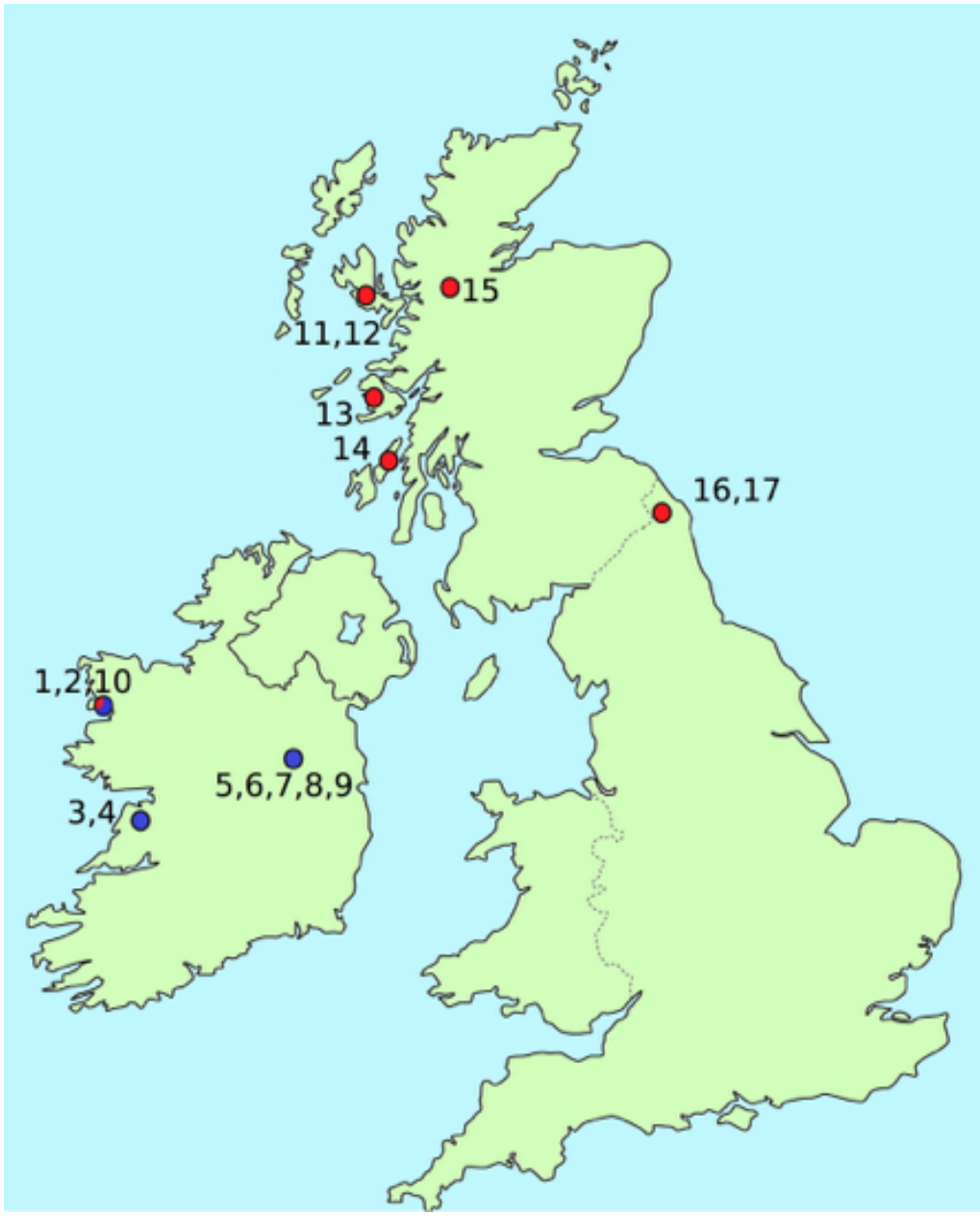


Figure S1. Geographic origins for 17 samples out of the total 24 sequenced. Red denotes historical, while blue represents modern samples. Map legend shown in Table 1.

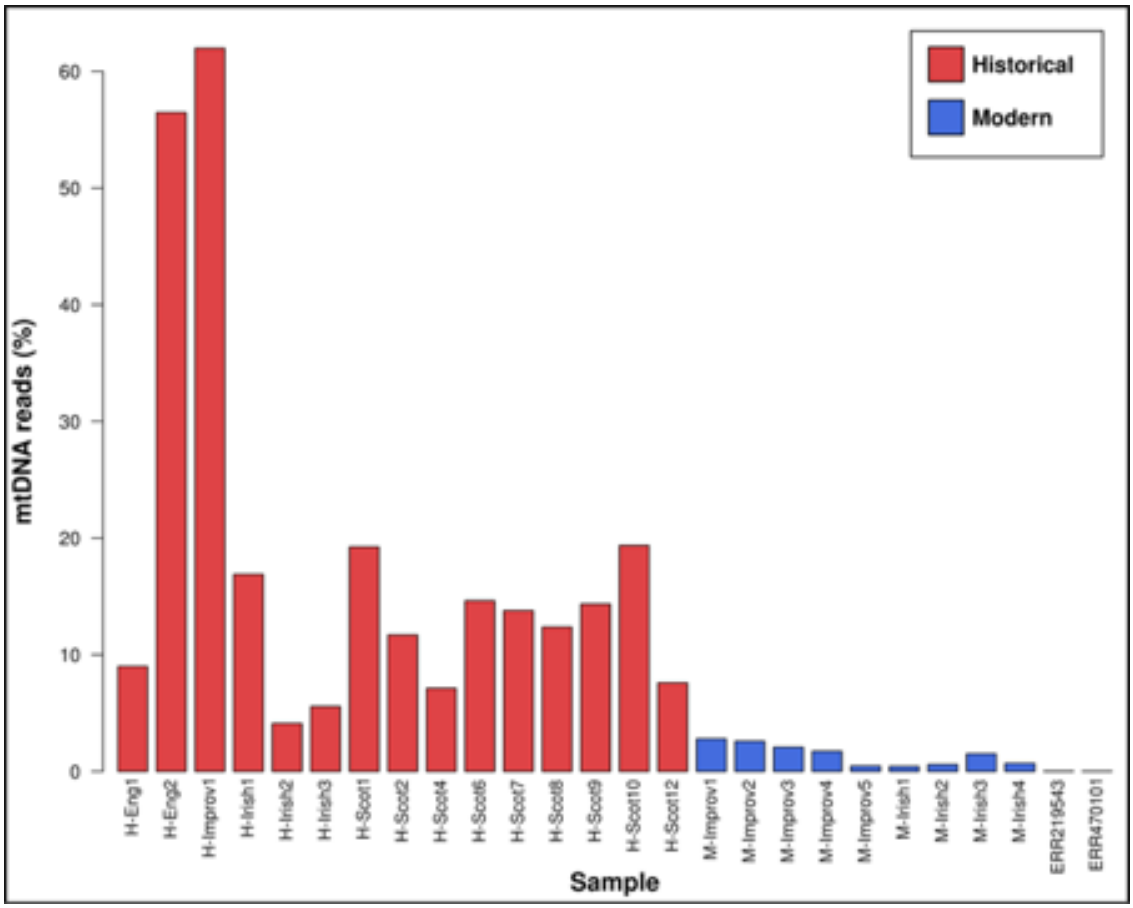


Figure S2: This chart highlights the percentage of mtDNA reads retrieved out of the total aligned reads from 26 goat samples when aligned to the complete goat reference genome (CHIR_1.0 with the mtDNA genome replaced by the spliced genome from this study). Samples H-Eng1 to M-Irish4 represent the samples selected for mtDNA capture in this study. ERR219543 (Capra hircus from Morocco) and ERR470101 (Capra hircus from France) are whole genome sequenced modern goat samples retrieved from the NCBI Sequence Read Archive and as such provide a base level of mtDNA retrieval from genomic samples. An increase in mtDNA alignment relative to the whole genome sequencing is seen in all captured samples.

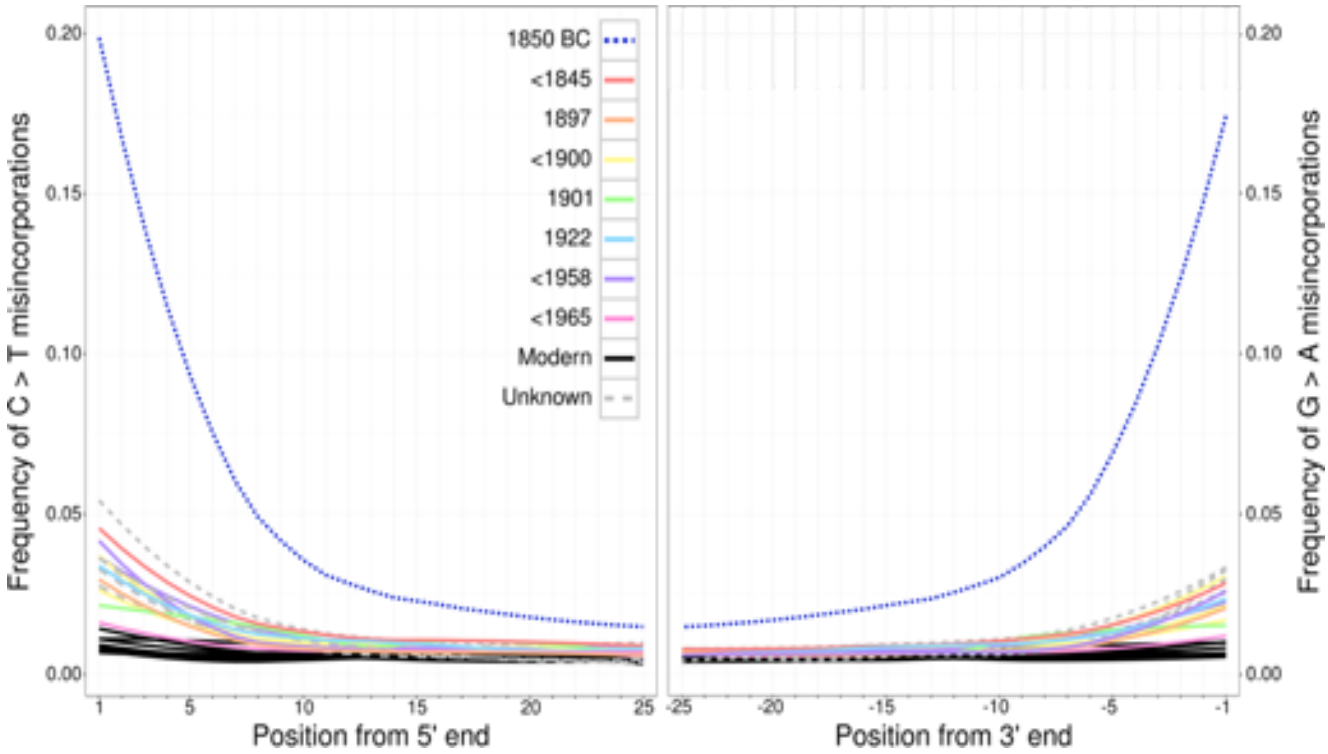


Figure S3: **Misincorporation patterns for modern and taxidermic samples.** These display the mismatch frequency relative to the reference genome as a function of position in the sequence r ead. Samples are coloured according to date. Samples of unknown age are represented by a grey dashed line. An ancient Irish bone sample (blue dotted line) is also displayed for comparative purposes.