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	2026-11-05 14:53:01
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], Seán 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.

Supplementary Methods

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.

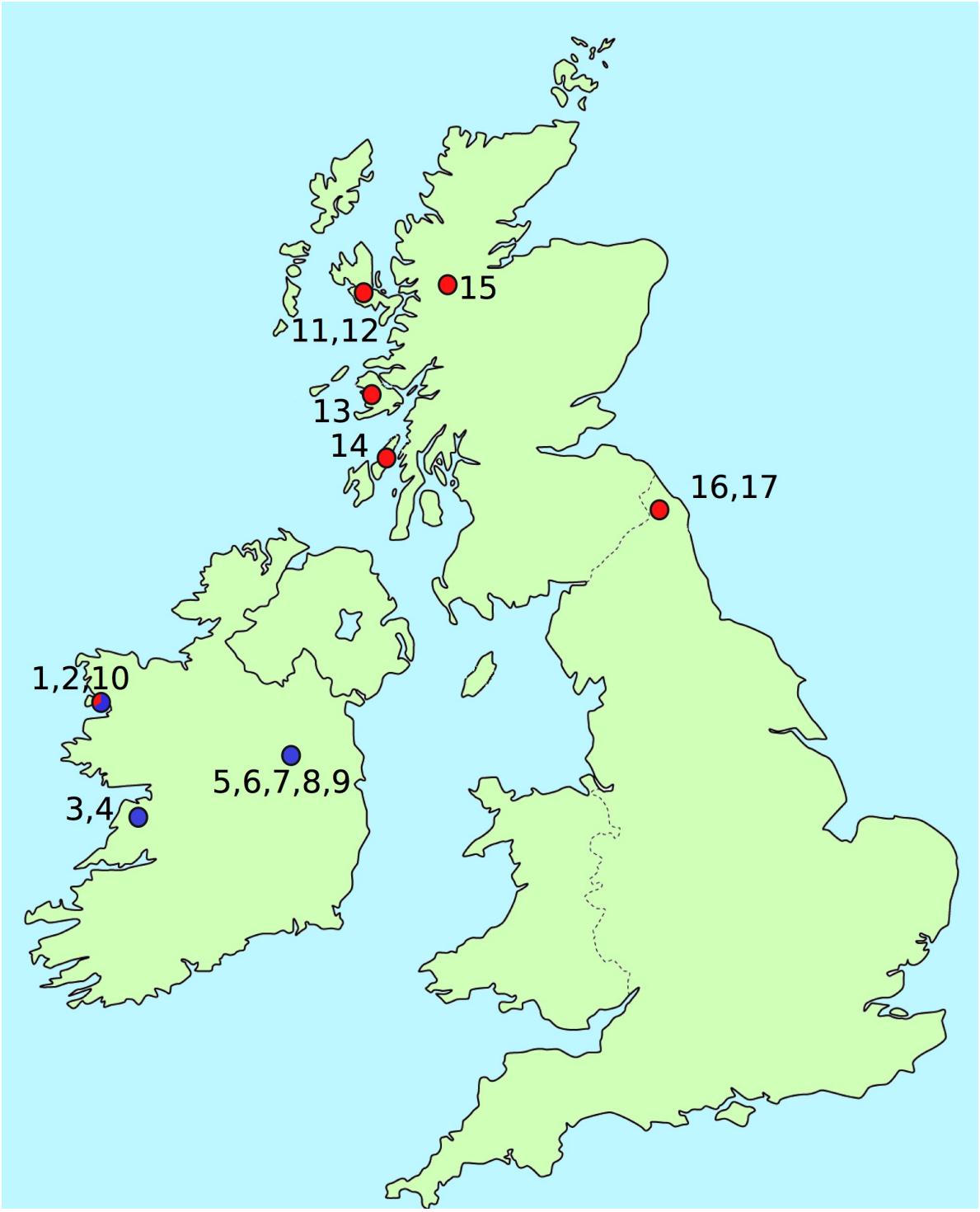
10. Peng M-S, Fan L, Shi N-N, Ning T, Yao Y-G, Murphy RW, et al. 2015 DomeTree: a canonical toolkit for mitochondrial DNA analyses in domesticated animals. *Mol Ecol Resour.* 15(5):1238–42.

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.



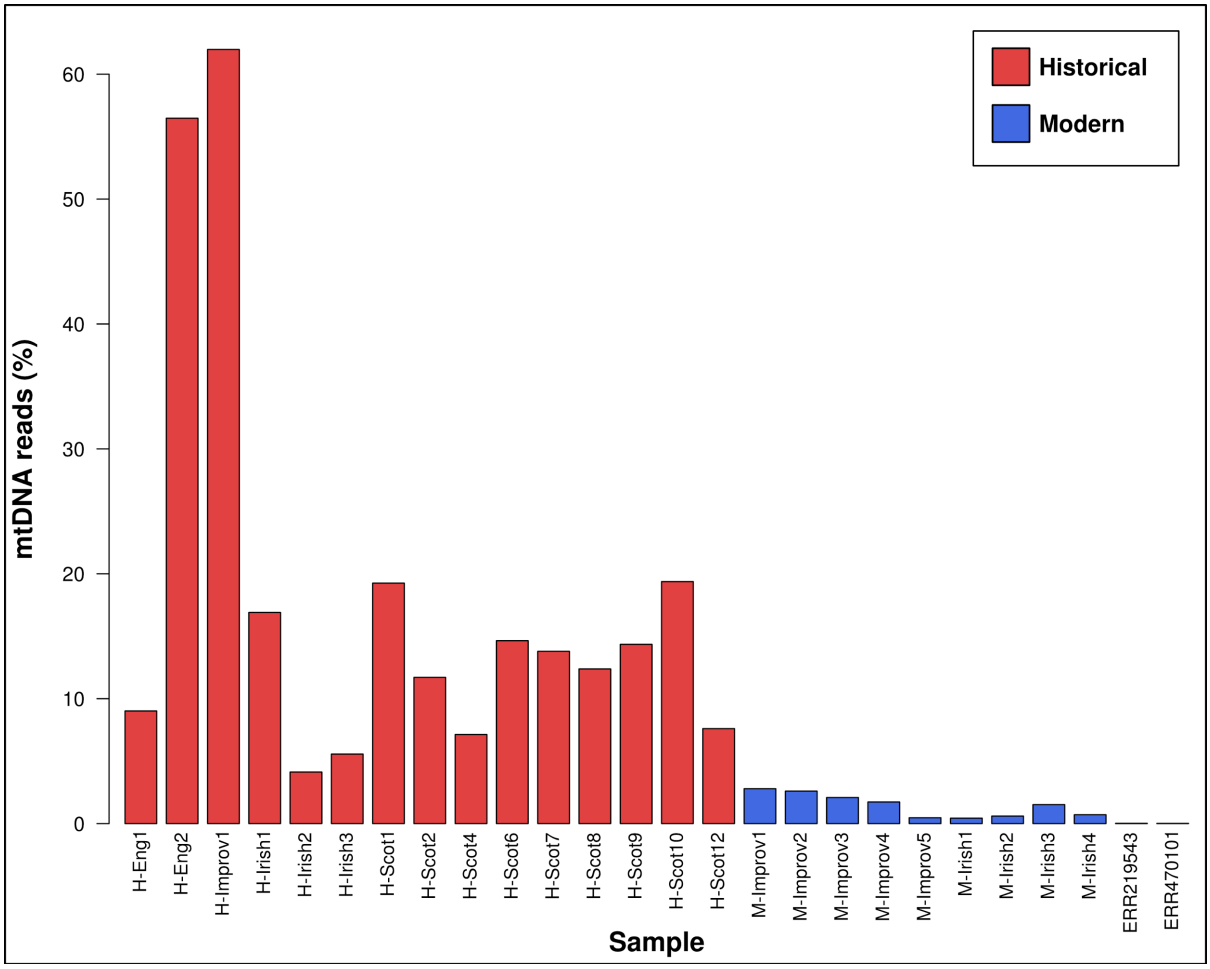


Figure S2: This chart highlights the percentage of mtDNA reads retrieved out of the total aligned reads from 26 goat samples when aligned to the complete goat reference genome (CHIR_1.0 with the mtDNA genome replaced by the spliced genome from this study). Samples H-Eng1 to M-Irish4 represent the samples 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.

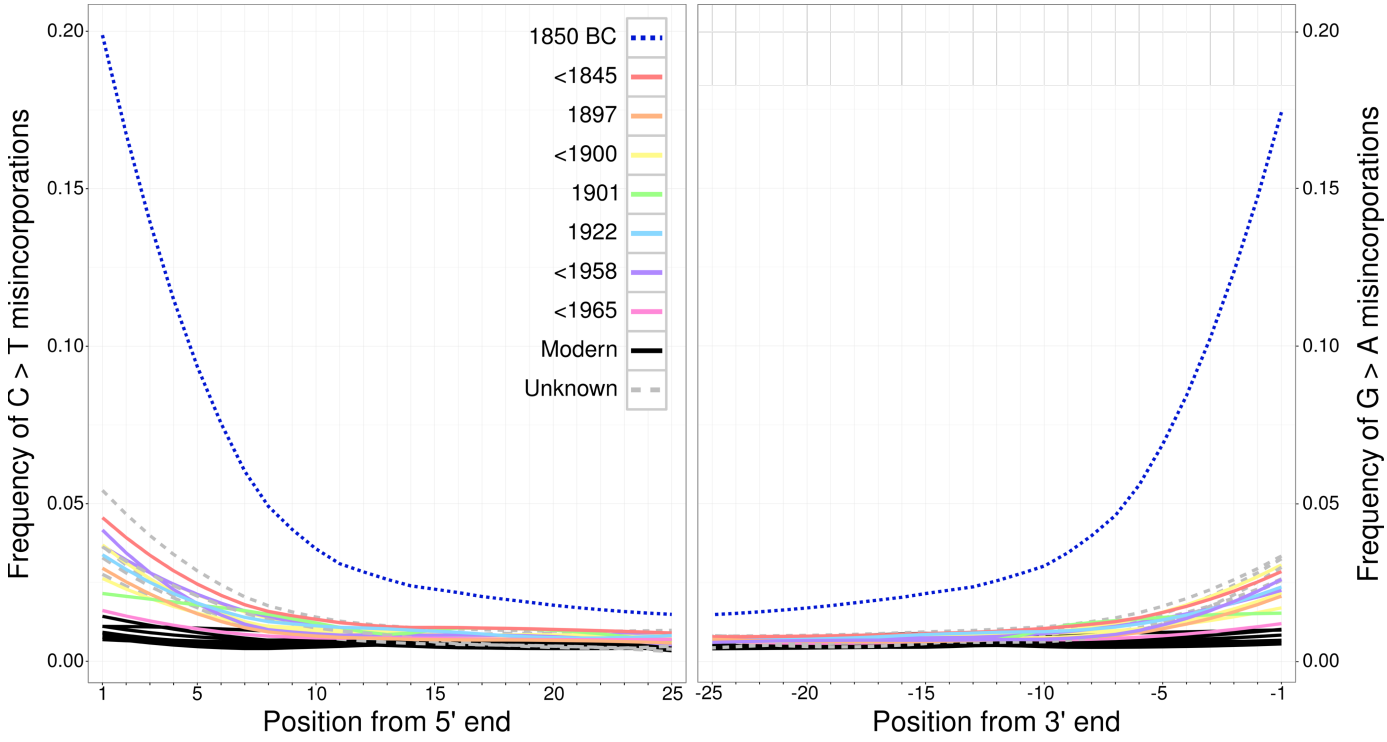


Figure S3: **Misincorporation patterns for modern and taxidermic samples.** These display the mismatch frequency relative to the reference genome as a function of position in the sequence read. Samples are coloured according to date. Samples of unknown age are represented by a grey dashed line. An ancient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