Ancient Biomolecules, Vol. 2, pp. 267–280
Reprints available directly from the publisher
Photocopying permitted by license only

Ancient and Modern DNA in a Study of Horse Domestication

ADRIAN M. LISTER[a,*], MIRANDA KADWELL[b], LAURA M. KAAGAN[a], WILLIAM C. JORDAN[b], MARTIN B. RICHARDS[c] and HELEN F. STANLEY[b]

[a]*Department of Biology, University College London, London WC1E 6BT, UK;* [b]*Institute of Zoology, Regent's Park, London NW1 4RY, UK;* [c]*Institute of Molecular Medicine, John Radcliffe Hospital, Headington, Oxford OX3 9DS, UK*

(Received 19 December 1997; In final form 30 June 1998)

A 608 bp sequence from the mitochondrial control region was obtained from 29 domestic horses spanning a wide range of different breeds. These show extensive variation implying divergence of haplotypes long before domestication, and little congruence of haplotype with breed. Przewalski's horse is no more distant from the domestic breeds than the latter are among themselves, but our data are consistent with it being a wild relict of the same extensive Holarctic Pleistocene horse population from which, elsewhere, the ancestors of the domestics were derived. A calculated coalescence date of approximately 0.7–0.9 million years ago for the modern haplotypes corresponds closely with palaeontological estimates for the first entry of caballine horses into Eurasia. To pursue the ancestry of the domestics, 52 fossil or subfossil horse specimens were extracted and amplified for ancient DNA. These spanned a wide range of ages and preservation types, including specimens from Siberian permafrost, British Late Pleistocene caves, the La Brea tar pits, the Central Asian Neolithic, and the supposedly wild 'tarpan' of Eastern Europe. No control region sequences have been obtained, but three specimens yielded a 90 bp fragment of 16S rRNA gene: a bone from Kent's Cavern, Britain (approximately 12,000 BP), skin from a frozen Siberian carcass, (approximately 30,000 BP), and bone from the lectotype tarpan (approximately 100 BP).

Keywords: Ancient DNA, Breed genetics, Domestication, Horses, Mitochondrial DNA

INTRODUCTION

Animal domestication is a topic with a dual interest. First, the process of domestication provides an example of rapid change and diversification which, although 'artificially' caused, is analogous in many ways to evolution by natural selection in the wild (Darwin, 1859). Second, domestic animals have played a crucial role in human history, with numerous different breeds and presence in every region of human habitation. Yet beyond general outlines, the nature of the wild ancestral population(s), the time(s) and place(s) of domestication, and the diversification of breeds, are very poorly understood for many important taxa. One such is the domestic horse, *Equus caballus* L.

In this paper, we follow Gentry *et al.* (1996) in naming domestic horses *Equus caballus* L 1758, wild ones *Equus ferus* Boddaert 1785. There are many unanswered questions about the origins of

*Corresponding author. E-mail: a.lister@ucl.ac.uk.

domestic horses, and their relationships to putative wild ancestors. First, what is the status of Przewalski's horse of Mongolia? This form, named *Equus przewalskii* Poliakov 1881 or *E. ferus przewalskii*, has generally been regarded as a relict population of wild horse, presumably descended directly from the Late Pleistocene wild horses of Eurasia. However, some authors have suggested that it might, in fact, be a feral population derived from local domestic stocks (Bökönyi, 1974). Przewalski's horse is now essentially extinct in the wild, but various captive populations are derived from a founder stock of 12 animals taken from the wild in 1899–1901, plus one domestic animal which was unfortunately included in the breeding programme. However, a complete pedigree record exists (Volf, 1994), and it is possible to select animals for study which are descended purely from the original 12 wild-caught animals.

A second population which has been posited as a relict wild horse was the 'tarpan' of eastern Europe, of which the last example died in captivity in 1918 (Groves, 1994). Supposedly divided into 'forest' and 'steppe' types, the integrity of the 'tarpan' as a biological entity is uncertain. Some (e.g. Heptner *et al.*, 1988) have regarded it as a wild relict, perhaps at the western end of a cline with Przewalski's horse at the eastern end; others (e.g. Nobis, 1971) as essentially feral.

Relevant sources of comparison with these latter-day wild contenders are the last certain wild horses in the Eurasian fossil record. During the last glacial period, horses were extremely widespread across northern Eurasia (Guthrie, 1990; Kahlke, 1995). This population remained until roughly 12,000 BP, when the 'steppe-tundra' habitat started to fragment with the transition to the present interglacial. The fate of these horses is essentially unknown, but remnant populations must have formed the stocks from which the first domestic animals were derived, and one may survive as Przewalski's horse.

Among modern domestic breeds, there are various hypotheses concerning relationships between different types, and the possibility that some breeds may be more 'primitive' than others (Nobis, 1974; Edwards, 1993). For example, there is a supposed division between 'heavy', 'cold-blooded' horses such as the Shire and Suffolk Punch, and 'light', 'hot-blooded' breeds such as Arab and Thoroughbred. Of particular interest is the suggestion (Baker, 1993) that some of the 'native' ponies such as the Exmoor represent endemic relicts of original wild populations in their respective areas. Linked to this is the question of whether there was a 'multi-regional' origin of domestic horses, involving multiple domestication events in different regions, or a single event, from which domestic horses were taken to many parts of the world. An important focus of this research is in Central Asia, where the earliest horse domestication is believed to have occurred (Levine, 1990, 1996).

Previously published molecular studies of horse relationships have utilised a variety of techniques. Early studies on protein polymorphisms included those of Ryder *et al.* (1979) and Blokhuis and Buis (1979). George and Ryder (1986) undertook restriction enzyme analysis of mitochondrial DNA (mtDNA) from Przewalski's horse, three domestic breeds, and other *Equus* species. They deduced up to 0.5% sequence divergence among *E. ferus/caballus*. Xu and Arnason (1994) determined the complete mtDNA sequence of a domestic horse, discovering extensive heteroplasmy resulting from variable numbers of repeats of a short motif in the control region. Ishida *et al.* (1995) sequenced approximately 270 bp of the mtDNA control region for a small array of domestic breeds, Przewalski's horse, and other *Equus* species, calculating substitution rates and finding Przewalski's horse to be within the range of variation of domestic breeds.

In this study, we set out to compare mtDNA sequence data from a larger range of modern domestic horse breeds than has previously been attempted, as well as the 'wild' Przewalski's horse. In addition, we planned to sample ancient

material representing a range of fossil and historical populations possibly implicated in domestic horse origins: Late Pleistocene wild horses from Europe and Siberia, archaeological material from Central Asia around the time of domestication, and supposedly wild relicts such as the tarpan.

SPECIMENS

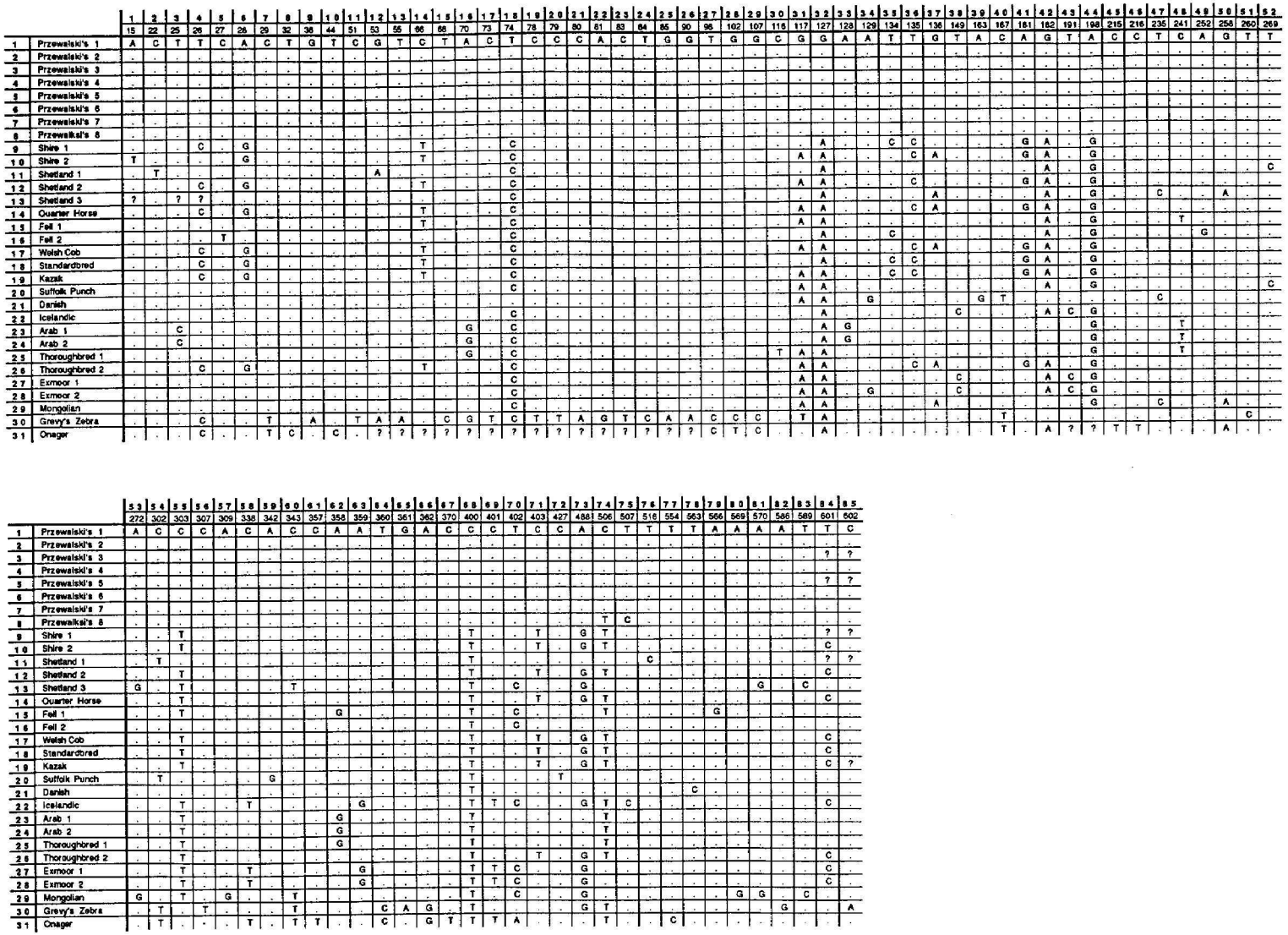
Modern samples were selected to cover as wide a range of modern breeds as possible, including 'native ponies' (Shetland, Icelandic, Fell and Exmoor), 'hot-blooded horses' (Arabs and Thoroughbreds), 'cold-blooded horses' (Shire and the Suffolk Punch), supposedly 'recent breeds' (Danish, Welsh Cob, Standardbred and Quarter Horse), and 'steppe horses' (from Mongolia and Kazakhstan). Blood or hair was obtained from living animals, taking care as far as possible to ensure correct breed identity. Twenty-one domestic animals were sampled in all, with sample sizes of between one and three per breed. In addition, we obtained blood from a sample of eight Przewalski's horses, plus one individual each of Grevy's zebra (*E. grevyi*) and onager (*E. hemionus onager*) as outgroups. Samples were obtained from the Institute of Zoology blood bank, by collecting from captive animals, and from colleagues who are equine specialists with access to particular breeds. Further information on specimen identity is available from LMK.

Because of the low expected success rate with ancient material, and the multiple hypotheses we wished to test, a total of 52 ancient specimens were collected, prepared and taken through the extraction and polymerase chain reaction (PCR) procedures. These were obtained from the following sources; full specimen details are available from AML:

1. *Late Pleistocene, permafrost, Siberia,* ranging in age from approximately 16–40 thousand years (ka). Three partial carcasses were sampled (two from Yakutia: Joger and Koch, 1994; one from Selerikan: Vereshchagin and Lazarev, 1977), frozen soon after death. Three other specimens were isolated limb bones from Kolyma (Sher *et al.*, 1979), also from frozen ground. The material spanned skin (five samples), hair (two samples), muscle (three samples), tendon (two samples) and bone (two samples).
2. *Late Pleistocene, cave deposits, Britain.* Tooth and bone from two sites, each dated approximately 12 ka BP: Gough's Cave, Somerset (Currant, 1986) and the Black Band at Kent's Cavern, Devon (Campbell and Sampson, 1971). The specimens were deposited in wet sludge but then held at relatively constant, cool temperature. From Gough's Cave, six teeth and four bones were sampled; from Kent's Cavern, eight teeth and one bone.
3. *Late Pleistocene, California.* Two bones and one tooth preserved in tar seeps at Rancho La Brea, approximately 14 ka BP (Marcus and Berger, 1984).
4. *Late Pleistocene, Kirillovka, Kazakhstan.* Fragments of skull bone preserved in sediment under dry but temperature-variable conditions, approximately 30–40 ka BP (Levine, personal communication).
5. *Bronze Age, Botai, Kazakhstan.* Five bones and six teeth from an archaeological site preserved under similar conditions to the above, approximately 5,500 BP (Levine, personal communication).
6. *Historical, European Russia.* Bone (three samples), tooth (one sample) and skin (one sample) of the holotype and lectotype Tarpan, approximately 100 BP (Groves, 1994).

RESULTS

Examination of a 608 bp sequence of the mitochondrial control region indicated considerable variation among the modern individuals (Fig. 1). Including the zebra and onager outgroups, a total of 85 variable sites were detected, of which four were transversions. Among the 29 living horses



		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52
		15	22	25	26	27	28	29	32	36	44	51	53	55	66	68	70	73	74	78	79	80	81	83	84	85	90	96	102	107	116	117	127	128	129	134	135	136	149	163	167	181	182	191	198	215	216	235	241	252	258	260	269
1	Przewalski's 1	A	C	T	T	C	A	C	T	G	T	C	G	T	C	T	A	C	T	C	C	C	A	C	T	G	G	T	G	G	C	G	G	A	A	T	T	G	T	A	C	A	G	T	A	C	C	T	C	A	G	T	T
2	Przewalski's 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
3	Przewalski's 3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
4	Przewalski's 4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
5	Przewalski's 5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
6	Przewalski's 6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
7	Przewalski's 7	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
8	Przewalksi's 8	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
9	Shire 1	.	.	.	C	.	G	.	.	.	.	.	.	.	T	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	C	C	.	.	.	G	A	.	G	.	.	.	.	.	.	.	.	.
10	Shire 2	T	.	.	.	.	G	.	.	.	.	.	.	.	T	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	A	A	.	.	.	C	A	.	.	.	G	A	.	G	.	.	.	.	.	.	.	.
11	Shetland 1	.	T	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	A	.	G	.	.	.	.	.	.	.	C
12	Shetland 2	.	.	.	C	.	G	.	.	.	.	.	.	.	T	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	A	A	.	.	.	C	.	.	.	G	A	.	G	.	.	.	.	A	.	.	.
13	Shetland 3	?	.	?	?	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	A	.	.	.	.	A	.	G	.	.	C	.	.	.	.	.
14	Quarter Horse	.	.	.	C	.	G	.	.	.	.	.	.	.	T	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	A	A	.	.	.	C	A	.	.	.	G	A	.	G	.	.	.	.	.	.	.	.
15	Fell 1	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	A	A	.	.	.	.	.	.	.	.	A	.	G	.	.	.	T	.	.	.	.
16	Fell 2	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	C	.	.	.	.	.	.	A	.	G	.	.	.	.	G	.	.	.
17	Welsh Cob	.	.	.	C	.	G	.	.	.	.	.	.	.	T	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	A	A	.	.	C	A	.	.	.	G	A	.	G	.	.	.	.	.	.	.	.
18	Standardbred	.	.	.	C	.	G	.	.	.	.	.	.	.	T	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	C	C	.	.	.	.	G	A	.	G	.	.	.	.	.	.	.	.
19	Kazak	.	.	.	C	.	G	.	.	.	.	.	.	.	T	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	A	A	.	C	C	.	.	.	.	G	A	.	G	.	.	.	.	.	.	.	C
20	Suffolk Punch	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	A	A	.	.	.	.	.	.	.	.	A	.	G	.	.	.	.	.	.	.	.
21	Danish	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	A	.	G	.	.	.	G	T	.	.	.	.	C	.	.	.	.	.	.	.
22	Icelandic	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	C	.	.	.	A	C	G	.	.	.	.	.	.	.	.
23	Arab 1	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	A	G	.	.	.	.	.	.	.	.	.	.	G	.	.	.	T	.	.	.	.
24	Arab 2	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	A	G	.	.	.	.	.	.	.	.	.	.	G	.	.	.	T	.	.	.	.
25	Thoroughbred 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	.	.	.	.	T	A	A	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	T	.	.	.	.
26	Thoroughbred 2	.	.	.	C	.	G	.	.	.	.	.	.	.	T	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	A	A	.	.	C	A	.	.	.	G	A	.	G	.	.	.	.	.	.	.	.
27	Exmoor 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	A	A	.	.	.	.	C	.	.	.	A	C	G	.	.	.	.	.	.	.	.
28	Exmoor 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	A	A	.	G	.	.	C	.	.	.	A	C	G	.	.	.	.	.	.	.	.
29	Mongolian	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	A	A	.	.	.	A	.	.	.	.	.	.	G	.	.	C	.	.	A	.	.
30	Grevy's Zebra	.	.	.	C	.	.	T	.	A	.	T	A	A	.	C	G	T	C	T	T	A	G	T	C	A	A	C	C	C	.	T	A	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	C	.
31	Onager	.	.	.	C	.	.	T	C	.	C	.	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	C	T	C	.	.	A	.	.	.	.	.	.	.	T	.	A	?	?	T	T	.	.	.	A	.	.

		53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85
		272	302	303	307	309	338	342	343	357	358	359	360	361	362	370	400	401	402	403	427	488	506	507	516	554	563	566	569	570	586	589	601	602
1	Przewalski's 1	A	C	C	C	A	C	A	C	C	A	A	T	G	A	C	C	C	T	C	C	A	C	T	T	T	T	A	A	A	A	T	T	C
2	Przewalski's 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	C	.	.	.	.	.	.	.	.	.	.
3	Przewalski's 3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?
4	Przewalski's 4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
5	Przewalski's 5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?	?
6	Przewalski's 6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
7	Przewalski's 7	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
8	Przewalksi's 8	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	C	.	.	.	.	.	.	.	.	.	.
9	Shire 1	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	T	.	G	T	.	.	.	.	.	.	.	.	.	?	?
10	Shire 2	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	T	.	G	T	.	.	.	.	.	.	.	.	.	C	.
11	Shetland 1	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	?	?
12	Shetland 2	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	T	.	G	T	.	.	.	.	.	.	.	.	.	C	.
13	Shetland 3	G	.	T	.	.	.	.	T	.	.	.	.	.	.	.	T	.	C	.	.	G	.	.	.	.	.	.	.	G	.	C	.	.
14	Quarter Horse	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	T	.	G	T	.	.	.	.	.	.	.	.	.	C	.
15	Fell 1	.	.	T	.	.	.	.	.	.	G	.	.	.	.	.	T	.	C	.	.	.	T	.	.	.	.	G	.	.	.	.	.	.
16	Fell 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
17	Welsh Cob	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	T	.	G	T	.	.	.	.	.	.	.	.	.	C	.
18	Standardbred	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	T	.	G	T	.	.	.	.	.	.	.	.	.	C	.
19	Kazak	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	T	.	G	T	.	.	.	.	.	.	.	.	.	C	?
20	Suffolk Punch	.	T	.	.	.	.	G	.	.	.	.	.	.	.	.	T	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.
21	Danish	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.
22	Icelandic	.	.	T	.	.	T	.	.	.	.	G	.	.	.	.	T	T	C	.	.	G	T	C	.	.	.	.	.	.	.	.	C	.
23	Arab 1	.	.	T	.	.	.	.	.	.	G	.	.	.	.	.	T	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
24	Arab 2	.	.	T	.	.	.	.	.	.	G	.	.	.	.	.	T	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
25	Thoroughbred 1	.	.	T	.	.	.	.	.	.	G	.	.	.	.	.	T	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
26	Thoroughbred 2	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	T	.	G	T	.	.	.	.	.	.	.	.	.	C	.
27	Exmoor 1	.	.	T	.	.	T	.	.	.	.	G	.	.	.	.	T	T	C	.	.	G	.	.	.	.	.	.	.	.	.	.	C	.
28	Exmoor 2	.	.	T	.	.	T	.	.	.	.	G	.	.	.	.	T	T	C	.	.	G	.	.	.	.	.	.	.	.	.	.	C	.
29	Mongolian	G	.	T	.	G	.	.	T	.	.	.	.	.	.	.	T	.	C	.	.	G	.	.	.	.	.	.	G	G	.	C	.	.
30	Grevy's Zebra	.	T	.	T	.	.	.	T	.	.	.	C	A	G	.	T	.	.	A	.	G	T	.	.	.	.	.	.	.	G	.	.	A
31	Onager	.	T	.	.	.	T	.	T	T	.	.	C	.	G	T	T	T	A	.	.	.	T	.	.	C	.	.	.	.	.	.	.	.

FIGURE 1 The 85 variable sites in 608 bp of the modern horse mtDNA control region. Top row, numbering of variable sites; second row, position from start of control region. '?' indicates missing data.

of the ingroup (Przewalski's plus domestics), there were 19 different haplotypes and 54 variable sites (including one transversion), representing 8.9% of the total sequence, of which 34 sites were phylogenetically informative for parsimony analysis (Fig. 1). Distribution of variable sites was not uniform along the 608 bp horse fragment (Fig. 1), 48% of them (26 sites) occurring within the first 32% (198 bp) of sequence. This indicates the existence of a 'hypervariable region' toward the 5′ end of the control region, as in humans, cattle and other species (Loftus *et al.*, 1994; Krings *et al.*, 1997).

Across the entire 608 bp horse fragment, average sequence divergence (Kimura 2-parameter distances) among all individuals in the ingroup was 1.9%, with a maximum between individuals of 3.4%, the genotypes differing by up to 21 substitutions.

Analysis by the maximum parsimony method produced 68 most-parsimonious trees of 117 steps (including outgroups). The trees all have similar topologies, and the network diagram (Fig. 2) includes many of the most parsimonious trees supported by the data (Bandelt *et al.*, 1995). A neighbour-joining tree (Fig. 3) has very similar structure to the network, while a maximum likelihood tree differs in a few points (Fig. 4). In all trees, the Przewalski's horse haplotypes and a Danish breed individual are closely allied, and

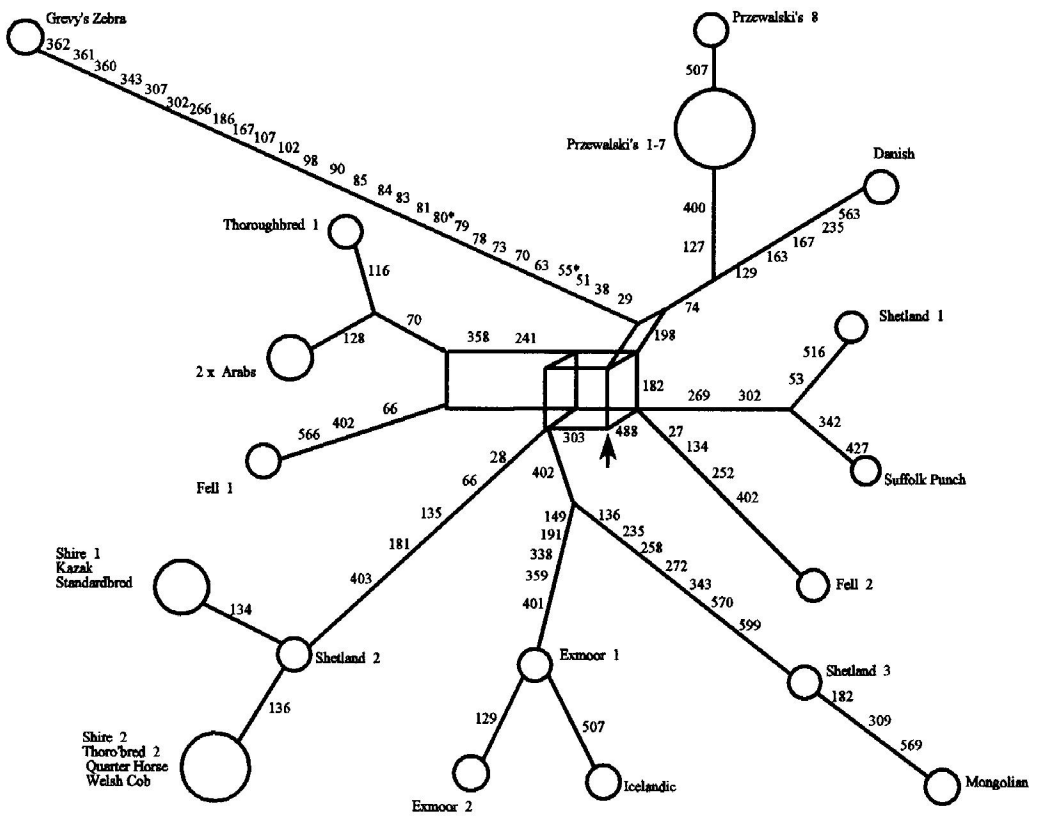


FIGURE 2 Reduced median network based on 563 bp of the mtDNA control region for living horses, with Grevy's zebra as outgroup, constructed after the method of Bandelt *et al.* (1995). The circles represent sequence haplotypes and the links are labelled according to base positions (cf. Fig. 1) where transitions have occurred; transversions are designated by an asterisk. The areas of the circles are proportional to the frequencies of the sequence haplotypes, but branch lengths are not to scale. The central cube in the network indicates alternative topologies depending on the position of the root, which may be placed at any of its nodes. The root position assumed for purposes of calculation is indicated by an arrow.

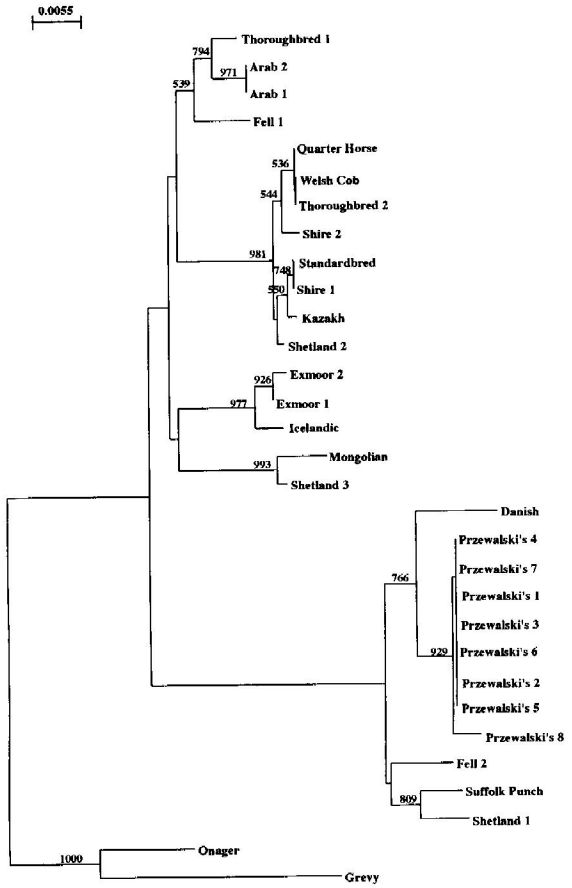


FIGURE 3 Neighbour-joining tree based on 608 bp of the mtDNA control region for living horses, with onager and Grevy's zebra as outgroups. Support values derived from 1000 bootstrap replicates are shown for values over 500 (= 50% support).

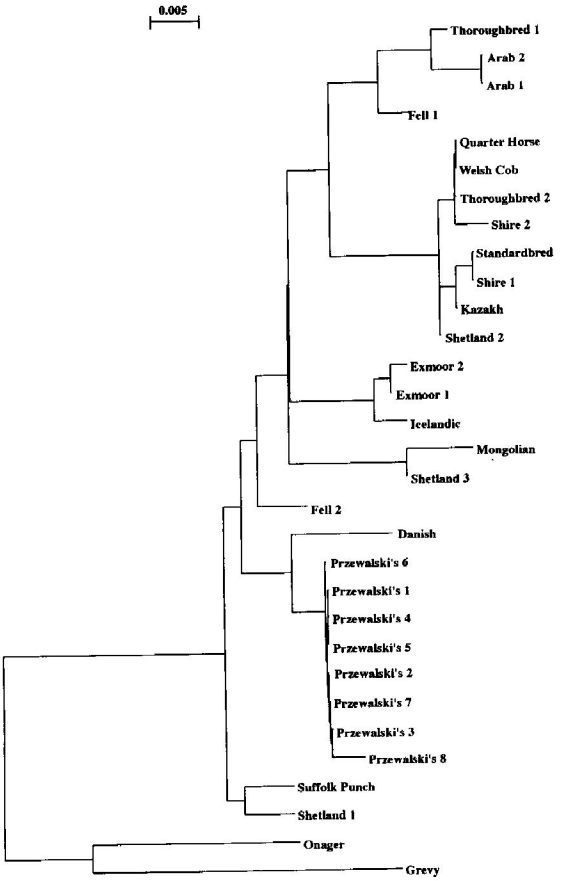


FIGURE 4 Maximum likelihood tree based on 608 bp of the mtDNA control region for Living horses, with onager and Grevy's zebra as outgroups.

in most trees are close to a Fell pony, a Suffolk Punch and a Shetland. All other individuals form a crown group with certain consistent subgroupings of individuals within it (Figs. 2–4; see later).

To calculate the haplotype divergence times within *E. caballus/ferus*, a palaeontological estimate of around 2 million years (Ma) has been taken for the zebra–horse split. This assumes that the divergence must have taken place some time shortly before the earliest fossil caballine at approximately 1.8 Ma (Forstén, 1992). The date cannot have been earlier than 4 Ma, the age of the earliest known fossil *Equus* (Forstén, 1992).

Divergence times within the ingroup were estimated using a molecular clock approach. The log likelihoods of maximum likelihood trees computed with and without a molecular clock assumption were −1472.87 and −1459.65, respectively, a non-significant difference at the 5% level. Assuming a molecular clock, UPGMA (unweighted pair-groups method) gave branch lengths of 0.03341 for the outgroup–ingroup node, and 0.01477 for the most recent common ancestor of the ingroup (domestic plus Przewalski's horses). The 2 Ma palaeontological estimate of the former thus yields a coalescence date of approximately 0.88 Ma for the ingroup.

Corrected transition/tranversion ratio in the horse study is 18.5 : 1, and the calculated substitution rate 1.67% per million years, surprisingly low in comparison with average figures from cattle (15%: Bradley *et al.*, 1996) and humans (11.8%: Stoneking *et al.*, 1992). The higher rates of the latter studies are partly accounted for by their concentration on the hypervariable region, and in the case of cattle may also be inflated by an unrealistically young estimate of 1 Ma for the *Bos–Bison* split as a calibration point (Sher, 1997), but the differences are still striking.

We also attempted Forster *et al.*'s (1996) method, calculating the average number of substitutions from the ingroup individuals to an assumed root in Fig. 2. The average of 6.9 transitions in the caballine branches, out of the 563 bp sampled (1.23%), in relation to the substitution rate calculated above, indicates a coalescence date of approximately 0.74 Ma, comparable to the UPGMA estimate.

Turning to our ancient extracts, we first tested amplification of cytochrome *b*, as these primers work well in many modern samples and the size seemed appropriate for ancient material. However, we had limited success: PCR product was obtained and cloned from three specimens (Yakutian frozen tendon, Selerikan frozen skin and historical tarpan bone), but no recognisable cytochrome *b* sequence was obtained. We therefore went on to design control region primers, targeting sequences ranging from 90–300 bp and therefore including smaller fragments than the cytochrome *b* ones. The primers were designed to amplify the most variable regions observed in the modern DNA, and were first tested on high-quality modern DNA to confirm they were successfully amplifying the expected fragment. Primers were also tested for their ability to amplify DNA of other species commonly in use in our laboratory (e.g. camelids), and a range of non-specific PCR products was observed.

The following control region primers were tested in extant samples (cf. Fig. 6); those giving the best results (i.e. a single, strong PCR product) are underlined: L110/HEq3; ThrL/Eq.H151; L19/Eq.H151; L19/H.Eq3; L292/HEq4; L292/HEq5; L292/DLH; L292/H436; L292/EqH16099.

Each of these five most promising control region primer pairs was tried on each of the 52 ancient specimens, but all amplifications gave either smears or multiple bands, none of which appeared to be of the correct size, apart from a faint band for the 130 bp product (L292/H436) on skin from the Selerikan frozen Siberian horse for which Taylor (1996) had published a short 16S rRNA sequence. We subsequently tried all our specimens with the primers described by Taylor, and of all the primer pairs we tested, the 16S primers were the most successful, amplifying product in a bone from Kent's Cavern, one of the Tarpans, and reproducibly amplifying some of the Yakutian and Selerikan Siberian frozen specimens. In view of their apparent sensitivity, the 16S primers were used to test for the presence of DNA in all of our extracts and, if positive with 16S, further attempts were made to amplify control region fragments, but without success.

In summary, the following sequence data were obtained from ancient specimens:

- Kent's Cavern (Devon) bone: 90 bp 16S rRNA sequence, fully homologous with horse.
- Tarpan (lectotype) bone: 90 bp 16S rRNA sequence, fully homologous with horse.
- Selerikan (Siberian frozen) skin: 90 bp 16S rRNA sequence, some sequence blocks homologous with horse.
- Yakutia (Siberian frozen) tendon: cytochrome *b* PCR leading to anonymous sequence.
- Selerikan (Siberian frozen) skin: cytochrome *b* PCR leading to anonymous sequence.

The general paucity of positive results means that only limited comments can be made on the relative success rate in different preservational conditions, tissues and genes. The following can be regarded as isolated data points contributing

to a wider picture to be developed from the results of different studies and laboratories:

- 16S primers consistently gave better product on the same samples than control region primers.
- Siberian frozen specimens (Yakutia and Selerikan) were the most reproducibly amplifiable.
- Within Siberian specimens, the order of success with different tissue types, by specimen, was as follows. Selerikan carcass: skin > muscle > hair; Yakutia carcass: tendon > skin > muscle; Kolyma (isolated body parts): skin > bone and tendon.
- From Gough's Cave, none of our ten specimens yielded product, despite the success of Bailey *et al.* (1996) with one aurochs bone from the same deposit.
- From Kent's Cavern, one of our nine specimens yielded product and sequence, despite the failure of Bailey *et al.* (1996) with an aurochs bone from the same deposit.
- From Kent's Cavern, eight teeth failed while one bone was successful, despite the generalisation of DeGusta *et al.* (1994) that mammalian tooth preserves DNA better than bone.
- From Botai, younger than Kent's Cavern but with a continental, dry climate, none of our eleven specimens yielded product.
- From Rancho La Brea (tar pits), none of our three specimens yielded product, despite the success of Janczewski *et al.* (1992) with three out of six sabre-tooth cat bones.

The Kent's Cavern bone from which sequence was obtained was radiocarbon dated by the Oxford AMS system, to give a precise age of the recovered DNA. The result, consistent with the provenance of the bone from the Late Upper Palaeolithic 'Black Band' (Campbell and Sampson, 1971) was 12,250 ± 110 BP (OxA-5692). Six clones produced good 16S rRNA sequence (Fig. 5), with four variable sites in approximately 90 bp. For the Tarpan, only two of six clones produced good sequence (Fig. 5), differing at three positions. The variations among clones of the same individual are most likely due to misincorporations during PCR, though some could be due to heteroplasmy. The Tarpan and Kent's sequences have been aligned with a Przewalski's sequence obtained in this study, a domestic horse from Xu and Arnason (1994), the Selerikan sequence from Siberia published by Taylor (1996), and as an outgroup, two published onager sequences (Fig. 5). The onagers both differ from the horses by only a single transition, indicating the slow evolution of this gene and its inappropriateness for intraspecific studies.

```
E.caballus                  GAACAAAACAACCTCCGAGTGATTTAAATCCAGACTAACCAGTCAAAATATATAAT-CACTTATTGATCCAAACCATTGAT-CAACGGAACAAGTTACCCT
Przewalski's 1               .A.....-.....-.....
Selerikan                     .....-.....-.....

Kent's Cavern 1855 D1         .....N.....N.....N.....N.....-..N.....-T.....
Kent's Cavern 1855 D9         .....N.....N.....N.....-..N.....-T.....
Kent's Cavern 1855 D5         ...N.....N.....A.....-.....-.....T.....
Kent's Cavern 1855 D10        .....N.....A.....-.....-.....T.....
Kent's Cavern 1855 D11        .....N.....C.....-.....-.....N.....
Kent's Cavern 1855 D12        .....N.....-.....-.....N.....

Tarpan T3 A8                  .....T.....N.....A.....-.....-.....
Tarpan T3 18                  .....T.....A.....G.....N.....G.....-.....

E hemionus (Milinkovitch)     .....T.....C.....-.....-.....
E hemionus (Pääbo)            .....T.....G...-.....T.....-.....
```

FIGURE 5 Aligned sequences of a 90 bp fragment of the mtDNA 16S rRNA gene in ancient and modern horses. From the top: reference sequence of domestic horse from Xu and Arnason (1994); Przewalski's horse 'Oska' (this paper); Selerikan (Siberian frozen) horse from Taylor (1996); six clones of bone 1855 from Kent's Cavern (British Pleistocene; this paper); two clones from Tarpan lectotype no. 94535 (this paper); two onager (*E. asinus*) outgroups from Milinkovitch *et al.* (1993) and Höss and Pääbo (1993).

Indeed, the invariant positions among the Kent's clones show no differences from the domestic horse or from the Selerikan specimen; only the Przewalski's horse has a single transition. The two tarpan clones both differ from other sequences at two positions (one a transversion), but these results need replicating. There is effectively, therefore, no phylogenetic information in these data, which were indeed collected only as an indicator of the preservation of ancient DNA.

DISCUSSION

Overall, the amount of sequence divergence among modern breeds is greater than that could have arisen within any plausible timescale of domestication (approximately 5 ka by current archaeological estimates: Levine, 1990), and clearly reflects more ancient haplotype diversity. The deep divergences are reflected in the long branches in Fig. 2, some of which are terminated by clusters of similar haplotypes indicating more recent radiation. The observed variation corresponds to the results of Marklund *et al.* (1995), who detected extensive mtDNA control-region polymorphism within each of five breeds of horse using PCR-SSCP (single strand conformation polymorphism) screening. A similar picture was found by Loftus *et al.* (1994) in the control region of domestic cattle.

A few apparently meaningful clusters of modern breeds occur in our trees and network (Figs. 2–4). Unsurprisingly, seven Przewalski's individuals are sequence identical, the eighth differing at only one position. Two Exmoor ponies group with one Iceland pony; two Arabs group with one Thoroughbred. However, the overriding picture is that haplotypes and breeds are not congruent. For example, three Shetland ponies (one miniature, two thought to be pure bred) have quite different haplotypes, and conversely, a single haplotype is found in individuals of quite different breeds (Shire, Thoroughbred, Quarter Horse and Welsh Cob), although a possible factor linking this group is partial Spanish ancestry (Edwards, 1993). There is no evidence in our results of a common genetic heritage among British pony breeds (Exmoor and Shetland), or among so-called 'heavy' or 'cold-blooded' horses (Shire and Suffolk Punch). The pony phenotype could be due to a semi-wild existence in similar conditions. The Shires and Suffolk Punch have largely different ancestry, the former from forestry horses of the Low Countries, the latter from the extinct Norfolk Roadsters, though both share some Flemish blood (Edwards, 1993).

Further, it is unsurprising that our trees do not show the same inter-breed topology as those produced by other authors who have included some of the same breeds. For example, Ishida *et al.* (1995) found a link between Przewalski's horse and Thoroughbreds, with Mongolian native horses (not necessarily the same breed as our Mongolian sample) forming a sister-group to these. George and Ryder's (1986) tree linked Przewalski's to Quarter Horse, with Arabs as sister-group to these. Possibly significant in their data is a close link between the American breeds: Quarter horse, Morgan and Peruvian Paso. In general, however, the marked intra-breed variation revealed by our study indicates that tree topologies will depend largely on the particular animals chosen.

There is no evidence in our data that Przewalski's horse had a long and separate history from domestics. Przewalski's horse appears close to the root in some analyses, but no more distant from domestic lineages than the latter are from each other. Moreover, maximum corrected sequence divergence between our Przewalski's sample and the domestic horses is 3.0%, similar to the 3.4% found among the domestics themselves. Nonetheless, this does not necessarily indicate that Przewalski's horse is feral as Ishida *et al.* (1995) have implied on more limited mtDNA data. Our data are consistent with the palaeontological evidence for a vast population of wild caballine horse across Europe, Asia and

North America through much of the Middle and Late Pleistocene. Caballine horses, as *E. ferus mosbachensis*, first entered Eurasia around 1 Ma–700 ka BP, and finally suffered population collapse around the Pleistocene/Holocene transition (approximately 15–8 ka BP: Forstén, 1992). Interestingly, our estimated coalescence dates correspond closely to the estimate for the former event, suggesting that genetic diversification began with the first expansion of *E. ferus* in Eurasia, and was not significantly depleted by the population crash at the Pleistocene/Holocene boundary. The Middle to Late Pleistocene population, doubtless with some regional structure, must have accumulated considerable mtDNA variation, and from 'samples' of this population were derived both Przewalski's horse and the ancestors of modern domestic breeds.

Bailey *et al.* (1996) came to a similar conclusion for cattle, stating that 'the haplotypes pre-date breed establishment'. Vila *et al.* (1997) similarly deduced for dogs that 'mitochondrial haplotype diversity could not be partitioned according to breeds', and 'many breeds contained representatives of more than one dog haplotype grouping... most breeds show evidence of a genetically diverse heritage'. This does not preclude the possibility that some domestic horse breeds remain phenotypically closer to the ancestral type than others. Breeds such as the 'native ponies', even if derived from domestic stock, may have originated relatively early and remained relatively untouched by outcrossing. For other breeds, extensive historical cross-breeding has contributed to present day genetic diversity and to the lack of congruence between haplotypes and breeds. It should be remembered that our study is based only on the female-inherited mitochondrial genome. Ideally, this should be complemented in future work by data from the nuclear genome. Since translocation of horses for cross-breeding purposes has historically entailed mostly males, mitochondrial markers should if anything de-emphasise the effect of cross-breeding and retain the imprint of local differentiation. The absence of the latter in our data is therefore the more striking.

The modern data do not distinguish single from multi-regional domestication events, though the extent of the modern haplotype diversity probably reflects an input of wild animals from different areas. On the one hand, it is unlikely that all domestic horses sprang from a single, small, local population. On the other, the independent domestication of wild animals in very distant parts of the world might have left a more coherent signature in the mitochondrial trees and network. Certainly, there is no support in our data for the widespread belief that modern phenotypic groupings are derived from divergent wild stocks such as a supposed 'tundra horse', 'plateau or desert horse', 'forest horse', 'steppe horse', and endemic ponies (Epstein, 1971; Edwards, 1993; Nobis, 1974). An intermediate model may therefore be most likely, domestic horses having arisen from wild stock distributed over a moderately extensive geographical region, large enough to have contained within it considerable pre-existing haplotype diversity.

The incorporation of ancient DNA sequences from potential ancestral populations is clearly desirable in addressing these issues. From hundreds of amplifications of over 50 ancient samples, we obtained usable sequence only for the 16S rRNA gene from a frozen Siberian skin specimen, a bone of a wild horse approximately 12,250 years old from a British cave deposit, and skull bone from the lectotype Tarpan approximately 100 years old. Other samples gave varying hints of DNA preservation, providing some data on the relative tractability for ancient DNA of different tissue types and preservation conditions. However, we have so far not succeeded in obtaining control region sequence which could be incorporated with the results from modern horses.

Our modern dataset provides a framework for more detailed analysis of between- and within-breed variation, and also for the incorporation of ancient sequences. This study needs to be

enlarged by accruing larger domestic samples of accurate pedigree (using both mtDNA and microsatellites), and by further pursuing those ancient specimens which have given indications of DNA preservation, and further samples from the same deposits.

MATERIALS AND METHODS

Modern DNA was extracted following the procedure of Pääbo (1986) but without dithiothreitol; extraction blanks were included in all analyses. In preliminary experiments, we found little variation among the domestic breeds in the mitochondrial cytochrome *b* gene. We therefore focussed on the control region, which in mammals generally has a higher evolutionary rate. A total of 31 individuals were analysed including two outgroups. Fragments of the 5′ end of the control region were routinely amplified using primers ThrL 16272 (numbers follow human reference sequence: modified from Hoelzel *et al.*, 1991) and EqH16099 (with respect to the horse sequence of Xu and Arnason 1994). Sequencing was carried out using primers EqH16099, H.Eq 4 or H. Eq5 and H.Eq3 (Fig. 6; primer sequences available from HFS). Both manual and automatic sequencing were employed and duplicate or triplicate reactions performed. Sequences of modern domestic and Przewalski's horses have been deposited in Genbank with accession numbers AF072975-996.

Early in the project we established a laboratory dedicated as an ancient DNA facility. The whole laboratory was thoroughly cleaned before each work session. Samples were handled under an enclosed hood, which was cleaned with bleach every time a new sample was to be introduced. The inside of the hood was regularly UV-irradiated overnight. Sterile, disposable plasticware was employed wherever possible, and used only once. All accessible parts of the freezer mill, and other non-disposable items, were bleached and washed in sterile, filtered water, and finally in ethanol, between each extraction. They were also periodically UV-irradiated. All solutions were prepared and stored under sterile conditions.

Bone or tooth surfaces were sanded, then roughly crushed with a hammer or cut into small pieces with a saw. External surfaces were washed with ethanol and water, air dried and crushed to a fine powder in a freezer mill (Glen Creston). Soft tissues were manually cut into small pieces. Bleach was not used on the samples as it is potentially damaging to DNA (H. Poinar, personal communication). Two main DNA extraction procedures were tested. The first entailed modifications of Pääbo (1986) and Pääbo *et al.* (1988), with high concentrations of proteinase K to digest tough ancient tissue, followed by phenol/chloroform purification. The second, which we found to be more successful (i.e. yielded DNA with a higher success rate in PCR), was the silica/guanidinium thiocyanate method of Höss and Pääbo (1993), which has strong DNA binding ability, fewer steps, and avoids potentially damaging reagents such as phenol. All extraction solutions were tested for contamination by running a 'blank' through the entire extraction and PCR procedure. During our experiments, contamination was never observed in any PCR blank.

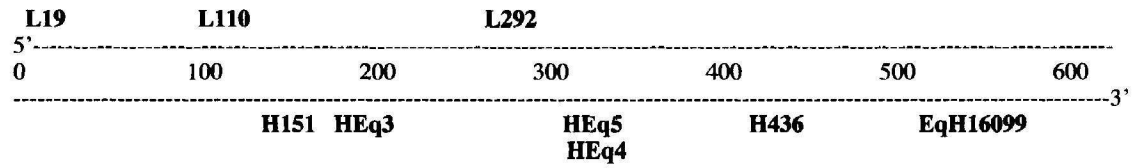


FIGURE 6 Positions of equid-specific primers designed for ancient DNA work. Numbering begins at the start of the mtDNA control region. L and H refer to the light and heavy strands of mtDNA, respectively.

Amplification of ancient DNA was performed using standard PCR procedures but with increased numbers of cycles (35) and cycle times, a high concentration of *Taq* DNA polymerase (IGI, 2 Units per sample), and the addition of bovine serum albumin (BSA). We experimented with various concentrations of *Taq*, BSA and Mg^{2+}, and with Amplitaq Gold polymerase (PE Applied Biosystems), but there was little influence on the results. Where there was no product, we tried in all cases varying the concentration of the extract twofold or fourfold. In a few cases (e.g. the Siberian metacarpal 3916-95) increased concentration gave improved PCR results, but in most cases it had no effect. It was found that DNA extracts deteriorated over several months, even at −20°C, so it is best to use an extract soon after extraction. Appropriate controls (i.e. PCR blank and extraction blank) were always performed in the PCR reaction, and the whole blank reaction was checked on an agarose gel to confirm the absence of contamination.

On ancient material, PCRs were carried out for three regions of mtDNA: a 176 bp fragment of the cytochrome *b* gene (primers L14724/H14900); short fragments of the control region using horse-specific primers (see Results); a 90 bp fragment of 16S rRNA (Taylor, 1996).

Because of the degraded nature of ancient DNA, PCR artefacts are a particular problem and direct manual sequencing of direct PCR product was unsatisfactory. We therefore cloned our ancient PCR products using the Stratagene PCR-Script Amp SK(+) cloning kit. For this approach, blunt-ended PCR products were generated using *Pfu* DNA polymerase, which has a lower misincorporation rate than *Taq* DNA polymerase. It was generally more difficult to generate PCR products with *Pfu* DNA polymerase. We therefore amplified with *Taq* DNA polymerase first, cut the product from an agarose gel, eluted the DNA and re-amplified with *Pfu* DNA polymerase before cloning.

The plasmid clones were sequenced using T3 or T7 primers (located within the vector) and a *Taq* FS Prism cycle sequencing kit (PE Applied Biosystems). Reactions were run on an automated sequencer and the data analysed using Sequencher software (Gene Codes). We sequenced between six and nine clones from each specimen, but not all gave recognisable sequence. Because specimen handling is the most likely source of contamination, all sequences were checked against human, but none was positive in this respect.

A reduced median network was constructed according to the method of Bandelt *et al.* (1995). The sequences were analysed between positions 27 and 589. An initial run with all characters suggested that positions 117 and 506 were incompatible with a large number of other positions, and the data were therefore re-run without them. The Grevy's zebra was included as outgroup.

Sequence data were also analysed by the neighbour-joining and maximum likelihood methods using PHYLIP 3.572c (Felsenstein, 1993), and the maximum parsimony method using PAUP 3.1.1 (Swofford, 1993). Distances were corrected using the Kimura two-parameter model. In addition, maximum likelihood trees were constructed using the Hasegawa–Kishino–Yano (1985) model (HKY85), with and without the imposition of a molecular clock, and a likelihood ratio test performed. The unweighted pair-groups method (UPGMA) was then used to obtain distances at various nodes in the trees. These analyses were performed using PAUP* (Swofford, 1998).

Acknowledgements

We thank Sue Baker, Gennady Baryshnikov, Kim Bryan, Andrew Currant, Daniel Debaye, Peter Lazarev, Marsha Levine, Ann Oakenfull, Igor Pavlinov, Mike Scott, Christopher Shaw, Andrei Sher and Alma Swan for help in obtaining samples, Vincent Macaulay and Cathy Kerr for assistance in preparing the network diagram, Dave Swofford for the use of the PAUP* test

version, and Mike Bruford and Kate Byrne for discussion.

References

Bailey, J.F., Richards, M.B., Macaulay, V.A., Colson, I.B., James, T., Bradley, D.G., Hedges, R.E.M. and Sykes, B.C. (1996) Ancient DNA suggests a recent expansion of European cattle from a diverse wild progenitor species. *Proc. R. Soc. Lond. ser. B* **263**, 1467–1473.

Baker, S. (1993) *Survival of the Fittest: A Natural History of the Exmoor Pony.* Exmoor Books.

Bandelt, H.J., Forster, P., Sykes, B.C. and Richards, M.B. (1995) Mitochondrial portraits of human populations using median networks. *Genetics* **141**, 743–753.

Blokhuis, H.J. and Buis, R.C. (1979) Genetic relationships between breeds of horses and ponies in The Netherlands. *Anim. Blood Grps. Biochem. Genet.* **10**, 27–38.

Bökönyi, S. (1974) *The Przewalsky Horse.* Souvenir Press, London.

Bradley, D.G., MacHugh, D.E., Cunningham, P. and Loftus, R.T. (1996) Mitochondrial diversity and the origins of African and European cattle. *Proc. Natl. Acad. Sci. USA* **93**, 5131–5135.

Campbell, J.B. and Sampson, C.G. (1971) A new analysis of Kent's Cavern, Devonshire, England. *Univ. Ore. Anthrop. Pap.* **3**.

Currant, A.P. (1986) The lateglacial mammal fauna of Gough's Cave, Cheddar, Somerset. *Proc. Univ. Bristol Spelaeol. Soc.* **17**, 286–304.

Darwin, C.R. (1859) *On the Origin of Species by Means of Natural Selection.* John Murray, London.

DeGusta, D., Cook, C. and Sensabaugh, G. (1994) Dentine as a source of ancient DNA. *Ancient DNA Newsl.* **2**(1), 13.

Edwards, E.H. (1993) *Horses.* Dorling Kindersley, London.

Epstein, H. (1971) *The Origin of the Domestic Animals of Africa*, Vol. 2. Africana, New York.

Felsenstein, J. (1993) *PHYLIP (Phylogeny Inference Package).* University of Washington, Seattle.

Forstén, A. (1992) Mitochondrial-DNA time-table and the evolution of *Equus*: comparison of molecular and paleontological evidence. *Ann. Zool. Fennici* **28**, 301–309.

Forster, P., Harding, R., Torroni, A. and Bandelt, H.-J. (1996) Origin and evolution of Native American mtDNA variation: a reappraisal. *Am. J. Hum. Genet.* **59**, 935–945.

Gentry, A., Clutton-Brock, J. and Groves, C.P. (1996) Proposed conservation of usage of 15 mammalian specific names based on wild species which are antedated by or contemporary with those based on domestic animals. *Bull. Zool. Nomencl.* **53**, 28–37.

George, M. and Ryder, O.A. (1986) Mitochondrial DNA evolution in the genus *Equus*. *Mol. Biol. Evol.* **3**, 535–546.

Groves, C.P. (1994) Morphology, habitat and taxonomy. In, *Przewalski's Horse – The History and Biology of an Endangered Species* (L. Boyd and K.A. Houpt, Eds.), pp. 39–59, SUNY Press, New York.

Guthrie, R.D. (1990) *Frozen Fauna of the Mammoth Steppe.* University of Chicago Press, Chicago.

Hasegawa, M., Kishino, H. and Yano, T. (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**, 160–174.

Heptner, V.G., Nasimovitch, A.A. and Bannikov, A.G. (1988) *Mammals of the Soviet Union. Vol. 1: Artiodactyla and Perissodactyla.* Smithsonian Institution and National Science Foundation, Washington D.C.

Hoelzel, D.R., Hancock, J.M. and Dover, G.A. (1991) Evolution of the cetacean mitochondrial D-loop region. *Mol. Biol. Evol.* **8**, 475–493.

Höss, M. and Pääbo, S. (1993) DNA extraction from Pleistocene bones by a silica-based purification method. *Nucl. Acids Res.* **21**, 3913–3914.

Ishida, N., Oyunsuren, T., Mashima, S., Mukoyama, H. and Saitou, N. (1995) Mitochondrial DNA sequences of various species of the genus *Equus* with special reference to the phylogenetic relationship between Przewalski's wild horse and domestic horse. *J. Mol. Evol.* **41**, 180–188.

Janczewski, D.N., Yukhi, N., Gilbert, D.A., Jefferson, G.T. and O'Brien, S.J. (1992) Molecular phylogenetic inference from saber-toothed cat fossils of Rancho La Brea. *Proc. Natl. Acad. Sci. USA* **89**, 9769–9773.

Joger, U. and Koch, U. (1994) *Mammuts aus Sibirien.* Hessisches Landesmuseum, Darmstadt.

Kahlke, R.-D. (1995) Die Entstehungs-, Entwicklungs- und Verbreitungsgeschichte des oberpleistozänen *Mammuthus-Coelodonta*-Faunenkomplexes in Eurasien (Großsäuger). *Abh. Senck. Naturforsch. Ges.* **546**, 1–164.

Krings, M., Stone, A., Schmitz, R.W., Krainitzi, H., Stoneking, M. and Pääbo, S. (1997) Neandertal DNA sequences and the origin of modern humans. *Cell* **90**, 19–30.

Levine, M.A. (1990) Dereivka and the problem of horse domestication. *Antiquity* **64**, 727–740.

Levine, M.A. (1996) Domestication of the horse. In, *The Oxford Companion to Archaeology* (B.M. Fagan, Ed), Oxford University Press, Oxford, pp. 315–317.

Loftus, R.T., McHugh, D.E., Bradley, D.G., Sharp, P.M. and Cunningham, E.P. (1994) Mitochondrial diversity and the origins of African and European cattle. *Proc. Natl. Acad. Sci. USA* **91**, 2757–2761.

Marcus, L.F. and Berger, R. (1984) The significance of radiocarbon dates for Rancho La Brea. In, *Quaternary Extinctions: A Prehistoric Revolution* (P.S. Martin and R.G. Klein, Eds.), pp. 159–183, University of Arizona Press, Tucson.

Marklund, S., Chaudhary, R., Marklund, L., Sandberg, K. and Andersson, L. (1995) Extensive mtDNA diversity in horses revealed by PCR-SSCP analysis. *Anim. Genet.* **26**, 193–196.

Milinkovitch, M.C., Ortig, G. and Meyer, A. (1993) Revised phylogeny of whales suggested by mitochondrial ribosomal DNA sequences. *Nature* **361**, 346–348.

Nobis, G. (1971) *Vom Wildpferd zum Hauspferd.* Bohlau, Köln.

Nobis, G. (1974) The origin, domestication and early history of domestic horses. *Vet. Med. Rev.* **3**, 211–225.

Pääbo, S. (1986) Molecular genetic investigation of ancient human remains. *C. S. H. Symp. Quant. Biol.* **L1**, 441–446.

Pääbo, S., Gifford, A. and Wilson, A.C. (1988) Mitochondrial DNA sequences from a 7000-year old brain. *Nucl. Acids Res.* **16**, 9775–9787.

Ryder, O.A., Starkes, R.S., Starkes, M.C. and Clegg, J.B. (1979) Hemoglobin polymorphism in *Equus przewalski* and *Equus caballus* analysed by isoelectric focusing. *Comp. Biochem. Physiol.* **62B**, 305–308.

Sher, A.V. (1997) An early Quaternary bison population from Untermassfeld: *Bison menneri* sp. nov. *Monographien des Römisch-Germanischen Zentralmuseums Mainz* **40**, 101–180.

Sher, A.V., Kaplina, T.N., Giterman, R.E., Lozhkin, A.V., Arkhangelov, A.A., Kiselyov, S.V., Kouznetsov, Y.V.,

Virina, E.I. and Zazhigin, V.S. (1979) *Scientfic Excursion on Problem "Late Cenozoic of the Kolyma Lowland" Tour XI: The Guide*. XIV Pacific Science Congress. Academy of Sciences of the USSR, Moscow.

Stoneking, M., Sherry, S.T., Redd, A.J. and Vigilant, L. (1992) New approaches to dating suggest a recent age for the human mtDNA ancestor. *Phil. Trans. R. Soc. Lond. ser. B* **337**, 167–175.

Swofford, D.L. (1993) *Phylogenetic Analysis Using Parsimony (PAUP), Version 3.1.1*. Illinois Natural History Survey, Champaign, Illinois.

Swofford, D.L. (1998) *PAUP*, Beta Test Version 64*. Sinauer Associates, Sunderland, Massachusetts.

Taylor, P.G. (1996) Reproducibility of ancient DNA sequences from extinct Pleistocene fauna. *Mol. Biol. Evol.* **13**, 283–285.

Vereshchagin, N.K. and Lazarev, P.A. (1977) Description of the parts of the carcass and skeletal remains of the Selerikan horse. *Trudy Zool. Inst. USSR* **63**, 85–185.

Vila, C., Savolainen, P., Maldonado, J.E., Amorim, I.R., Rice, J.E., Honeycutt, R.L., Crandall, K.A., Lundeberg, J. and Wayne, R.K. (1997) Multiple and ancient origins of the domestic dog. *Science* **276**, 1687–1689.

Volf, J. (1994) The studbook. In, *Przewalski's Horse – The History and Biology of an Endangered Species* (L. Boyd and K.A. Houpt, Eds.), pp. 61–73, SUNY Press, New York.

Xu, X. and Arnason, U. (1994) The complete mitochondrial DNA sequence of the horse, *Equus caballus*: extensive heteroplasmy of the control region. *Gene* **148**, 357–362.