

Mitochondrial DNA sequence diversity in extant Irish horse populations and in ancient horses

A. M. McGahern*, C. J. Edwards[†], M. A. Bower[‡], A. Heffernan^{§,¶}, S. D. E. Park*, P. O. Brophy*, D. G. Bradley[†], D. E. MacHugh* and E. W. Hill*

*Animal Genomics Laboratory, School of Agriculture, Food Science and Veterinary Medicine, College of Life Sciences, University College Dublin, Belfield, Dublin 4, Ireland. [†]Smurfit Institute of Genetics, Trinity College, Dublin 2, Ireland. [‡]McDonald Institute for Archaeological Research, Downing Street, Cambridge, CB2 3ER, UK. [§]Weatherbys Ireland Laboratory, c/o The Irish Equine Centre, Johnstown, Naas, Co. Kildare, Ireland. [¶]School of Biology and Environmental Science, College of Life Sciences, University College Dublin, Belfield, Dublin 4, Ireland

Summary

Equine mitochondrial DNA sequence variation was investigated in three indigenous Irish horse populations (Irish Draught Horse, Kerry Bog Pony and Connemara Pony) and, for context, in 69 other horse populations. There was no evidence of Irish Draught Horse or Connemara Pony sequence clustering, although the majority of Irish Draught Horse sequences (47%) were assigned to haplogroup D. Conversely, 31% of the Kerry Bog Pony sequences were assigned to the rare haplogroup E. In addition to the extant population analyses, ancient DNA sequences were generated from three out of four Irish archaeological specimens, all of which were assigned to haplogroup A.

Keywords ancient DNA, biomolecular archaeology, endangered breed, Irish horse, mitochondrial DNA.

Ireland is home to three native horse populations: the Irish Draught Horse (ID), the Kerry Bog Pony (KB) and the Connemara Pony (CON). Although pedigrees in these populations have been recorded since the turn of the last century, little is known about their genetic make-up or the genetic contributions of their founder populations. These indigenous Irish horses have played an integral role in native Irish culture, economics and social history. In the past, ID horses were used for agricultural purposes, hunting, carriage transport, war and barge towing. In recent times, the ID has found a niche as a competition animal (Storey 2005); however, its presence is being progressively diluted, and the selection of sire lines has been seriously limited. Numbers of ID are now so greatly reduced in Ireland, with approximately 900 mares and 94 approved stallions (Slavin 2005), that the breed is considered endangered (FAO 2000). KB ponies were used mainly to carry bog peat for domestic fuel, but the modernization of farming and transport has dramatically reduced the need for this horsepower. The KB pony was officially recognized as a breed in 2006

but is almost extinct; in 1994 there were only 20 KB ponies remaining in Ireland. Volume 1 of the Connemara Pony Breeders Society Stud Book was published in 1926 and contained details of nine stallions and 93 mares. The CON is renowned for its hardiness, vigour, stamina and gentle temperament and is a popular riding horse. This is reflected in the current population – the latest published volume of the CON Stud Book (Volume 21, 1999) documents breeding stallion numbers at 1043 and breeding mare numbers at 11,621.

Mitochondrial DNA (mtDNA) has been used extensively in the last three decades as a tool for inferring the evolutionary and demographic past of populations and closely related species. Previous molecular studies have used mtDNA nucleotide sequence for the assessment of horse diversity (Ishida *et al.* 1994; Kavar *et al.* 1999; Bowling *et al.* 2000; Vila *et al.* 2001; Hill *et al.* 2002; Mirol *et al.* 2002) and for conservation purposes (Oakenfull *et al.* 2000; Luis *et al.* 2002). In this study we examined the mtDNA diversity of the ID, KB and CON populations, and placed them within the context of 69 other horse populations (Table S1). Additionally, we used sequences from ancient Irish horses in order to make temporal inference about the present diversity found in Irish horse breeds.

Wild horses were present in Ireland around 28 000 years ago (Woodman *et al.* 1997), but the increase of ice cover at the last glacial maximum led to their extinction. When the glaciers retreated, they did not re-establish themselves on

Address for correspondence

E. W. Hill, Animal Genomics Laboratory, School of Agriculture, Food Science and Veterinary Medicine, College of Life Sciences, University College Dublin, Belfield, Dublin 4, Ireland.
E-mail: emmeline.hill@ucd.ie

Accepted for publication 25 June 2006

the island. The earliest evidence for the presence of domesticated horses in Ireland is from early Bronze Age archaeological contexts at Newgrange, Co. Meath, dating to about 4400 BP (years before present; F. McCormick, personal communication). In this paper, we present data from three ancient Irish horse samples, varying in age from 32 255 to 1244 calibrated BP, thus encompassing both wild and domestic Irish horses. These sequences provided an archaeogenetic context for the living populations included in the study.

Pedigree analysis of the ID population identified 414 founder females for a population of 878 males contained within the Irish Horse Register (2002). Fifty-nine male descendants were randomly selected from 120 of the founders that had at least three generations recorded in the pedigree. Thirty-three individuals from the KB pony population were selected on phenotypic traits. Because there is no pedigree history for the KB population, animals were DNA-typed using the StockMarks[®] for Horses Equine Genotyping System (Applied Biosystems, CA, USA) to determine relatedness. No parent-offspring pairs were found. To increase the sample size, six other KB ponies (with dams not among the original 33 animals) were included, giving a total sample size of 39. Twelve CON pony sequences, selected to represent CON founder female families, have been previously described (Hill *et al.* 2002) (AF481247–AF481258) and were included in this study to complete the set of extant native Irish horse populations.

Remains from four ancient Irish horses were obtained from sites in Edenvale, Co. Clare [metatarsus; NMING:F21099; 1595 ± 78 cal BP (OxA-3700)], Kesh, Co. Sligo [tooth; NMING:F21122; 1472 ± 56 cal BP (OxA-3707)], Lough Shad, Co. Roscommon [distal tibia; NMING:F20927; 578 ± 48 cal BP (OxA-3710)] and Shandon, Co. Waterford [scapula; NMING:F21158; 32,255 ± 712 cal BP (OxA-4244)] (Woodman *et al.* 1997). In addition, one piece of fragmentary bone from Carsington Pasture Cave, Derbyshire, England was added to the analysis. This sample was radiocarbon dated to 1244 ± 41 cal BP (uncalibrated radiocarbon age 1314 ± 31; KIA26353).

Total genomic DNA was isolated from 59 ID horses and 39 KB ponies, and mtDNA amplified as described in Appendix S1. DNA extractions from ancient material were performed at Trinity College Dublin from either bone or tooth samples. Three of the four ancient Irish samples generated reproducible endogenous DNA, whereas the sample from Lough Shad, Co. Roscommon, contained no amplifiable DNA. Samples were subsequently validated in an independent laboratory at the McDonald Institute for Archaeological Research, Cambridge. Amplicons were directly sequenced, and a representative sampling of PCR products were subsequently cloned (Appendix S1). A consensus sequence was constructed, which reached greater than 95% confidence levels at all bases and aligned perfectly to the direct sequence.

Sequences were deposited in GenBank (ID = DQ327891–DQ327949; KB = DQ327852–DQ327890; Ancient = DQ327848–DQ327851) and, for analyses, were truncated to 247 bp (nucleotide positions 15494–15740 of X79547; Xu & Arnason 1994). Twenty-eight distinct maternal lineages were identified in the 59 founder ID sequences, while 17 haplotypes were found in the 39 KB pony founder female lineages and 11 haplotypes in the 12 CON founders (Table 1).

The probability that two individuals in a population shared an identical haplotype (the probability of identity, PI) was calculated as the sum of the square of the frequency of each haplotype in that population. The estimated PI in KB was 0.0900, CON 0.097 and ID 0.058. Sequence sharing among the ID founders was threefold less than that observed in the thoroughbred (0.15; Hill *et al.* 2002) and is similar to estimates in the Arab horse (0.05; Bowling *et al.* 2000; McGahern *et al.* 2006). This suggests that ID mtDNA sequences were derived from diverse foundation stock, whereas KB and CON matrilineages were more limited.

Using the nomenclature from Jansen *et al.* (2002), 28 ID individuals grouped to haplogroup D (48%), with the remaining sequences in haplogroups A, C and B (32%, 12% and 8% respectively). The CON sequences were assigned to haplogroups D, A, B and C (42%, 25%, 17% and 17% respectively). These four haplogroups are the most common among European horse sequences. In contrast, the KB samples had an unusual sequence distribution, with the majority assigned to haplogroups E, C and A (31%, 26% and 23% respectively).

The ancient Irish horse remains were all assigned to haplogroup A, with a significant clustering evident when ancient sequences were compared with modern Irish ($P = 0.026$) and modern British ($P = 0.020$) populations (Table S1) using Fisher exact tests (Sokal & Rohlf 1995). Thus, the ancient sequences did not reflect the distribution of modern day Irish horse populations. By comparison a single ancient British sample was found in haplogroup D. The clustering of ancient Irish horse sequences within haplogroup A suggests that it may represent an ancient haplogroup. However, other studies that included data from ancient horses have found a more varied spread of samples across the seven haplogroups (A to G; Vila *et al.* 2001; Keyser-Tracqui *et al.* 2005). In addition, the position of these ancient Irish samples may be insecure because only three samples were assessed.

Unlike the situation for many other large domesticated animal species that demonstrate remarkably clear geographic patterning (MacHugh & Bradley 2001), caution must be used when inferring population structure from equine mtDNA data. Discrete horse populations share mtDNA sequence with other populations despite geographic distance. For example, in the ID the most common sequence haplogroup was also the most common globally, with 28 of 59 sequences (47%) represented in haplogroup D. Historical

Table 1 Mitochondrial control region sequence variation in Irish horses. A 247-bp segment of the control region from 59 ID, 39 KB, 12 CON, 3 ancient Irish and 1 ancient English horse samples are aligned to the root haplogroup (A6; as determined by Jansen *et al.* 2002). Variable nucleotides are indicated. A period denotes identity; * indicates a deletion; ? indicates missing sequence data (ancient sample F21122). Sequence positions based on the reference sequence (X79547; Xu & Amason 1994) are given above each column. The three ancient Irish sequences (F21099; Clare 1595 ± 78 cal BP; F21122; Sligo 1472 ± 56 cal BP; F21158; Waterford 32,255 ± 712 cal BP) plus the one English sequence (KIA26353; Derbyshire 1244 ± 41 cal BP) are shown in bold type. All samples have been assigned to one of the seven main haplogroups by means of its relative position in the median network (McGahern *et al.* 2006).

Name	15494	15495	15496	15521	15526	15528	15532	15534	15538	15540	15542	15584	15585	15597	15600	15601	15602	15603	15604	15615	15617	15631	15632	15635	15649	15650	15659	15666	15683	15703	15709	15718	15720	15726	15737	15740	Haplogroup
A6	T	C	A	G	T	C	C	C	A	A	C	C	G	A	G	T	T	T	G	A	T	A	T	C	A	C	A	G	T	C	C	A	G	T	A	A	
ID09, ID19, ID22, ID40, ID54, CON10	A
ID10	T	.	A	G	.	.	.	T	A
ID14	A
X79547, ID20, ID65, ID66, KB21, KB27, KB38, F21099	.	T	C	A	G	.	.	.	A	
ID21	T	.	.	G	A
ID26	C	A	A	
ID33	T	.	A	G	G	A	G	.	.	.	A	
ID34, KB01, KB29, KB44	.	T	C	G	.	.	.	A	G	.	.	.	A	
ID38	G	.	.	A	G	.	.	G	A	A	
ID47	G	.	.	A	G	A	A	
ID48, ID61	T	.	A	.	.	C	A	A	
ID59	T	.	A	.	A	T	.	.	A	A	
KB03	T	.	A	.	A	T	.	.	A	A	
KB11, KB16, CON04	A	.	A	A
CON03	C	G	A	A	
F21122	T	.	.	.	G	T	.	.	.	A	A	A	
F21158	T	T	.	.	A	A	A	
ID15, ID71, KB28	G	.	.	A	.	A	B	
ID16, ID70, CON14	G	.	.	.	A	B	
ID29, KB13	G	.	.	A	.	A	C	B	
CON08	G	.	.	A	.	A	A	B	
ID08, ID63, KB30, CON06	C	A	C	
ID36, KB02, KB14, KB15, KB32	A	C
ID57, ID72, KB09, KB10, KB33, KB40, KB42	A	C
ID69	T	A	A	C
ID73, CON05	A	A	C
ID01, ID07, ID32, ID46, CON11	C	G	A	C	A	A	D
ID02, ID04, ID06, ID11, ID39, ID45, ID53, ID56, ID67, ID68, ID74, CON09, CON12, KIA26353	C	G	D

Table 1 Continued.

Name	Haplogroup
15494	C
15495	C
15496	G
15521	G
15526	.
15528	.
15532	*
15534	T
15538	.
15540	.
15542	.
15584	.
15585	A
15597	.
15600	.
15601	.
15602	C
15603	C
15604	A
15615	.
15617	.
15631	.
15632	.
15635	.
15649	G
15650	A
15659	.
15666	.
15683	.
15703	.
15709	T
15718	.
15720	.
15726	.
15737	.
15740	D
ID03, ID05, ID17	D
ID12, ID18, ID24, ID52, ID64, KB08, KB34	D
ID23, ID25	D
ID51	D
ID60, KB37, KB46	D
ID62	D
KB07, CON13	D
CON02	D
KB04, KB06, KB12, KB20, KB39, KB41, KB45	E
KB05, KB17, KB43	E
KB18	E
KB35	E
KB22	F

documentation has suggested a Spanish origin for many of the founders of the ID (Fell 1991; O'Toole 2001), which is the putative origin for haplogroup D (Lopes *et al.* 2005). However, many breeds have contributed to the ID history, and it is therefore not surprising to find a broad genetic base for the ID population.

Most striking was the observation that only a few of the KB sequences belonged to the common European haplogroup D (Europe: 35%; KB Pony: 13%), and a high proportion were within haplogroup E (31%). Haplogroup E is rare globally and no other population analysed to date has been so highly represented herein (McGahern *et al.* 2006). Within all published horse data, haplogroup E contains 29 of 844 mtDNA sequences of which 66% are from populations in the British Isles (12 KB pony, six Shetland and one Thoroughbred). Haplogroup E also contains one Icelandic pony and two Cheju mtDNA sequences. Fisher exact tests of independence for haplogroup E concordance with British and European pony populations were both significant ($P < 0.0001$). Therefore, this haplogroup has a propensity to contain a high frequency of mtDNA sequences from small ponies distributed principally on the western fringe of Europe.

Acknowledgements

The authors thank the Irish Draught Horse Society, the Kerry Bog Pony Society, the Irish Horse Board and horse owners for access to samples. The Irish Draught Horse and the Kerry Bog Pony studies were financed independently by The Irish Department of Agriculture & Food. We thank Emma Teeling and John Flynn for access to samples and helpful comments, and Christopher Howe and Matthieu Visuete-Forster for comments and technical support. We thank Nigel Monaghan for provision of the ancient Irish horse specimens and Andrew Chamberlain for access to the CPC horse sample. C.J.E. is supported by the Irish Research Council for Science Engineering and Technology Basic Research Grant Scheme (project numbers SC/1999/409 and SC/2002/510). M.A.B. is supported by the McDonald Institute for Archaeological Research, University of Cambridge, and the Isaac Newton Trust. E.W.H. is supported by a Science Foundation Ireland, President of Ireland Young Researcher Award.

References

Bowling A.T., Del Valle A. & Bowling M. (2000) A pedigree-based study of mitochondrial D-loop DNA sequence variation among Arabian horses. *Animal Genetics* 31, 1–7.
 FAO (2000) *World Watch List for Domestic Animal Diversity*. Food and Agriculture Organisation of the United Nations (FAO), Rome.
 Fell A. (1991) *The Irish Draught Horse*. J.A. Allen & Company Ltd., London.
 Hill E.W., Bradley D.G., Al-Barody M., Ertugrul O., Splan R.K., Zakharov I. & Cunningham E.P. (2002) History and integrity of

- thoroughbred dam lines revealed in equine mtDNA variation. *Animal Genetics* **33**, 287–94.
- Ishida N., Hirano T. & Mukoyama H. (1994) Detection of aberrant alleles in the D-loop region of equine mitochondrial DNA by single-strand conformation polymorphism (SSCP) analysis. *Animal Genetics* **25**, 287.
- Jansen T., Forster P., Levine M.A., Oelke H., Hurler M., Renfrew C., Weber J. & Olek K. (2002) Mitochondrial DNA and the origins of the domestic horse. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 10905–10.
- Kavar T., Habe F., Brem G. & Dovc P. (1999) Mitochondrial D-loop sequence variation among the 16 maternal lines of the Lipizzan horse breed. *Animal Genetics* **30**, 423–30.
- Keyser-Tracqui C., Blandin-Frappin P., Francfort H.P., Ricaut F.X., Lepetz S., Crubez E., Samashev Z. & Ludes B. (2005) Mitochondrial DNA analysis of horses recovered from a frozen tomb (Berel site, Kazakhstan, 3rd Century BC). *Animal Genetics* **36**, 203–9.
- Lopes M.S., Mendonca D., Cymbron T., Valera M., da Costa-Ferreira J. & Machado Ada C. (2005) The Lusitano horse maternal lineage based on mitochondrial D-loop sequence variation. *Animal Genetics* **36**, 196–202.
- Luis C., Bastos-Silveira C., Cothran E.G. & Oom M.M. (2002) Variation in the mitochondrial control region sequence between the two maternal lines of the Sorraia horse breed. *Genetics and Molecular Biology* **25**, 309–11.
- MacHugh D.E. & Bradley D.G. (2001) Livestock genetic origins: goats buck the trend. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 5382–4.
- McGahern A., Bower M.A.M., Edwards, C.J., Brophy P.O., Sulimova G., Zakharov I., Vizuete-Forster M., Levine M., Li S., MacHugh D.E. & Hill E.W. (2006) Evidence for biogeographic patterning of mitochondrial DNA sequences in Eastern horse populations. *Animal Genetics*, doi: 10.1111/j.1365-2052.2006.01495.x
- Mirol P.E.A., Peral Garcia P. & Dulout F.N. (2002) Mitochondrial variability in the D-Loop of four equine breeds shown by PCR-SSCP analysis. *Genetics and Molecular Biology* **25**, 25–8.
- O'Toole H. (2001) *Characterisation of the Irish Draught Horse population in Ireland*. Thesis (MAGSc), University College Dublin, Dublin.
- Oakenfull A.E., Lim A.N. & Ryder O.A. (2000) A survey of equid mitochondrial DNA: implications for the evolution, genetic diversity and conservation of *Equus*. *Conservation Genetics* **1**, 341–55.
- Slavin M. (2005) RID inbreeding at danger levels. *Irish Farmers Journal*, 5 February, p. 5.
- Sokal R.R. & Rohlf F.J. (1995) *Biometry: The Principles and Practice of Statistics in Biological Research*. Freeman, New York.
- Storey N. (2005) *The Irish Draught Horse: A History*. The Collins Press, Cork, Ireland.
- Vila C., Leonard J.A., Gotherstrom A., Marklund S., Sandberg K., Liden K., Wayne R.K. & Ellegren H. (2001) Widespread origins of domestic horse lineages. *Science* **291**, 474–7.
- Woodman P., McCarthy M. & Monaghan N. (1997) The Irish Quaternary Fauna Project. *Quaternary Science Reviews* **16**, 129–59.
- Xu X. & Arnason U. (1994) The complete mitochondrial DNA sequence of the horse, *Equus caballus*: extensive heteroplasmy of the control region. *Gene* **148**, 357–62.

Supplementary Material

The following supplementary material is available for this article online from <http://www.blackwell-synergy.com/doi/full/10.1111/j.1365-2052.2006.01506.x>

Appendix S1 Supplementary methods.

Table S1 Horse populations included in mtDNA comparative analyses.

Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors.