Animal Genetics, 2000, **31**, 1–7

A pedigree-based study of mitochondrial D-loop DNA sequence variation among Arabian horses

A T Bowling, A Del Valle, M Bowling

Summary

Through DNA sequence comparisons of a mitochondrial D-loop hypervariable region, we investigated matrilineal diversity for Arabian horses in the United States. Sixty-two horses were tested. From published pedigrees they traced in the maternal line to 34 mares acquired primarily in the mid to late 19th century from nomadic Bedouin tribes. Compared with the reference sequence (GenBank X79547), these samples showed 27 haplotypes with altogether 31 base substitution sites within 397 bp of sequence. Based on examination of pedigrees from a random sampling of 200 horses in current studbooks of the Arabian Horse Registry of America, we estimated that this study defined the expected mtDNA haplotypes for at least 89% of Arabian horses registered in the US. The reliability of the studbook recorded maternal lineages of Arabian pedigrees was demonstrated by haplotype concordance among multiple samplings in 14 lines. Single base differences observed within two maternal lines were interpreted as representing alternative fixations of past heteroplasmy. The study also demonstrated the utility of mtDNA sequence studies to resolve historical maternity questions without access to biological material from the horses whose relationship was in question, provided that representatives of the relevant female lines were available for comparison. The data call into question the traditional assumption that Arabian horses of the same strain necessarily share a common maternal ancestry.

Keywords: Arabian horse, horse, horse genetics, maternal lineage, mitochondrial DNA, mtDNA

Introduction

Identification of DNA sequence polymorphism in the mitochondrial genome has unique applications to genetic studies of domestic animals not provided by nuclear genes. The mitochondrial genome is maternally inherited, haploid, and its genes do not recombine. Cells have hundreds to thousands of mitochondria, clonal descendants from the maternally derived set present in the egg cytoplasm at conception. The D-loop hypervariable region of mitochondrial DNA (mtDNA) is of particular interest because, unlike the protein-coding gene regions, it has a high level of sequence variation (Aquadro & Greenberg 1983). It has a moderate mutation rate, estimated for humans of one site every 6000 years (Stoneking *et al*. 1992). The D-loop sequence variation combined with lack of recombination produces a highly informative tool for matrilineal (dam line) relationship studies within a species (Vigilant *et al*. 1989).

We studied D-loop mtDNA variation among Arabian horses, a breed with well-known and wide spread influence on breeds throughout the world. Our study was in the context of the studbook of the Arabian Horse Registry of America (AHRA), founded in 1908. Although over 550 000 horses have been registered, living AHRA horses trace only to about 100 dam lines (potential mtDNA sources), primarily founded by mares exported from the Middle East in the mid to late 19th century (RJ Cadranell II, personal communication). The relationships among the source mares are generally not known; however, traditional tribal horse breeders categorized horses based on their dam line or 'strain' (Bökönyi 1984).

Nearly every Arabian horse exported from the Middle East and subsequently recorded in an Arabian horse studbook was accompanied by a strain designation. Arabian strain nomenclature is a binomial system, composed of strain and substrain designators. Occasionally no substrain is indicated. The designators are assigned based solely on the strain designation of its mother. In general it has been assumed in the West that horses sharing a strain and substrain designation are likely to be derived historically from a common maternal ancestor, and that different substrains of a particular main strain (e.g. Seglawi Jedran and Seglawi al Abd) reflect a female-line connection more distant in time. Arabian horse strain theory (e.g. Raswan 1930) describes breed evolution in terms of 'related' strains, i.e. over time a substrain may become

A T Bowling
A Del Valle
M Bowling
Veterinary Genetics Laboratory, School of Veterinary Medicine, University of California, Davis CA, USA

Correspondence: Ann T Bowling.

Accepted 15 May 1999

sufficiently numerous and widespread to acquire its own main strain designation. Since both strain designations and mtDNA are maternally inherited, through mtDNA studies it is possible to compare oral history and biological models of the breed's history and structure.

We used the D-loop hypervariable region of mtDNA to investigate matrilineal genetic diversity of Arabian horses and the plausibility of matrilineal relationship among horses with the same strain designations. Other aims of this study included providing a diversity database that subsequently may be useful for questions of Arabian identification, parentage and forensics and for exploring questions of evolution, phylogeny, domestication and breed origins.

Materials and methods

Samples

Sixty-two Arabian horses were tested. They were not randomly selected, but rather were chosen by pedigree to provide information about breed diversity contributions of female founder lines and about haplotype uniformity within lines. The horses traced by pedigree in the direct maternal line to 34 mares, primarily acquired from nomadic Bedouin tribes of Middle East in the mid-19th to the mid-20th century (Table 1). Among these 34 matrilineal founders, 31 were designated to one of eight traditional Bedouin strains and three had no recorded strain (Table 1).

The DNA source was peripheral blood or hair roots. Total DNA was extracted from blood using standard protocols. For obtaining DNA from hair, four hair root bulbs (2–3 mm from terminal portion of pulled mane hairs) were placed in 50 μl of 200 mM NaOH and heated at 97 °C for 10 min. The hair root reactions were neutralized using 50 μl of 200 mM HCl and 100 mM Tris–HCl, pH 8·5, with thorough mixing.

DNA Analysis

PCR primers for the hypervariable region of the D-loop between $tRNA^{Pro}$ and the large central conserved sequence block were designed based on horse sequence (Ishida *et al.* 1994a; Xu & Árnason 1994). The primer sequences amplified a 397-bp fragment between sites 15427 and 15735 (GenBank X79547) (forward primer 5′AACGTTTCCTCCCAAGGACT3′, reverse primer 5′GTAGTTGGGAGGGTTGCTGA3′). PCR reactions were prepared as follows: 100–730 ng total genomic DNA or 4 μl of hair lysate, 10 mM Tris–HCl pH 8·3, 50 mM KCl, 2·5 mM $MgCl_2$, 0·2 μM each primer, 2 μM each dNTP, 2 U Perkin Elmer *Ampli-Taq* DNA polymerase (Perkin Elmer Corp., Norwalk, CT) and sterile dH_2O to a total volume of 100 μl. Thermocycling was perfomed as follows: 94 °C 'hot start' for 5 min, followed by 30 cycles each consisting of 1 min at 94 °C, 45 s at 60 °C and 1 min at 74 °C, final extension at 72 °C for 30 min and 4 °C hold.

DNA sequencing was done on both strands from three to five PCR-generated templates with a Perkin-Elmer ABI 377 DNA Sequencer using the PCR primers as sequencing primers and dideoxy terminator chemistry. For some horses PCR products from the hypervariable regions were also cloned with a TA cloning kit (InVitrogen Corp., Carlsbad, CA) using protocols recommended by the manufacturer. Sequencing templates were generated from PCR products of five to eight TA clones from each animal using M13 primers. These PCR products were subsequently used in sequencing reactions as described above using the M13 primers as sequencing primers.

The GenBank accession numbers for the 27 sequences are AF132568–AF132594.

Analysis of variation

DNA sequences were compared with a reference sequence from 'a Swedish horse' (GenBank X79547). Frequencies of founder dam lines among current US Arabians were obtained using published pedigrees. The contemporary pedigrees were chosen using a random number generator to select 200 horses among 5000 horses with AHRA registration numbers between 495001 and 500000 (primarily horses born in 1993). Mitochondrial DNA haplotype frequencies for US Arabians were then approximated based on the proportions of the dam lines in the random studbook pedigree sample. That two individuals by chance have the same haplotype (probability of identity, PI) was calculated as the sum of the product of the frequency of the haplotypes ($PI = \Sigma a_k^2$ where a is the frequency of the kth genotype). A dendrogram (not shown) based on haplotype data was generated using NEIGHBOR in the PHYLIP package program (Felsenstein 1993). Ass mtDNA sequence (Xu *et al.* 1996) was used as an outgroup equid taxon.

Results

To confirm the reliability of our protocol for determining mtDNA haplotypes from hair or

Table 1. Source mares for mtDNA diversity study of Arabian horses registered with the AHRA, in order by matrilineal frequency from random pedigree sampling

Matrilineal line founder	Freq[1]	Country of importation, date[2]	Strain (substrain)[2]	N (g)[3]	mtDNA haplotype[4]
Rodania	0·105	To England 1881	Kehilan (Rodan)	4 (7,9,9,10)	A01
Ghazieh	0·090	To Egypt, *c.* 1855	Seglawi (Jedran)	5 (9, 10,10,13,13)	A11
Nejdme	0·075	To US 1893	Kehilan (Ajuz)	2 (8,10)	A20
Urfah	0·065	To US 1906	Seglawi (al Abd)	5 (7,7,7,8,8)	A06(4)/A01(1)
Wadduda[5]	0·060	To US 1906	Seglawi (al Abd)	4 (5,6,8,9)	A01
Werdi	0·050	To US 1906	Kehilan (Krush)	2 (6,7)	A22
Basilisk[5]	0·045	To England 1879	Seglawi (Jedran)	3 (9,10,10)	A03
Mlecha	0·040	To Poland 1845	Kehilan (Dajani)	1 (11)	A13
Dajania	0·040	To England 1878	Kehilan (Dajani)	4 (9,9,10,12)	A17(2)/A28(2)
Ferida	0·035	To England 1891	Maneghi (Sbaili)	2 (6,11)	A10
Zulima	0·035	To Spain 1905	Seglawi (al Abd)	1 (6)	A21
Gazella	0·035	To Poland 1845	Kehilan (Ajuz)	3 (9,11,14)	A19
El Dahma	0·030	Foaled in Egypt, *c.* 1880	Dahman (Shahwan)	1 (8)	A12
Selma	0·030	To Egypt, *c.* 1855	Hamdani (Simri)	1 (8)	A05
Abeyah[5]	0·025	To US 1906	Abeyan (Sherak)	2 (8,8)	A02
Reshan	0·020	To US 1906	Kehilan (Heyfi)	1 (9)	A03
Haidee	0·015	To England 1874	Maneghi (Hedruj)	1 (11)	A24
Zaalee	0·015	To France 1882	Seglawi (Jedran)	2 (9,9)	A09
Hadba	0·015	To US 1906	Hadban (Enzahi)	1 (5)	A17
Milordka	0·015	Foaled in Poland 1816	(no recorded strain)	1 (11)	A25
Sahara	0·010	To Poland 1845	Kehilan (Moradi)	1 (13)	A04
Queen of Sheba	0·010	To England 1879	Abeyan (Sherak)	1 (6)	A08
Elsissa	0·010	To Poland 1874	Hadban (Enzahi)	1 (6)	A14
Sheha	0·010	To Egypt, before 1924	Kehilan (Nowak)	1 (4)	A07
Balkis	0·005	To France 1880	Kehilan (Ajuz)	1 (11)	A15
Jellabiet Feysul	0·005	To Egypt, *c.* 1846	Kehilan (Jellabi)	2 (12,13)	A23
Thorayyah	nf[6]	To US 1950	Tuwaysan	1 (2)	A27
Dahma	nf	To England 1881	Dahman (om Amr)	1 (11)	A13
Murana I	nf	To Germany 1816	(no recorded strain)	1 (17)	A19
Lisa	nf	To US, before 1912	Seglawi (Jedran)	2 (6,6)	A21
Hamra Johara	nf	To US 1961	(no recorded strain)	1 (2)	A24
Dafina	nf	To England 1927	Kehilan (Krush)	1 (4)	A16
Nuhra	nf	To England 1939	Kehilan (Wadnan)	1 (5)	A26
Noura	nf	To US 1928	Maneghi	1 (7)	A01

[1]Calculated from 200 randomly selected AHRA pedigrees.
[2]Information from studbooks and other published and unpublished sources.
[3]Number of samples tested for this line (N) and generations from matrilineal founder for each sample (g).
[4]If more than one haplotype, number of examples of each is in parentheses.
[5]Sequence previously reported (Bowling *et al.* 1998).
[6]Not found in selected studbook pedigree sample. Haplotype frequency likely to be < 0·005.

blood by direct PCR product sequencing, for seven horses both tissue sources were used and PCR products were also cloned. Consensus sequences obtained from multiple cloned PCR products of an individual horse did not differ from results obtained by direct sequencing of multiple PCR products, nor did hair and blood results differ (data not shown).

The 62 samples representing 34 maternal lines and eight strains provided evidence for 27 haplotypes of the D-loop hypervariable region in Arabian horses, each differing from the reference sequence ('a Swedish horse') by 1–11 sites and from each other by 1–14 sites within the 397 bp of sequence (Table 2). Compared with the reference sequence, these samples showed altogether 31 sites of base substitution (transition mutations). No deletions or insertions were observed. Most substitution sites were shared between or among haplotypes, but five haplotypes (A09, A17, A22, A23, A24) had unique substitution sites. Haplotypes A01, A02 and A03 and samples for Wadduda, Abeyah and Basilisk lines were previously reported (Bowling *et al.* 1998). Additional examples of the A01 and A03 types were provided in this study from other maternal source lines. None of the Arabian haplotypes

Table 2. Arabian horse haplotype sequences of mtDNA D-loop hypervariable region compared to the reference sequence GenBank X79547. Haplotypes are listed according to an order suggested by a neighbor-joining algorithm using PHYLIP. Dots indicate concordance with the X79547 sequence

Haplo-type	Position of nucleotide substitutions relative to GenBank X79577 (15382–15778)																									
	15477	15478	15493	15494	15495	15496	15534	15538	15542	15585	15596	15597	15598	15602	15603	15604	15615	15616	15617	15630	15635	15649	15650	15666	15667	15683	15703	15709	15720	15726	15740
	C	C	T	T	T	A	C	A	C	G	A	A	T	C	T	G	A	A	T	A	C	A	A	G	A	C	T	C	G	G	A
A02	.	.	.	.	C	.	.	.	T	A	.	G	.	T	.	.	G	.	.	.	T	.	G	A	.	.	C	.	A	.	.
A11	.	.	.	.	C	.	.	.	T	.	.	G	.	T	.	.	.	.	.	.	T	.	G	A	.	.	C	.	A	.	.
A26	.	.	.	.	C	.	.	.	T	.	.	.	.	T	.	.	.	.	.	.	T	.	G	A	.	.	C	.	A	.	.
A12	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	A	.	.	.	.	A	.	.
A25	.	.	.	.	C	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	G	A	.	.	.	.	A	.	.
A10	.	.	.	.	C	.	.	.	.	.	.	G	.	T	.	.	.	.	.	.	.	.	G	.	.	.	.	.	A	.	.
A09	T	.	.	.	C	.	.	.	.	.	.	G	.	T	.	.	.	.	.	.	.	.	G	.	.	.	.	.	A	.	.
A17	.	.	C	.	C	.	.	G	.	.	G	.	.	T	.	.	.	.	.	.	.	.	G	.	.	.	.	T	A	.	.
A28	.	.	C	.	C	.	.	G	.	.	.	.	.	T	.	.	.	.	.	.	.	.	G	.	.	.	.	T	A	.	.
A19	.	.	.	.	C	.	.	G	.	A	.	.	.	T	.	.	.	.	.	.	.	.	G	.	.	.	.	T	A	.	.
A20	.	.	.	.	C	.	.	G	.	A	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	T	A	.	.
A23	.	.	.	.	C	.	.	.	.	.	.	.	.	T	.	.	.	.	.	G	.	.	.	.	.	T	.	.	A	.	.
A22	.	.	.	.	C	.	.	.	.	.	.	.	.	T	.	.	.	.	C	.	.	.	.	.	.	.	.	.	A	.	.
A21	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A14	.	T	.	C	C	G	T	.	.	A	.	.	.	.	C	.	.	.	.	.	.	G	.	.	.	.	.	.	A	.	.
A03	.	.	.	C	C	G	T	.	.	A	.	.	.	.	C	.	.	.	.	.	.	G	.	.	.	.	.	.	A	.	.
A13	.	.	.	C	C	G	T	.	.	A	.	.	.	T	C	.	.	.	.	.	.	G	.	.	.	.	.	.	A	.	.
A16	.	.	.	C	C	G	T	.	.	A	.	.	.	T	C	A	.	.	.	.	.	G	.	.	.	.	.	.	A	.	.
A15	.	.	.	C	C	G	T	.	.	.	.	.	.	T	C	.	.	.	.	.	.	G	.	.	.	.	.	.	A	.	.
A24	.	.	.	C	.	.	.	.	.	.	.	.	C	T	.	.	G	G	.	.	.	.	.	.	.	.	C	.	A	.	.
A08	.	T	.	C	.	.	.	.	.	.	.	G	.	T	.	.	.	.	.	.	.	.	.	.	.	.	C	.	A	.	.
A07	.	.	.	C	.	.	.	.	.	.	.	G	.	T	.	.	.	.	.	.	.	.	.	.	G	.	C	.	A	.	.
A01	.	.	.	C	.	.	.	.	.	.	.	G	.	T	.	A	.	.	.	.	.	.	.	.	G	.	C	.	A	.	.
A06	.	.	.	C	.	.	T	.	.	.	.	G	.	T	.	A	.	.	.	.	.	.	.	.	G	.	C	.	A	.	.
A05	.	.	.	C	.	.	.	.	.	.	.	G	.	T	.	A	.	.	.	.	.	.	.	.	.	.	C	.	A	A	G
A27	.	.	.	C	.	.	.	.	.	A	.	.	.	T	.	A	.	.	.	.	.	.	.	.	.	.	C	.	A	A	G
A04	.	.	.	C	.	.	.	.	.	A	.	.	.	T	.	A	.	.	.	.	.	.	.	.	.	.	C	.	A	.	G

Table 3. Chart of female line connections of seven Arabian horses tested to resolve maternity of Bint Yamama. Tested horses are coded at right side of table (D-loop hypervariable sequence haplotype assignment in parentheses)

				Tested horse	
Line	Branch	Daughter	Grand-daughter	Generations from branch	Name (haplotype)
? (Ghazieh or Jellabiet Feysul)	Bint Yamama	Negma	Roda	7	MM (A11)
			Mahroussa	6	AY (A11)
Ghazieh	Bint Helwa	Hamasa	Hazna	7	UL (A11)
		Ghazala	Gulnare	10	AM (A11)
			Jamila	10	SYJ (A11)
Jellabiet Feysul	Makbula	Kibla	Kabila	8; 7	ET (A23); OM (A23)

matched those provided for three Thoroughbreds, for two Przewalski's horses or for single examples of Mongolian and Japanese native horses (Ishida *et al.* 1994a; Ishida *et al.* 1995; Oakenfull & Ryder 1998). Haplotypes are reported in a tabular array suggested by 100 iterations of a neighbor-joining algorithm (Table 2), but no dendrogram is provided because major subdivisions were not supported by significant bootstrap values over 50%.

Horses in the random AHRA studbook sampling of 200 pedigrees traced to 41 maternal lines (1–21 horses per line, data not shown) including 26 of the 34 lines of this study. Estimated population frequencies for the 26 lines, based on the random pedigree sampling, are shown in column two of Table 1. All 16 of the lines encountered in the random pedigree sampling at a frequency $\geqslant 2\%$ are represented in this study. In addition, our data include examples of less frequent and rare lines. Ten of the lines of this study occurred at a population frequency $< 2\%$ and eight were not represented in the random pedigree set (by implication existing at a frequency $< 0{\cdot}5\%$). Summing population frequency estimates for the 34 maternal lines of this study, we defined the expected mtDNA haplotypes for at least 89% of Arabian horses registered in the US. The remaining potential variation not accounted for here is defined by the ≈ 66 infrequent maternal lines not specifically addressed. While additional Arabian horse haplotype variants are likely to be found among those lines, some haplotypes are likely to match those already described.

Nearly three-fourths (20/27) of the haplotypes were unique to a single source mare. Seven haplotypes (A01, A03, A13, A17, A19, A21, A24) were shared by two or three source mares. Matching mtDNA haplotypes were observed for mare lines with both the same and different strain designations. Twelve haplotype pairs differed by only one base. Again, among the pairs differing only by a single base were examples of both the same and different strains.

For 14 lines, samples were tested from two or more animals, altogether 42 horses. The haplotypes within 12 of these lines were identical despite separation loops by pedigree from common matrilineal ancestors of up to 80 years and 18 generations. For the Dajania line, both A17 and A28 haplotypes were found. These haplotypes differ at one site, position 15596 (two horses with A17 and two with A28). The sequence difference observed by direct sequencing of PCR products was supported by cloning studies (five to 12 clones per horse), using samples of both blood and hair as a source of mtDNA. For the Urfah line both A01 and A06 were found. These haplotypes differ at one site, position 15534. In neither the Dajania horses nor the Urfah horses was evidence of heteroplasmy observed.

The power of derived mitochondrial haplotypes to answer maternity questions for horses from which no biological material is directly available was demonstrated (Table 3). The maternal line pedigree for Bint Yamama, foaled about 1893, traces either to Ghazieh (A11; Blunt 1986) or to Jellabiet Feysul (A23; Dickinson 1937). These lines differ by eight bases within the 397 bp of this study and are clearly distinguishable. The haplotype of Bint Yamama was derived using samples from two lines of her descent, both with identical haplotypes. In comparison with the Ghazieh line and the Jellabiet Feysul line, the Bint Yamama pedigree can be excluded as sharing a matrilineal source with the Jellabiet Feysul line. The derived haplotype for Bint Yamama is A11, like that of the Ghazieh line, providing substantial support for that pedigree connection.

Discussion

Our data for D-loop mtDNA sequence of Arabian horses confirmed the reliability of the studbook recorded female lines of Arabian pedigrees, since haplotypes are concordant across long-separated branches of descent. The stability of the sequence through multiple maternal generations allows reliable derivation of matrilineal haplotypes within the timeframe of studbook records. Although the necessity for sequencing precludes D-loop analysis from being a routine procedure for identification or parentage testing, mtDNA has an advantage over tests of nuclear genes in that it can be applied to historical questions of maternity in the absence of biological material from putative dams or even from offspring, provided that representatives of the relevant female lines are available for comparison.

Our data provided evidence for considerable mitochondrial diversity within Arabian horses. Two previously published reports for D-loop hypervariable region polymorphisms of horses (neither one including Arabians) provide conflicting conclusions about diversity. Ishida *et al*. (1994b) found only three variants in a sample of 40 horses from different breeds, while Marklund *et al*. (1995) concluded that the D-loop region provided evidence of extensive polymorphism, finding 15 variants among 78 horses of five breeds. These reports based their conclusions not on sequence information, but on the single strand conformational polymorphism (SSCP) technique.

For the 397 bp D-loop sequence defined in this study, Arabian horses recorded with the AHRA have one major haplotype (A01; Table 1). An animal of this haplotype would match at least 16 of 100 randomly chosen Arabian horses (sum of frequencies of Rodania, Wadduda and Noura dam lines found in random sampling of studbook pedigrees). However, between any randomly chosen pair of AHRA animals, the chance of a match for D-loop sequences (probability of identity, PI) would be about 0·047, as calculated from the haplotype frequencies provided by the random pedigree study.

In practice, PI would be slightly less effective than calculated. Based on the two examples provided in our study, 'identity' conclusions should include the possibility of single base differences between individuals within the same maternal line. These single-base differences could be the result of incorrect pedigrees. However, the more likely explanation is that they present resolution of heteroplasmy (multiple mitochondrial genotypes within an individual) into its alternative haplotypes, as previously reported for other species, including cattle and humans (Hauswirth & Laipis 1982; Ashley *et al*. 1989; Gill *et al*. 1994). In our study, neither direct sequencing of PCR products, nor cloning analysis, nor use of a second tissue source provided evidence for heteroplasmy as an immediate explanation for the within-line differences, but none of these alternatives would necessarily have been sensitive enough to detect variants at low frequency within the clonal mitochondrial collections of the tested horses.

The Arabian is thought to be among the oldest of horse breeds, and by all traditional accounts there is minimal evidence of crossbreeding with recognized contemporary breeds. While this study provides evidence of matrilineal diversity for Arabians (27 mtDNA haplotypes among 34 sources), in the haplotype collection are at least ten pairs that differ by only a single base in addition to the within-line haplotype pairs previously described. These breed data may be evidence of accumulated mutational diversity arising from alternative fixations of resolved heteroplasmy among descendants of common founders.

Some interpretations of Bedouin breeding traditions have highlighted the subdivision of the desert-bred Arabians into strains, based on lines of matrilineal descent. Strain characteristics and relationships have been described in detail (e.g. Raswan 1930). Constructing genetic relationships from mtDNA haplotypes using the neighbor-joining algorithm provided no evidence that the Arabian breed has major subdivisions or that named strains, as recorded for maternal lines in current pedigrees, represent monophyletic matrilineal groups. Although identity or near identity of mtDNA sequence within the collection of source mares did suggest close maternal relationships for some pairs, these identities were not necessarily between representatives of the same strain or what have been recorded to be related strains. Likewise, some lines descending from mares recorded to have the same strain proved to have different mtDNA haplotypes.

Acknowledgements

The authors thank the owners who provided samples from their horses for this study and RJ Cadranell II of *Arabian Visions* magazine for providing scholarly assistance with historical documentation and registry records. The authors gratefully acknowledge a gift from the Arabian Horse Registry of America to the first author in partial support of this research.

References

Aquadro C.F. & Greenberg B.D. (1983) Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. *Genetics* **103,** 287–312.

Ashley M.V., Lapis P.J. & Hauswirth W.W. (1989) Rapid segregation of heteroplasmic bovine mitochondria. *Nucleic Acid Research* **17,** 7325–31.

Blunt Lady A. (1986) *Journals and Correspondence: 1878–1917* (ed. by R. Archer & J. Fleming). Alexander Heriot, Cheltenham, England, pp. 333–352.

Bökönyi S. (1984) Horse. In: *Evolution of Domesticated Animals* (ed. by I.L. Mason), Longman Group Limited, Essex, England, pp. 162–173.

Bowling A.T., Del Valle A. & Bowling M. (1998) Verification of horse maternal lineage based on derived mitochondrial DNA sequence. *Journal of Animal Breeding and Genetics* **115,** 351–6.

Dickinson J.M. (1937) *A Catalog of Travelers Rest Arabian Horses*. Privately published, Nashville.

Felsenstein J. (1993) *PHYLIP (Phylogeny Inference Package), Version 3.5c.* Distributed by the author. Department of Genetics, University of Washington, Seattle.

Gill P., Ivanov P.L., Kimpton C. *et al.* (1994) Identification of the remains of the Romanov family by DNA analysis. *Nature Genetics* **6,** 130–5.

Hauswirth W.W. & Laipis P.J. (1982) Mitochondrial DNA polymorphism in a maternal lineage of Holstein cows. *Proceedings of the National Academy of Sciences, USA* **79,** 4686–90.

Ishida N., Hasegawa T., Takeda K. *et al.* (1994a) Polymorphic sequence in the D-loop region of equine mitochondrial DNA. *Animal Genetics* **25,** 215–21.

Ishida N., Hirano T. & Mukoyama H. (1994b) Detection of aberrant alleles in the D-loop region of equine mitochondrial DNA by single strand conformation polymorphism (SSCP) analysis. *Animal Genetics* **25,** 287.

Ishida N., Oyunsuren T., Mashima S., Mukoyama H. & Saitou N. (1995) Mitochondrial DNA sequences of various species of the genus Equus with special reference to the phylogenetic relationship between Przewalskii's wild horse and domestic horse. *Journal of Molecular Evolution* **41,** 190–8.

Marklund S., Chaudhary R., Marklund L., Sandberg K. & Andersson L. (1995) Extensive mtDNA diversity in horses revealed by PCR-SSCP analysis. *Animal Genetics* **26,** 193–6.

Oakenfull E.A. & Ryder O.A. (1998) Mitochondrial control region and 12S rRNA variation in Przewalski's horse (*Equus przewalskii*). *Animal Genetics* **29,** 456–459.

Raswan C.R. (1930) *Der Araber und sein Pferd*. Von Schickhardt & Ebner (Konrad Wittwer), Stuttgart.

Stoneking M., Sherry S.T., Redd A.J. & Vigilant L. (1992) New approaches to dating suggest a recent age for the human mtDNA ancestor. *Philosophical Transactions of the Royal Society of London, B* **337,** 167–75.

Vigilant L., Pennington R., Harpending H., Kocher T.D. & Wilson A.C. (1989) Mitochondrial DNA sequences in single hairs from a southern African population. *Proceedings of the National Academy of Sciences, USA* **86,** 9350–4.

Xu X. & Árnason Ú. (1994) The complete mitochondrial DNA sequence of the horse, Equus caballus: extensive heteroplasmy of the control region. *Gene* **148,** 357–62.

Xu X., Gulberg A. & Árnason Ú. (1996) The complete mitochondrial DNA (mtDNA) of the donkey and mtDNA comparisons among four closely related mammalian species-pairs. *Journal of Molecular Evolution* **43,** 438–46.