Original article

History of Lipizzan horse maternal lines as revealed by mtDNA analysis

Tatjana KAVAR[a], Gottfried BREM[b],
Franc HABE[a], Johann SÖLKNER[c], Peter DOVČ[a*]

[a] Department of Animal Science, Biotechnical Faculty, University of Ljubljana,
Groblje 3, 1230 Domžale, Slovenia
[b] Ludwig Boltzmann-Institut für immuno-, zyto- & molekulargenetische Forschung,
Veterinärmedizinishe Universität Wien, Veterinärplatz 1, 1210 Wien, Austria
[c] Department of Livestock Science, University of Agricultural Sciences,
Gregor Mendel Strasse 33, 1180 Wien, Austria

Abstract – Sequencing of the mtDNA control region (385 or 695 bp) of 212 Lipizzans from eight studs revealed 37 haplotypes. Distribution of haplotypes among studs was biased, including many private haplotypes but only one haplotype was present in all the studs. According to historical data, numerous Lipizzan maternal lines originating from founder mares of different breeds have been established during the breed's history, so the broad genetic base of the Lipizzan maternal lines was expected. A comparison of Lipizzan sequences with 136 sequences of domestic- and wild-horses from GenBank showed a clustering of Lipizzan haplotypes in the majority of haplotype subgroups present in other domestic horses. We assume that haplotypes identical to haplotypes of early domesticated horses can be found in several Lipizzan maternal lines as well as in other breeds. Therefore, domestic horses could arise either from a single large population or from several populations provided there were strong migrations during the early phase after domestication. A comparison of Lipizzan haplotypes with 56 maternal lines (according to the pedigrees) showed a disagreement of biological parentage with pedigree data for at least 11% of the Lipizzans. A distribution of haplotype-frequencies was unequal (0.2%–26%), mainly due to pedigree errors and haplotype sharing among founder mares.

Lipizzan horse / control region / origin / phylogeny / pedigree

1. INTRODUCTION

The Lipizzan horse breed was established in the Habsburg court stud at Lipica in 1580 [12] . This baroque horse breed is considered as the oldest European cultural horse breed. In the 17th century, Lipizzans began to spread out from Lipica to the wide area of central and eastern Europe. Numerous

---

* Correspondence and reprints
E-mail: peter.dovc@bfro.uni-lj.si

maternal lines that follow very strict breeding rules have been developed in addition to six classical stallion lines, but some of them have died out during the history of the breed. Founder mares of existing Lipizzan maternal lines were born in the 18th, 19th and 20th century and they originated from many different breeds, including Karst-, Spanish-, Italian-, Kladruber- and Arabian horses [12]. Only maternal lines established on the stud of Lipica before the 2nd World war are recognised as classical maternal lines. All the other maternal lines originate from studs established on the territory of the former Austro-Hungarian Empire [12]. The nowaday Lipizzan studs in this region are considered as traditional studs and they represent important centres for the preservation of the Lipizzan breed.

The history of the Lipizzan maternal lines has been well described [12], however, these descriptions are based mainly on historical- and pedigree data. Additional information elucidating the history of maternal lines could be provided by examining the nucleotide sequences of the mtDNA control region. In our previous study [7] we showed that the sequence variability of the mtDNA control region within the Lipizzan horse breed is sufficient for differentiation of the majority of the Lipizzan maternal lines, and that the control region is a suitable genetic marker for tracing back the history of the Lipizzan maternal lines.

In this study, we sequenced the mtDNA control region of representatives of the Lipizzan maternal lines from eight traditional Lipizzan studs (Lipica, Slovenia; Piber, Austria; Monterotondo, Italy; Szilvasvarad, Hungary; Beclean and Fagaras, Romania; Đakovo, Croatia; Topolcianky, Slovakia). First, we described Lipizzan maternal lines in terms of mtDNA haplotypes, and estimated matrilineal diversity in the Lipizzan horse breed. Second, we examined the relationship among Lipizzan-, domestic- and wild-horse haplotypes, in order to provide new information about the origin of the Lipizzan maternal lines. Finally, we compared the sequence data obtained with those of the Lipizzan maternal line pedigrees and described the genetic structure in eight Lipizzan studs with the aim to prove reliability of pedigrees and to reconstruct the recent history of the Lipizzan maternal lines.

2. MATERIALS AND METHODS

The upstream part of the control region (GenBank X79547 [19], nt 15450–nt 15834) was sequenced for 212 Lipizzans representing 56 maternal lines from eight traditional Lipizzan studs: Lipica, Piber, Monterotondo, Szilvasvarad, Beclean, Fagaras, Đakovo and Topolcianky. We selected 1–6 blood samples per maternal line in order to cover wide portions of the pedigree. In the case of ambiguity (discrepancy of genetic and pedigree data), we sequenced additional samples from maternally related animals or, whenever such samples were not

available, we analysed questionable samples again, to minimise the chance of errors during the processing of samples in the laboratory. To improve the estimation of the genetic relationship among mtDNA haplotypes, we sequenced the downstream part of the control region (GenBank X79547 [19], nt 16351–nt 16660) of one sample per each distinct haplotype, according to the upstream part of the control region.

mtDNA was extracted following a standard procedure [18]. PCR amplification of the upstream part of the control region was performed on an MJ Research PTC100 thermal cycler, with an annealing temperature of 52 °C for 30 s. Reactions (20 μL) contained template DNA (50 ng), 1 × PCR buffer II (Perkin Elmer), 1.5 mM MgCl$_2$, 20 μM dNTPs, 0.4 U *Taq* Polymerase (Perkin Elmer) and 10 pmol of each primer (HDF: 5′-AGTCTCACCATCAACACCCAAAGC-3′ and HF: 5′-CCTGAAGTAGGAACCAGATG-3′). PCR fragments were sequenced using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit and the ABI PRISM 310 DNA Sequencer (PE Applied Biosystems). For the sequence analysis of the downstream part of the control region, amplification of the entire control region was performed using the HDF and HDR primer pair, (HDR: 5′-ACTCATCTAGGCATTTTCAGTG-3′) and an annealing temperature of 49 °C for 60 s.

Sequences of the upstream (385 bp) and downstream (310 bp) part of the control region were aligned using the reference sequence (GenBank X79547, [19]). Kimura 2-parameter distances were calculated [9] and a Neighbour-joining (NJ) tree [15] was drawn with the Phylip program packages [4]. The same program was used to perform bootstrap analysis on 1 000 data sets. For the graphical presentation of the tree, the Treeview program [14] was applied.

The Lipizzan haplotypes obtained were compared to equine control region sequences retrieved from the GenBank database: AF064627-32, AF326635-86 [17], AF072975-96 [10], AF169009-10, AF014405-17 [8], AF056071 [8], D23665-6 [5], D14991 [5], AF055876-9 [13], AF132568-94 [2], AF431965-9, X79547 [19], representing sequences of domestic horses (a variety of breeds from all over the world), sequences of Przewalski's wild horses and sequences of late Pleistocene horses from Alaska. Multiple alignment was performed using Clustal W [16]. Identical sequences were joined into the same haplotype, regardless of the sequence length. The number of sequences joined in each haplotype was accounted for as the haplotype frequency. The NJ tree was constructed as described above. Median networks were constructed using an algorithm for speedy construction by hand [1]. They were generated from the binary data matrix, which comprises zeros at positions where the sequence haplotype in question matches the consensus type and ones where a (usually transitional) variant is present. In the process of construction, compatibilities between characters become manifest in simple branching, whereas

incompatibilities increase dimensionality by doubling parts of the network. Unnecessarily large networks were avoided by reductions of some of the most obvious recurrent mutations (by splitting characters that account for hypothetical multiple hits).

Pedigree data were recorded during stud visits, and pedigrees were reconstructed by tracing-back the maternal line to the individual's founder mare. An animal with an unknown mother was defined as a founder mare. Haplotype frequencies were calculated for a sample of 416 breeding mares, which were present at eight traditional Lipizzan studs at the time of sample collection. For 212 mares, genetic data were obtained by sequencing and for the rest only pedigree data were used.

3. RESULTS

Sequence analysis of 212 Lipizzan horses revealed 37 distinct mtDNA haplotypes. Twenty-four of them are presented in Table I, the remaining 13 haplotypes were described previously [7]. The alignment of 37 distinct Lipizzan sequences with the reference sequence (GenBank X79547, [19]) showed that the majority of polymorphic sites (47) was found in the upstream part of the control region, but only 14 in the downstream part. Many haplotypes with sequence differences in the upstream part have identical sequences in the downstream part. Lipizzan haplotypes differ from each other by 1 to 24 nucleotides or from 0.14 to 3.5%. They are clustered into four groups: C1, C2, C3 and C4 (Fig. 1). The bootstrap values were high for groups C2, C3 and C4 but values for the C1 group were slightly lower. However, the integrity of the C1 group is supported by high sequence identity in the downstream region. Within the C1 group several low bootstrap values were observed, therefore, we did not define subgroups in this group, although at least two subgroups were well supported by relatively high bootstrap values. One of them consisted of the Slavina, Dubovina and X haplotypes and the other one of the G, F and Trompeta haplotypes (Fig. 1). On the contrary, the C2 and C3 groups could be further divided into C2a, C2b and C3a, and C3b subgroups respectively. Such division is well supported by high bootstrap values, and almost identical sequences in the downstream part of the control region.

A comparison of the Lipizzan haplotypes with other equine haplotypes showed that the other haplotypes cluster into the same four groups (C1–C4) including several subgroups (Fig. 2). Lipizzan haplotypes can be found in almost all subgroups with only a few exceptions. For example, they are not present in the C3c group, which consists of five haplotypes determined in wild horses from the late Pleistocene from Alaska [17]. However, these five haplotypes are related to the C3b group of haplotypes, to which the Lipizzan haplotype Thais also belongs.

Table I. Polymorphic sites within the upstream (nt 15450–nt 15834) and downstream (nt 16351–nt 16660) part of the control region for 24 Lipizzan mtDNA haplotypes and the reference sequence (GenBank X79547, [19]). The additional 13 Lipizzan haplotypes were described previously [7]. The DNA sequences of all 37 Lipizzan mtDNA haplotypes are available in GenBank under Acc. No. Y057408-34, AY057435-6, AF168689-98 and AF168699-705.

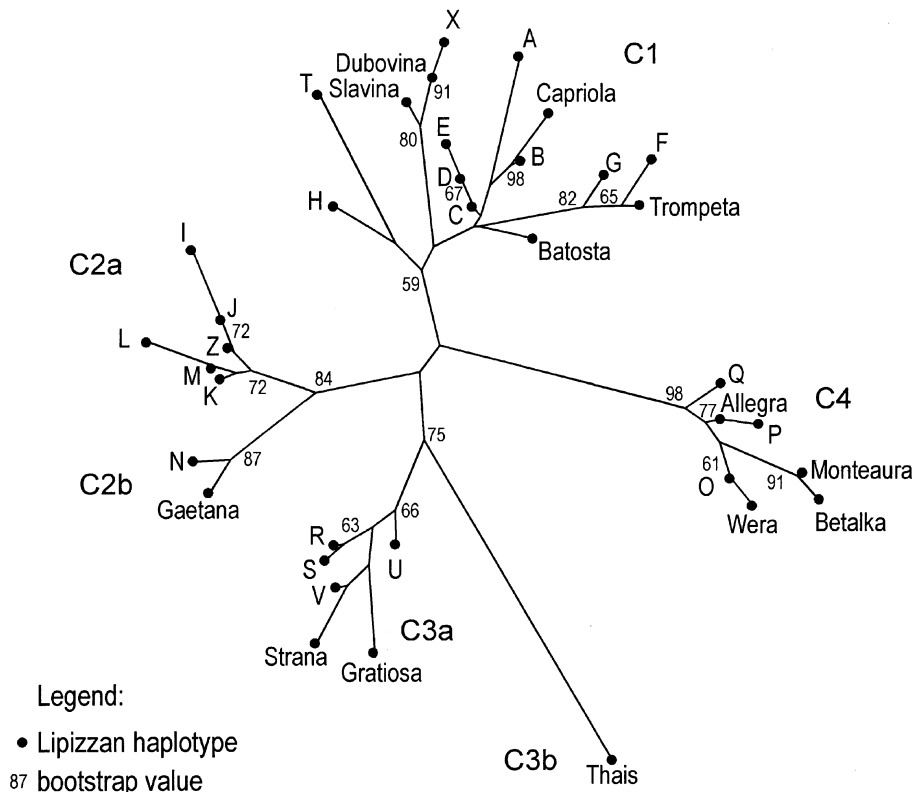| Lipizzan haplotypes	15494	15495	15496	15532	15534	15538	15542	15585	15595	15597	15601	15602	15603	15604	15617	15625	15635	15649	15650	15659	15666	15667	15672	15683	15703	15709	15718	15720	15726	15740	15771	15777	15806	15807	15809	15811	15821	15826	15827	16371	16407	16439	16543	16546	16559	16563A	16629	16644
																																										Downstream part						
X79547	T	T	A	C	C	A	C	G	A	A	T	C	T	G	T	T	C	A	A	T	G	A	G	C	T	C	C	G	G	A	C	A	C	C	A	C	T	A	A	T	C	C	T	T	C	-	A	T
A	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.
B	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	A	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	.
C	.	C	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	.
D	.	C	.	.	.	.	.	.	.	.	.	G	.	T	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	.
X	.	C	.	.	.	T	.	.	G	.	T	.	.	.	.	.	T	.	G	.	A	.	.	.	.	C	.	T	A	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.
E	.	C	.	.	.	.	.	.	.	.	.	G	.	T	.	.	.	C	.	.	G	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	.
F	.	C	.	.	.	G	.	A	.	.	.	.	.	T	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	T	.	A	.	.	T	.	.	G	G	C	.	T	.	.	.	.	.	.
G	.	C	.	.	.	G	.	A	.	G	.	.	.	T	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	T	.	A	.	.	T	.	.	G	G	C	.	.	.	.	.	.	.	.
H	.	C	.	.	.	.	T	A	.	.	.	.	.	T	.	.	.	.	G	.	A	.	.	.	.	.	.	.	.	.	A	.	.	T	.	.	.	.	.	C	.	.	.	.	.	.	.	C
T	.	C	.	-	.	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	T	.	T	.	G	.	C	.	.	.	.	.	.	.	C
I	.	C	.	.	.	.	.	.	G	G	.	T	.	A	.	.	.	.	.	.	.	.	.	.	.	C	.	.	A	.	G	T	G	.	.	T	C	.	.	C	.	A	.	.	.	.	G	.
J	.	C	.	.	.	.	.	.	.	G	.	T	.	A	.	.	.	.	.	.	.	.	.	.	.	C	.	.	A	.	G	T	G	.	.	T	.	.	.	C	.	A	.	.	.	.	G	.
K	.	C	.	.	.	.	.	.	.	.	.	T	.	A	.	.	.	.	.	.	.	.	.	.	.	C	.	A	A	.	T	G	.	.	T	.	.	.	.	C	.	A	.	.	.	.	G	.
Z	.	C	.	.	.	.	.	.	.	.	.	T	.	A	.	.	.	.	.	.	.	.	.	.	.	C	.	A	G	T	G	.	.	T	.	.	.	.	.	C	.	A	.	.	.	.	G	.
L	.	C	.	.	.	.	.	.	.	.	.	T	.	A	.	.	.	.	.	.	A	.	C	.	.	C	.	A	A	G	T	G	.	.	T	.	.	.	.	C	.	A	.	.	.	.	G	.
M	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	C	.	.	C	.	A	A	G	T	G	.	.	T	.	.	.	.	C	.	A	.	.	.	.	G	.
N	.	C	.	.	.	.	.	A	.	.	.	T	.	A	.	.	.	.	.	.	G	.	.	C	.	.	A	.	T	G	.	G	.	.	.	.	.	.	.	C	.	A	.	.	C	.	G	.
S	.	C	.	.	.	.	.	A	.	.	C	T	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	T	.	T	.	.	.	.	.	.	G	.	C	.	A	C	T	.	.	G	.
R	.	C	.	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	T	.	T	.	.	.	.	.	.	G	.	C	.	A	C	T	.	.	G	.
U	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	C	.	.	.	.	A	.	.	T	.	T	.	.	.	.	.	.	.	.	C	.	A	C	T	.	.	G	.
V	.	C	.	.	.	.	.	.	.	.	.	T	.	.	C	.	.	.	.	.	C	.	.	.	.	A	.	.	T	.	T	.	.	.	.	.	.	G	.	C	.	A	C	T	.	.	G	.
P	C	C	G	.	T	.	.	.	.	.	.	.	T	C	.	.	.	G	.	.	.	.	.	.	.	T	.	.	.	.	A	.	.	T	.	.	.	.	.	C	-	.	.	.	.	.	C	G
Q	C	C	G	.	T	.	.	.	.	.	.	.	T	C	.	.	.	G	.	.	.	.	.	.	C	.	A	.	.	T	.	.	.	.	.	.	T	.	.	C	-	.	.	.	.	.	C	G
O	C	C	G	.	T	.	.	A	.	.	.	.	.	C	.	.	.	G	.	.	.	.	.	.	.	A	.	.	T	.	.	.	.	.	.	.	.	.	.	C	-	.	.	.	.	.	C	G

Figure 1. Relationship among 37 Lipizzan mtDNA haplotypes presented by the NJ tree. Bootstrap values higher than 59 were entered into the tree. Haplotypes were clustered into four main groups (C1–C4) and into the C2a, C2b, C3a and C3b subgroups.

Lipizzan haplotypes are in general similar or even identical to other domestic horse haplotypes (Fig. 3). From the networks we can see that some haplotypes are more frequent than the others. They are common for several horse breeds and usually do not have unique polymorphic sites. Such haplotypes, for example, the Gaetana, Monteaura, Allegra and O haplotypes, are present in the Lipizzan breed as well as in the three to four other horse breeds. In the networks, more frequent haplotypes are usually surrounded by haplotypes which differ from them by one to two nucleotides (Fig. 3). Therefore, for each of these haplotypes one or two unique polymorphic sites are characteristic.

The distribution of the mtDNA haplotypes among Lipizzan studs is biased (Tab. II). Only the Batosta haplotype was present in all the studs but many haplotypes were observed only in one or two studs. The haplotype frequency distribution in the Lipizzan breed was unequal, reaching from 26% for the most
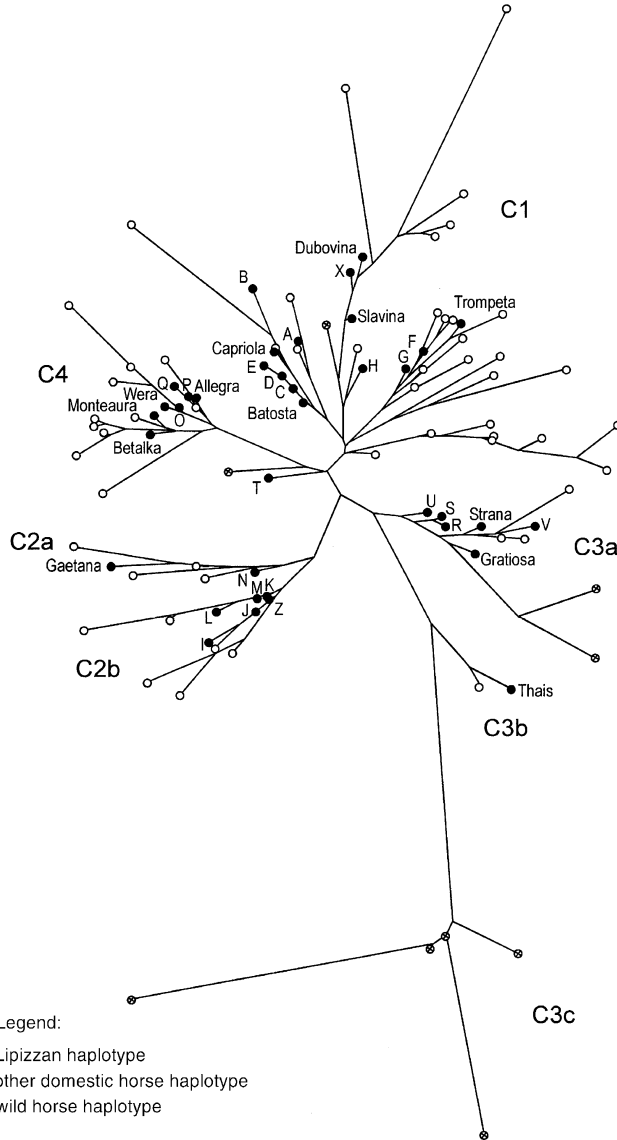
Figure 2. Relationship among Lipizzan haplotypes; some other domestic horse haplotypes and wild horse haplotypes are represented by the NJ tree. The names of the Lipizzan haplotypes are shown. Acc. No. of other sequences included in this tree are: AF014406-8, AF014411-6, AF056071, AF064627-30, AF064632, AF072976-7, AF072979, AF072981, AF072983-4, AF072986-7, AF072989-90, AF072992-3, AF072996, AF169009-10, AF326637, AF326639, AF326641, AF326643-45, AF326647-49, AF326653, AF326655, AF326657-61, AF326663, AF326666, AF326668-72, AF326674-5, AF326678-9, D14991, D23665, D23666, AF055877-8, X79547, AF055878-9 and AF431969.
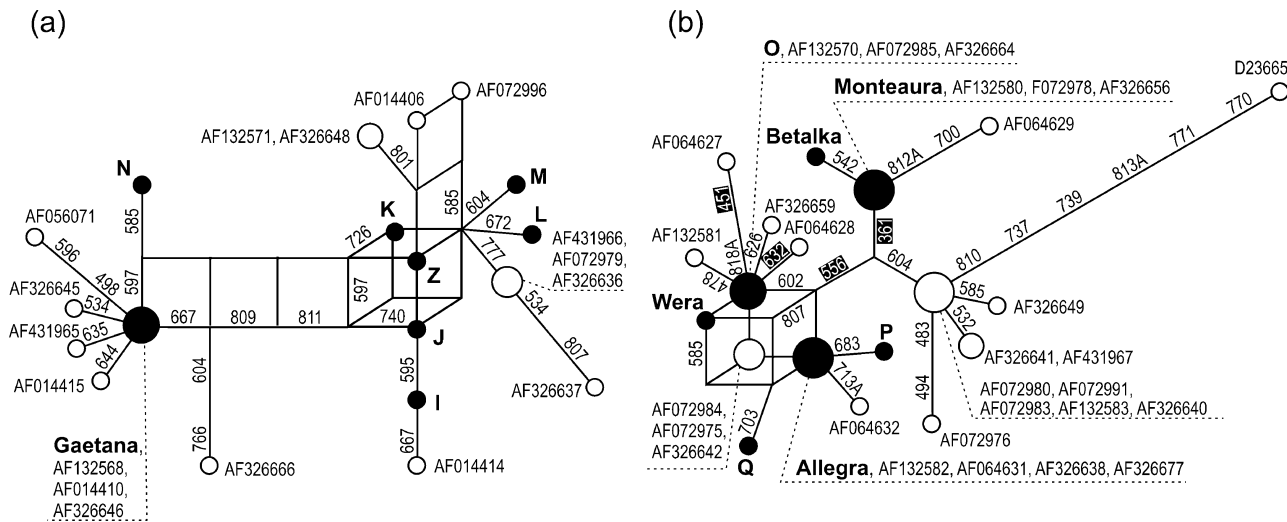
Figure 3. Median network of the C2 main group haplotypes (Fig. 3a) and median network of the C4 main group haplotypes (Fig. 3b). Data matrix for the median network of the C2 main group haplotypes is based on sequence data of the upstream part of the control region, while a data matrix for the network of the C4 main group is based on the sequence data from both parts of the control region. Haplotypes are represented by circles, proportional to haplotype frequencies. Circles representing haplotypes found in the Lipizzan horse breed are black. The numbers on the branches denote nucleotide positions where mutations have occurred (positions > 16 000 are in black boxes). Reticulations within the network indicate ambiguity in the topology and parallel lines in a single reticulation represent the same mutation.

Table II. Distribution and frequencies of 37 Lipizzan mtDNA haplotypes among eight Lipizzan studs. Data include 416 breeding mares present in eight Lipizzan studs at the time of sample collection. For 212 breeding mares, haplotypes were obtained by sequencing whereas the remaining 204 haplotypes were derived from pedigree.

Lipizzan stud	37 Lipizzan mtDNA haplotypes																																				
	Batosta	Capriola	Slavina	Betalka	Allegra	Dubovina	Gaetana	Gratiosa	Monteaura	Wera	P	M	G	B	J	U	X	O	I	Q	R	V	Strana	Thais	Trompeta	L	H	D	E	A	K	S	T	Z	C	F	N
Piber	4	18	2	11	3	6	3	1		3	1	1	2	2	2	9		3	1																		
Monterotondo	3	6	5	1	6	5		1	5					2	5	2	5																				
Đakovo	3		5		1	4	5		1																	2	3	2	13								
Lipica	4	9	5	4	5	2	2	2	2	3													2	2	2												
Szilvasvarad	12	20			8		1					5							18	6	2	1															
Beclean	1	5									4		3																	1	4	1	1				
Fagaras	3	34									3	8	15																	3	5	1	1	1	1	4	5
Topolcianky	11	17	1				3			6																											
Total	41	109	18	16	23	17	14	4	8	12	8	14	20	4	7	11	5	3	19	6	2	1	2	2	2	2	3	2	13	4	9	2	2	1	1	4	5
Frequency	0.099	0.262	0.043	0.038	0.055	0.041	0.034	0.010	0.019	0.029	0.019	0.034	0.048	0.010	0.017	0.026	0.012	0.007	0.046	0.014	0.005	0.002	0.005	0.005	0.005	0.005	0.007	0.005	0.031	0.010	0.022	0.005	0.005	0.002	0.002	0.010	0.012

Table III. Distribution of Lipizzan mtDNA haplotypes among 56 Lipizzan maternal lines (according to the pedigree data). Lipizzan maternal lines are labelled from 1 to 56.

ML[a]	Hap[b]	ML	Hap	ML	Hap	ML	Hap
1	Allegra, Capriola, Slavina	15	Capriola, I	29	Capriola	43	H
2	Allegra, Capriola, V	16	Capriola, X	30	Gaetana	44	I
3	Allegra, U,Wera,	17	Dubovina, L	31	Gaetana	45	J
4	Batosta, Dubovina, Monteaura	18	G, Z	32	Gratiosa	46	K
5	Betalka, Capriola, Monteaura	19	Allegra	33	Dubovina	47	M
6	Dubovina, L, Slavina	20	Allegra	34	Trompeta	48	N
7	Slavina, Gaetana, Strana	21	Batosta	35	Thais	49	M
8	Allegra, Capriola	22	Batosta	36	Wera	50	O
9	Allegra, Capriola	23	Batosta	37	A	51	P
10	Allegra, I	24	Betalka	38	A	52	Q
11	Allegra Slavina,	25	Capriola	39	B	53	R
12	Batosta, Monteaura	26	Capriola	40	D	54	S
13	C, N	27	Capriola	41	E	55	T
14	Capriola, F	28	Capriola	42	G	56	U

[a] Lipizzan maternal line.
[b] Lipizzan mtDNA haplotype.

common haplotype, Capriola, present in 13 lines (Tabs. II and III) to 0.2% for the V, C and Z haplotypes (Tab. II).

According to the pedigree data, Lipizzans from eight traditional studs belong to 56 maternal lines. In 38 lines only one haplotype was found, showing no discrepancy between genetic and pedigree data (Tab. III). However, in 11 lines two haplotypes were identified and in 7 lines three haplotypes were identified. Different haplotypes found within the same maternal line indicate pedigree errors. Therefore, at least 25 pedigree errors have occurred within the Lipizzan breed in the past. Because of the vertical transmission of errors through generations, we estimate that the biological origin of at least 11% of the Lipizzans is in disagreement with their pedigree data.

4. DISCUSSION

This study confirmed our preliminary assumption that the information collected by sequencing the upstream part of the mtDNA control region (nt 15450–nt 15834) is sufficient for characterisation of the mtDNA haplotypes. However, due to the high level of homoplasy observed in the upstream part, additional

sequence information from the less variable downstream part (nt 16351–
nt 16660) of the control region contributed considerably to the estimation of
the genetic relationship among haplotypes, reflected in higher bootstrap values
in the phylogenetic tree (Fig. 1) and in a slightly different relationship among
groups and subgroups (Fig. 1 and Fig. 2) than in NJ trees based only on the
upstream part of the control region [10, 17].

The presence of 37 mtDNA haplotypes reflects the broad genetic base of
Lipizzan maternal lines. According to historical data, numerous maternal lines
from founder mares of different breeds were established during the Lipizzan
breed history, therefore, a high number of mtDNA haplotypes in Lipizzans
met our expectations. The number of haplotypes identified in the Lipizzans
exceeds even the number of haplotypes found in Arabians (29)(AF132568-94,
AF064627/29). High matrilineal diversity was also observed in the Lipizzan
breed as in other domestic horse breeds [2, 10, 17] and in 1000 to 2000 years
old Viking Age horse bones found in a restricted area in northern Europe [17].
The reasons for this are an ancient origin of horse maternal lineages [10, 17],
incorporation of numerous matrilines into the gene pool of domestic horses
during the process of domestication (multiple origins scenario) [17], and little
congruence of haplotype with breed [10, 17].

Only the haplotypes (or their derivatives) which survived two recent popula-
tion bottlenecks (one was the mass extinction of large-bodied animals around
8 000 B.C., and the second one was horse domestication which took place
around 4 000 B.C. [3, 11]) can be found in present day domestic horses.
According to the control region mutation rate ($2–4 \times 10^{-8}$ per site per year) [6],
we assume that haplotypes identical to the haplotypes of early domesticated
horses, so called "ancestral" haplotypes, can still be found in several Lipizzan
maternal lines and in other breeds. We suggest that more frequent haplotypes,
for example, Gaetana, Allegra, Monteaura and O (Fig. 3), which are character-
ised by the lack of unique polymorphic sites, have not changed at least since
horse domestication. Haplotypes which can be found around those "ancestral"
haplotypes in the network probably represent their evolutionary derivatives.
However, the boundary among "ancestral" haplotypes and their "derivatives" is
often unclear, because of the chance that horses with closely related haplotypes,
differing only by one or two polymorphic sites, have been domesticated, too
(Fig. 3b).

Clutton-Brock's model of domestication [3] suggests that the wild stock from
which all domestic animals were bred, inhabited the plains of southern Russia,
from the Ukraine to the region of Turkestan, and the earliest domesticated
horses spread out from this arc and all the different types and breeds of horse
that are known today were developed as a result of artificial selection. However,
an alternative model was also suggested by the same author [3]: this is based
on the assumption that there was a geographical cline in the populations of

wild horses with those in the northern part of the range being smaller and more sturdy than those in the south [3]. No obvious phylogeographic structure in domestic horses has been observed [10] and "ancestral" haplotypes can be found in many different horse breeds, therefore, domestic horses could arise either from a single large population or from several populations provided strong migrations during the early phase after domestication. The first evidence that the wild horse populations might have been phylogeographicaly structured before domestication has already been introduced [17] (all mtDNA haplotypes determined in the late Pleistocene horses belong to the C3 group), but probably this single evidence is not sufficient to confirm or completely reject either one of the horse domestication models [3].

A comparison of mtDNA sequence data with the pedigrees of Lipizzan maternal lines has revealed only 37 different mtDNA haplotypes in 56 maternal lines. This is the consequence of haplotype sharing among maternal lines (Tab. III). Some of the maternal lines with identical haplotypes possibly belong to the same maternal line which is, due to the incomplete pedigree data, split into different maternal lines. The joining of such lines into the same maternal line could be supported by additional pedigree data which would close the gaps in the pedigree.

For the present genetic structure in eight Lipizzan studs, biased distribution of haplotypes (Tab. II), with many private haplotypes and only one haplotype found in all studs, is characteristic, thus suggesting that the independent status of studs has been well preserved up to now. It seems that the exchange of mares among studs was quite restricted. Consequently, private haplotypes remained mainly in the country of the founder mare origin, although some exceptions have been observed.

A comparison of genetic data with pedigree data showed some inconsistencies in the distribution of 37 mtDNA haplotypes within 56 maternal lines (Tab. III). A relatively short time of Lipizzan maternal lines existence (not more than 400 years) in combination with a high degree of nucleotide sequence differences (at least at two sites) among haplotypes within each maternal line (according to the pedigree) suggests that all inconsistencies could be explained solely by pedigree errors. In this context, heteroplasmy and subsequent purifying selection/drift, appear to be a very unlikely alternative explanation for the presence of more than one haplotype within the same maternal line. Therefore, these results showed that pedigrees of Lipizzan maternal lines were not completely reliable. We identified 25 pedigree errors, but it is likely that some additional errors remained hidden, due to the restriction that pedigree errors cannot be found in the uni-lineal pedigrees. Unfortunately, such parts of the pedigree occurred in the majority of 56 Lipizzan matrilineal pedigrees. Therefore, biological founder mares could be different from founder mares identified by pedigree data. On the contrary, it is possible that some of the

Lipizzan maternal lines, which were already assumed to be extinct according to the pedigree data (*e.g.* Rozca, Khel il Massaid, Mersucha), biologically survived as a consequence of pedigree errors. Pedigree errors and the fact that some maternal lines shared the same haplotype, had an impact on the unequal distribution of haplotype frequencies in the Lipizzan population (Tab. II). Several haplotypes were found only in one or two breeding mares. These rare haplotypes could disappear very soon, if the discrepancy among genetic and pedigree data is not taken into consideration.

In spite of the fact that genetic and pedigree data are in some cases in disagreement, we do not recommend radical revision of Lipizzan maternal line pedigrees. In addition, genetic data obtained by the analysis of present horses would not allow reliable corrections of the pedigrees. Therefore, we suggest that the pedigree information about the maternal origin of each breeding mare should be accompanied by the information about its mtDNA haplotype. This would not require typing of all of the animals in order to avoid negative consequences of pedigree errors (*e.g.* extinction of rare maternal lines); occasional testing of some randomly selected animals from each line would be sufficient.

REFERENCES

[1] Bandelt H.-J., Macaulay V., Richards M., Median Networks: Speedy construction and greedy reduction. One simulation and two case studies from human mtDNA, Mol. Phylogenet. Evol. 16 (2000) 8–28.

[2] Bowling A.T., Del Valle A., Bowling M., A pedigree-based study of mitochondrial D-loop DNA sequence variation among Arabian horses, Anim. Genet. 31 (2000) 1–7.

[3] Clutton-Brock J., A natural history of domesticated mammals, 2nd ed., Cambridge University Press, Cambridge, New York, Melbourne, 1999.

[4] Felsenstein J., PHYLIP (Phylogeny Inference Package) version 3.5c., University of Washington, Department of Genetics, Seattle, 1993.

[5] Ishida N., Hasegawa T., Takeda K., Sakagami M., Onishi A., Inumaru S., Komatsu M., Mukoyama H., Polymorphic sequence in the D-loop region of equine mitochondrial DNA, Anim. Genet. 25 (1994) 215–221.

[6] Ishida N., Oyunsuren T., Mashima S., Mukoyama H., Saitou N., 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, J. Mol. Evol. 41 (1995) 180–188.

[7] Kavar T., Habe F., Brem G., Dovč P., Mitochondrial D-loop sequence variation among the 16 maternal lines of the Lipizzan horse breed, Anim. Genet. 30 (1999) 423–430.

[8] Kim K.I., Yang Y.H., Lee S.S., Park C.K., Ma R., Lewin H.A., Phylogenetic relationships of Cheju horses to other horse breeds as determined by mtDNA D-loop sequence polymorphism, Anim. Genet. 30 (1999) 102–108.

[9] Kimura M., A simple model for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences, J. Mol. Evol. 16 (1980) 111–120.

[10] Lister A.M., Kadwell M., Kaagan L.M., Jordan W.C., Richards M.B., Stanley H.E., Ancient and modern DNA in a study of horse domestication, Ancient Biomolecules 2 (1998) 267–280.

[11] MacFadden B.J., Fossil horses: systematics, paleobiology, and evolution of the family *Equidae*, Cambridge University Press, Cambridge, 1992.

[12] Nürnberg H., Der Lipizzaner: mit einem Anhang über den Kladruber, Westarp Wissenschaften, Magdenburg, 1993.

[13] Oakenfull E.A., Ryder O.A., Mitochondrial control region and 12S rRNA variation in Przewalski's horse (*Equus przewalskii*), Anim. Genet. 29 (1998) 456–459.

[14] Page R.D.M., Treeview (Win32), University of Glasgow, Division of Environmental and Evolutionary biology, Institute of Biomedical and Life Sciences, Glasgow, 1998.

[15] Saitou N., Nei M., The Neighbor-joining method: A new method for reconstructing phylogenetic trees, Mol. Biol. Evol. 4 (1987) 406–425.

[16] Thompson J.D., Higgins D.G., Gibson T.J., CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Res. 22 (1994) 4673–4680.

[17] Vila C., Leonard J.A, Gotherstrom A., Marklund S., Sandberg K., Liden K., Wayne R.K., Ellegren H., Widespread origins of domestic horse lineages, Science 291 (2001) 474–477.

[18] White P.S., Densmore L.D., Mitochondrial DNA isolation, in: Hoelzel A.R. (Ed.), Molecular genetic analysis of populations. A practical approach, Oxford University Press, Oxford, 1992, pp. 37–43.

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