Mitochondrial D-loop sequence variation among Hucul horse

V. Czerneková[1], T. Kott[1], I. Majzlík[2]

[1]Institute of Animal Science, Prague-Uhříněves, Czech Republic
[2]Czech University of Life Sciences Prague, Prague, Czech Republic

ABSTRACT: Genetic variation in the Czech Hucul horse population was analyzed using a sequence analysis of the D-loop region of mitochondrial DNA. One hundred and sixty-five Hucul horses were tested. Sequencing of the 700-base pairs fragment of the mitochondrial DNA D-loop region revealed 38 mutation sites representing 14 haplotypes, which were clustered into six haplogroups. The genetic information obtained from the mitochondrial DNA typing is of utmost importance for the future breed-conservation strategies.

Keywords: genetic diversity; horse; mitochondrial DNA; phylogeny

Hucul is a mountain horse, an indigenous breed of the Carpathians. Thanks to solid body and limbs frame, good health and endurance, this breed is remarkable in its ability to maintain high trotting speed on hills and over tight mountain trails (Mason, 1996). The first Hucul stud was established in 1856, however, the world population of this breed comprises only about one thousand horses, and it is included in the protected gene fund of original and primitive animal breeds of FAO. In addition, the Hucul International Federation has been providing uniform methodology and breeding goals since 1994 (http://www.hucul-hif.eu/index_uk.html).

The conservation strategy benefits from analyses of molecular genetics data describing this breed (Cothran et al., 2005; Židek et al., 2009; Priskin et al., 2010). In this work, we concentrate on an application of the mitochondrial DNA (mtDNA) for the description of matrilineal genetics diversity in the Huculs. The identification of the DNA sequence polymorphism in a mitochondrial genome has unique applications to genetic studies of domestic animals. The mitochondrial genome is maternally inherited, haploid, and its genes do not recombine. The D-loop hypervariable region of the mtDNA is of particular interest because, unlike the protein-coding gene regions, it has a high level of sequence variation (Aquadro and Greenberg, 1983). Thus, the D-loop sequence variation combined with the lack of recombinations produce a highly informative tool for matrilineal relationship studies in horses (Bowling et al., 2000; Hill et al., 2002; Kavar et al., 2002; Mirol et al., 2002; Cozzi et al., 2004; Priskin et al., 2010). Here the D-loop hypervariable region of the mtDNA is investigated in as many as 15 maternal lines of the Hucul breed, for the total of 165 horses. Based on this, the matrilineal diversity is estimated and the statistical evaluation of the genetic variability of the studied population is performed. Moreover, the sequence data are compared to the 26 horse sequences from the NCBI GenBank database, and their relation is discussed.

MATERIAL AND METHODS

The mtDNA samples of 165 Hucul horses from the Czech Republic were tested. The horses were traced by pedigree in the direct maternal line to 15 mares

Supported by the Ministry of Agriculture of the Czech Republic (Project No. MZE 0002701404) and by the National Programme of Preservation and Usage of Genetic Resources of Farm Animals.

Table 1. Name, year of birth, country of origin, and number of samples of the Hucul horse line founder

Matrilineal line founder	Born	Origin	No. of samples
Agla	1941	Slovakia	13
Aglalia	1917	Romania	2
Barna	1941	Slovakia	12
Bukovina	1935	Romania	31
Dagmar	1944	Slovakia	27
Gelnica	1947	Slovakia	9
Paskana	unknown	Romania	2
Klapta	1942	unknown	15
Nakoneczna	1934	Poland	3
Polanka	1926	Poland	5
Rumina	1943	unknown	9
Sroczka	1947	Poland	6
Sekacka	1945	Slovakia	14
Valuta (Irma)	1949	Slovakia	16
Zuza	1938	Slovakia	1

(Table 1). The mtDNA source was peripheral blood, collected in the time period 2002–2010. The prime pairs DLF (i.e. 5'- TCTAGCTCCACCATCAAC) and HDR (5'-ACTCATCTAGGCATTTTCAGTG) (Kavar et al., 2002) were used for an amplification of the control region. Importantly, an isolation of mtDNA from blood was avoided by applying the phusion polymerase. Thus, the reaction set-up (20 µl) was as follows: 0.5 µl undiluted blood, 10 µl Phusion™ Flash High-Fidelity PCR Master Mix (Finnzymes, Espoo, Finland), and 0.5 µl of each primer (10 mol/µl). The thermal cycling procedure was carried out on the TGradient 96 Thermocycler (Whatman Biometra, Göttingen, Germany). The protocol included the pre-denaturation at 95°C (5 min), followed by 40 cycles at 95°C (10 s), 62°C (20 s), 72°C (30 s), and the final extension at 72°C for 5 min. PCR fragments were sequenced using ABI PRISM BigDye™ Terminator Cycle Sequencing Ready Reaction Kit and ABI PRISM 3130 Genetic Analyzer (Applied Biosystems, Foster City, USA).

The multiple alignment of DNA sequences was performed using the ClustalW program (Thompson et al., 1994). Population diversity parameters were calculated using DnaSP v5 software (Librado and Rozas, 2009). Genetic distances were calculated and the neighbour-joining tree was drawn with the PHYLIP program package (Felsenstein, 1993), which was also used to perform the bootstrap analysis on 1000 data sets. To estimate the intraspecies sequence divergence, the data were compared against 13 nucleotide sequences of the

Table 2. Diversity indices, mutation sites, and haplogroup assignments of the Hucul samples analyzed in the study

	15494	15495	15496	15521	15526	15528	15534	15538	15542	15585	15597	15602	15603	15604	15617	15635	15649	15650	15659	15666
X79547	T	T	A	G	T	C	C	A	C	G	A	C	T	G	T	C	A	A	T	G
Dagmar	.	C	.	.	.	.	.	.	T	.	G	T	.	.	.	T	.	G	.	A
Gelnica	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	A
Sekacka	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.
Agla	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Aglalia/Bukovina	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Paskana	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Rumina	.	C	.	.	.	.	.	G	.	.	G	T	.	.	.	.	.	G	.	.
Valuta	.	C	.	.	.	.	.	G	.	.	G	T	.	.	.	.	.	G	.	.
Klapta	.	C	.	.	.	del	.	.	.	A	.	T	.	.	.	.	.	.	.	.
Nakoneczna	C	C	G	.	.	.	T	.	.	A	.	.	C	.	.	.	G	.	.	.
Barna	.	C	.	A	C	.	.	.	.	.	.	T	.	.	C	.	.	.	C	.
Polanka	.	C	.	.	.	.	.	.	.	.	G	T	.	A	.	.	.	.	.	.
Sroczka	.	C	.	.	.	.	.	.	.	A	.	T	.	A	.	.	.	.	.	.
Zuza	.	C	.	.	.	.	.	.	.	.	.	T	.	A	.	.	.	G	.	.
Total	165 samples							38 segregating sites												

a = nomenclature by Jansen et al. (2002), b = nomenclature by Achilli et al. (2012)

mtDNA control region from the NCBI GenBank database (http://www.ncbi.nlm.nih.gov/GenBank), with the *Equus asinus* sequence X97337 as the outgroup.

GenBank Accession Nos. for the data presented here are as follows: JF951834–JF951847.

RESULTS AND DISCUSSION

In comparison with the previous studies of Hucul horse (see below), much longer mtDNA fragment was investigated. Thus, the 700-bp fragment of the mtDNA D-loop region (positions 15430–16129) was sequenced. The 165 samples representing 15 maternal lines featured 14 haplotypes of the D-loop hypervariable region (Table 2). The alignment of 14 distinct Hucul sequences with the reference sequence (GenBank X79547) showed altogether 38 sites (37 substitutions). One deletion, also reported in the sequence AF014407 by Kim et al. (1999), was identified in one maternal line. Two substitutions (positions 15874 and 16113) are reported for the first time. It appears that the part of the D-loop region, which includes the position 15874, was studied only by Gurney et al. (2010), however, the substitution was not described. As for the position 16113, it is not contained in the related analyses by Cothran et al. (2005), Priskin et al. (2010), or Georgescu et al. (2011).

Table 2 summarizes the data obtained for the 165 Hucul horses. Namely, 14 haplotypes belonging to 15 families were identified. The Hucul haplotypes thus found differ from the reference sequence (i.e. GenBank X79547) by 2–12 nucleotides, and by 1–11 nucleotides between each other. In all the haplotypes, there is the substitution of T by C in the position 15495. The Aglalia and Bukovina families have only one additional substitution (in the position 15826). The Agla and Paskana families feature 3 substitutions, and are followed by the Sekacka and Gelnica with 4 and 5 substitutions, respectively. The deletion in the position 15528 together with eight substitutions were found for the Klapta family members, while 10 substitutions were described for Rumina (Table 2). The highest number of substitutions exhibited the families Dagmar, Valuta, Nakoneczna, Barna, Polanka, Zuza (11 substitutions), and Sroczka (12). Clearly, there is a fairly high genetic variability (h = 0.895), and low nucleotide diversity (π = 0.01), in the investigated Hucul horses. Overall, the majority of the studied population can be distinguished by assigning one of the haplotypes to an individual.

Importantly, the Agla family shows a new characteristic nucleotide substitution site at position

Table 2 to be continued

	15667	15703	15709	15720	15726	15740	15771	15776	15777	15806	15809	15811	15826	15827	15870	15874	15956	16113	Haplogroup	
																			a	b
X79547	A	T	C	G	G	A	C	T	A	C	A	C	A	A	C	C	A	G	A5	A
Dagmar	.	C	.	A	.	.	.	.	.	.	.	.	.	.	T	.	.	.	A1	B
Gelnica	.	.	.	A	.	.	.	.	.	.	.	.	G	.			.	.	A3	B
Sekacka	.	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.	.	.	A3	B
Agla	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	T	.	.	A4	A
Aglalia/Bukovina	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	A4	A
Paskana	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	A	A4	A
Rumina	.	.	T	A	.	.	T	.	.	.	.	.	G	.	T	.	.	.	B2	I
Valuta	.	.	T	A	.	.	T	.	.	.	.	.	G	.	T	.	.	A	B2	I
Klapta	.	.	.	A	.	.	T	.	.	T	.	.	G	.	T	.	.	.	C	J
Nakoneczna	.	.	.	A	.	.	T	.	.	.	.	.	.	.	T	.	G	.	D2	L
Barna	.	.	.	A	.	.	T	C	.	T	.	.		G	.	.	.	.	E	M
Polanka	G	C	.	A	.	.	T	.	G	.	G	.	.	.	.	.	G	.	F1	P
Sroczka	.	C	.	A	A	G	T	.	G	.	.	T	.	.	.	.	G	.	F2	Q
Zuza	.	C	.	A	.	G	T	.	G	.	.	T	.	.	.	.	G	.	F2	Q
Total	haplotype diversity (h) 0.895 (0.010)												nucleotide diversity (π) 0.0146 (0.0006)							

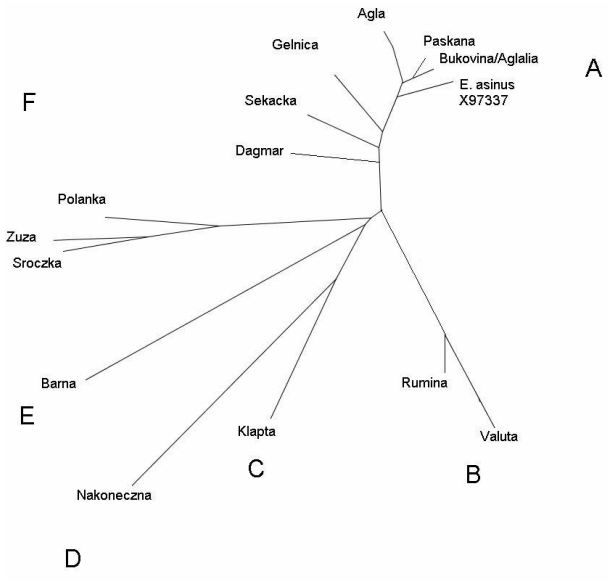

Figure 1. Neighbour-joining tree relating mtDNA haplotypes in Hucul horses. *Equus asinus* sequence (X97337) was chosen as the outgroup

15874, whereas families Paskana and Valuta presented a new substitution at position 16113 (see Table 2). Using the nomenclature from Jansen et al. (2002), haplotypes are clustered into six groups A–F (Figure 1). More than half of the individuals (59.3%) were grouped in the haplogroup A. It consists of representatives of the following families: Agla, Aglaia, Bukovina, Dagmar, Gelnica, Paskana, and Sekacka. Haplogroup B is represented by two maternal families (Rumina and Valuta) and has 15.3% of individuals. Huculs from the family Klapta, Nakoneczna, and Barna were grouped to haplogroups C, D, and E, respectively. Importantly, the haplogroup C has a characteristic deletion at position 15528. The very rare haplogroup is D with 1.8% of individuals. Finally, the haplogroup F (7.3%) joins families Polanka, Sroczka, and Zuza.

It can be of interest to describe the studied population according to the scheme recently proposed by Achilli et al. (2012). Thus, the haplotypes are clustered into eight groups (A, B, I, J, L, M, P, Q). The haplogroup A comprises the members of families Agla, Aglaia, Bukovina, and Paskana, and contains 29.1% of individuals. The group B is the most populated (30.3%) and contains 3 families: Dagmar, Gelnica, and Sekacka. The haplogroup I is represented by Rumina and Valuta maternal lines, and has 15.2% of individuals. Families Klapta, Nakoneczna, Barna, and Polanka belong to the groups denoted as J, L, M, and P, respectively, and contain accordingly 9.1, 1.8, 7.3, and 3.0% of

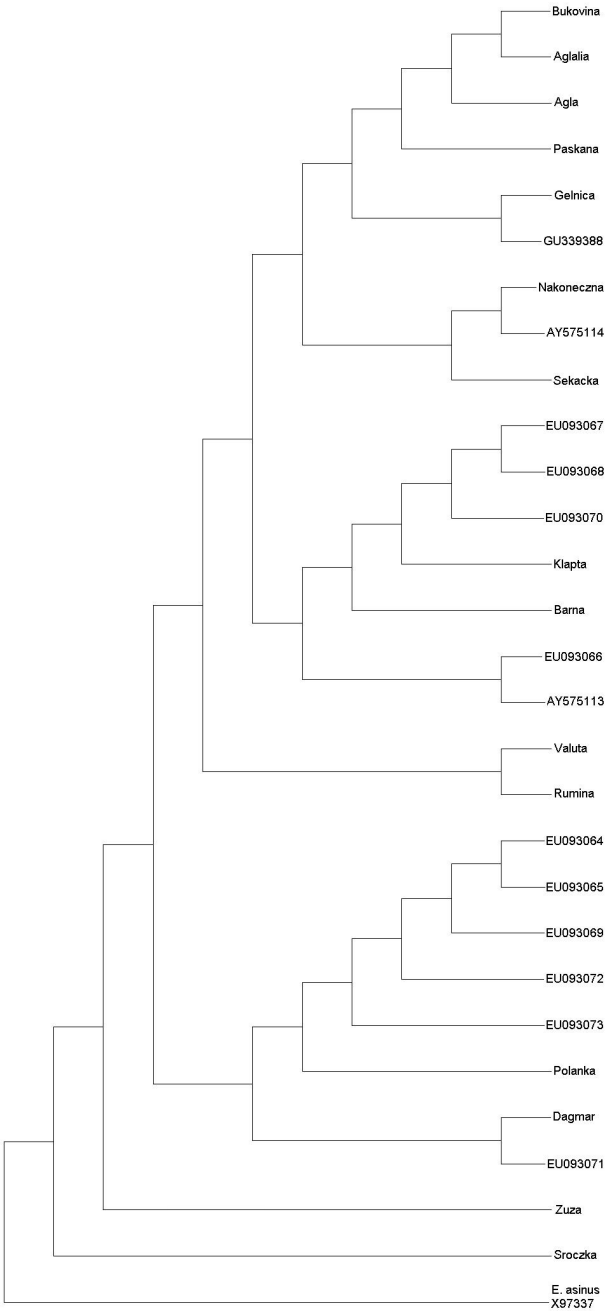

Figure 2. The neighbour-joining tree showing the relationship between partial mtDNA D-loop sequences from this study and from the 13 Hucul sequences available in the GenBank database

the population. Finally, the group Q combines the Sroczka and Zuza families, with 4.2% of the total number of individuals.

Figure 2 compares the present results with 13 sequences, as taken from GenBank database, of the Hucul horse. It is a well-known fact (Bowling et al., 2002; Hill et al., 2002; Kavar et al., 2002; Mirol et al., 2002; Cozzi et al., 2004; McGahern et al.,

2006; Priskin et al., 2010) that the mtDNA D-loop region is very polymorphic. This also holds for the Huculs, as their genetic lines are not isolated (see Cieslak et al., 2010 for the most recent discussion of this topic). However, a relatively low degree of polymorphism was found in previous mtDNA studies of Hucul horses (Cothran et al., 2005; Priskin et al., 2010; Georgescu et al., 2011), but this finding presumably originated from strongly limited numbers of investigated individuals. In particular, based on the analysis of ten samples, Priskin et al. (2010) reported that half of Hucul sequences belonged to the haplogroup F, and the remaining to the haplogroups A and C. Further, Georgescu et al. (2011) found 4 haplotypes in 29 samples from Lucina Stud, Romania, while Cothran et al. (2005) described 2 haplotypes only.

CONCLUSION

As mentioned above, the Hucul horse is close to extinction, but it features some remarkable characteristics, which have made it the object of ongoing preservation efforts. In order to choose optimal preservation strategies, it is crucial to establish the parameters describing the genetic diversity of this breed. These results are obtained here by studying the mitochondrial D-loop DNA polymorphism of the Hucul horses from the Czech Republic, and are likely important for the remaining Hucul populations (in Austria, Hungary, Poland, Romania, and Slovakia) as well.

REFERENCES

Achilli A., Olivieri A., Soares P., Lancioni H., Kashani B.H., Perego U.A., Nergadze S.G., Carossa V., Santagostino M., Capomaccio S., Felicetti M., Al-Achkar W., Penedo M.C.T., Verini-Supplizi A., Houshmand M., Woodward S.R., Semino O., Silvestrelli M., Giulotto E., Pereira L., Bandelt H.J., Torroni A. (2012): Mitochondrial genomes from modern horses reveal the major haplogroups that underwent domestication. Proceedings of the National Academy of Sciences of the United States of America, 109, 2449–2454.

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

Bowling A.T., Del Valle A., Bowling M. (2000): A pedigree-based study of mitochondrial D-loop DNA sequence variation among Arabian horses. Animal Genetics, 31, 1–7.

Cieslak M., Pruvost M., Benecke N., Hofreiter M., Morales A., Reissmann M., Ludwig A. (2010): Origin and history of mitochondrial DNA lineages in domestic horses. PloS One, 5, e15311.

Cothran E.G., Juras R., Macijauskiene V. (2005): Mitochondrial DNA D-loop sequence variation among 5 maternal lines of the Zemaitukai horse breed. Genetics and Molecular Biology, 28, 677–681.

Cozzi M.C., Strillacci M.G., Valianti P., Bighignoli B., Cancedda M., Zanotti M. (2004): Mitochondrial D-loop sequence variation among Italian horse breeds. Genetics Selection Evolution, 36, 663–672.

Felsenstein J. (1993): PHYLIP – Phylogeny Inference Package v3.2. Cladistics, 5, 164–166.

Georgescu S.E., Manea M.A., Dudu A., Costache M. (2011): Phylogenetic relationships of the Hucul horse from Romania inferred from mitochondrial D-loop variation. Genetics and Molecular Research, 10, 4104–4113.

Gurney S.M., Schneider S., Pflugradt R., Barrett E., Forster A.C., Brinkmann B., Jansen T., Forster P. (2010): Developing equine mtDNA profiling for forensic application. International Journal of Legal Medicine, 124, 617–622.

Hill E.W., Bradley D.G., Al-Barodz M., Ertugrul O., Splan R.K., Zakharov I., Cunningham E.P. (2002): History and integrity of thoroughbred dam lines revealed in equine mtDNA variation. Animal Genetics, 33, 287–294.

Jansen T., Forster P., Levine M.A., Oelke H., Hurles M., Renfrew C., Weber J., Olek K. (2002): Mitochondrial DNA and the origins of the domestic horse. Proceedings of the National Academy of Sciences of the United States of America, 99, 10905–10910.

Kavar T., Brem G., Habe F., Solkner J., Dovc P. (2002): History of Lipizzan horse materna lines as revealed by mtDNA analysis. Genetics Selection Evolution, 34, 635–648.

Kim K.I., Yang Y.H., Lee S.S., Park C., Ma R., Bouzat J.L., Lewin H.A. (1999): Phylogenetic relationships of Cheju horses to other horse breeds as determined by mtDNA D-loop sequence polymorphism. Animal Genetics, 30, 102–108.

Librado P., Rozas J. (2009): DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25, 1451–1452.

Mason I.L. (ed.) (1996): A World Dictionary of Livestock Breeds, Types and Varieties. 4th Ed. CAB International, Wallingford, UK.

McGahern A.M., Edwards C.J., Bower M.A., Heffernan A., Park S.D., Brophy P.O., Bradley D.G., MacHugh D.E., Hill E.W. (2006): Mitochondrial DNA sequence diversity in extant Irish horse populations and in ancient horses. Animal Genetics, 37, 498–502.

Mirol P.M., Garcia P.P., Dulout F.N. (2002): Mitochondrial variability in the D-loop of four equine breeds shown by PCR-SSCP analysis. Genetics and Molecular Biology, 25, 25–28.

Priskin K., Szabo K., Tomory G., Bogacsi-Szabo E., Csanyi B., Eordogh R., Downes C.S., Rasko I. (2010): Mitochondrial sequence variation in ancient horses from the Carpathian basin and possible modern relatives. Genetica, 138, 211–218.

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

Židek R., Jakabová D., Trandžík J., Gralak B., Burócziová M., Buleca Jr. J., Maróti-Agóts Á., Massányi P., Dvořák J., Říha J., Zöldág L. (2009): Diversity analysis of Hucul horse population based on molecular genetic data. Hungarian Veterinary Journal, 131, 685–691.

Received: 2013–01–02
Accepted after corrections: 2013–05–23

Corresponding Author

RNDr. Vladimíra Czerneková, Institute of Animal Science, Department of Molecular Genetics, Přátelství 815, 104 00 Prague 10-Uhříněves, Czech Republic
Tel. +420 267 009 743, fax +420 267 710 779, e-mail: czernekova.vladimira@vuzv.cz