MtDNA Haplotype Analysis in Dam Families of the Thoroughbred Riding Horses

Lyudmila A. KHARBROVA 1*, Alexander M. ZAITSEV 1, Larisa L. VIKULOVA 1, Marina V. ADAMKOVSKAYA 1, Nina V. BLOKHINA 1, and Sergey I. SOROKIN 2

1 All-Russian Scientific Research Institute of Horse Breeding, 91105, Divovo, Ryazan Region, Russian Federation
2 Laboratory “Horse Gene”, Moscow, Russian Federation
* Correspondence: l.khrabrova@yandex.ru

Abstract: The high level of diversity of the genetic structure of mitochondrial DNA and maternal inheritance through the maternal line make it a unique object for studying evolutionary processes, phylogenetic relationships and assessing population diversity. 64 thoroughbred horse mares were selected from 20 main dam families and the most promising branches for analysis of the non-coding region of the mtDNA D-loop. Primers for amplification of the mtDNA D-loop region were selected taking into account the reference sequence of the fossil Swedish horse X79547. GeneBank data was used for 17 haplogroups located under the access number (n = 83): JN398377-JN398457, EF597513-EF597514 for identifying the obtained nucleotide sequences. The 530 mtDNA D-loop region analysis was performed using the MEGA4 program. 23 haplotypes corresponding to 8 haplogroups A, B, G, H, I, L, M and N according to the classification of Achilli et al. (2012) were identified in the process of studying the polymorphism of the hypervariable region of the mtDNA D-loop in dam families of thoroughbred horse breed. Among mares from 20 dam families 16 were identified as belonging to only one haplogroup. 2 haplogroups were identified in 4 dam families of this breed (4, 5, 9, 19) due mainly to transitions at different breeding periods of thoroughbred riding horses. The genetic analysis of the mitochondrial genome of the main dam lines of thoroughbred horse breed indicates a rather high level of diversity of mtDNA haplotypes and a clear differentiation of female families in the domestic population of this breed.

Keywords: mtDNA haplogroup, genetic analysis, horse, dam families, thoroughbred horse breed.

1. Introduction

The high level of diversity of the genetic structure of mitochondrial DNA and maternal inheritance make it a unique object for studying evolutionary processes, phylogenetic analysis and assessing population diversity. The study of features of the nuclear and mitochondrial DNA of horses of different breeds and ranges, including the found remains of ancient horses, made it possible to clarify many important issues in the evolution of equids. In particular, a high variability of mitochondrial DNA was revealed, indicating the presence of several wild ancestors in domestic horses and the existence of different
regions of domestication [1, 2]. Thanks to the study of the mitochondrial genome it was finally established that the species Equus saballus is not a direct descendant of the Przewalski wild horse.

A comparative analysis of mitochondrial DNA in different populations and horse breeds of Europe and Asia and the phylogenetic reconstruction made on its basis showed the presence of complex variability of mitochondrial haplogroups [3–6] in horses. This was not observed in other domesticated species. Comparison of European and Asian horse breeds revealed a differentiated distribution of variants of mtDNA haplogroups and evidence of a biogeographic wedge in Asian populations, including the association of “eastern” mtDNA genotypes with haplogroup F. A number of additional mtDNA sequences were found when typing horses of Akhal-Teke, Vyatka, Mezen, Oryol and Yakut breeds. The closest resemblance to the European population was noted in the Mezen horse [4].

In 2012 a research team led by A. Achilli [7] performed a complete sequencing of 83 mitochondrial genomes of modern horses in Europe, Asia, the Middle East and America. The high molecular resolution phylogenetic analysis revealed 18 major haplogroups (A–R) with their diagnostic mutational motifs that arose during the Neolithic period. The researchers concluded that the proposed classification of the coded and control sections of the mitochondrial genome can be used to study the remains of ancient horses, phylogenetic relationships of modern breeds, intra-breed diversity and to evaluate the possible relationship of mtDNA with race performance.

The study of the hypervariable region sequence of the mtDNA D-loop makes it possible to evaluate the intrabreed diversity of horses by female lines and to compare the similarity of dam families by the mitochondrial genome [8–12]. The analysis of polymorphism of the mtDNA D-loop in 100 purebred horse mares belonging to 19 historical female lines made by Irish scientists [10] revealed the presence of 17 haplotypes of mtDNA. 39 polymorphic sites including 3 inserts, 35 transitions and 1 transversion were revealed in the sequenced fragment of the D-loop of the tested mares with a size of 381 bp. Only 11 female families were represented by one haplotype, while the thoroughbreds mares of the remaining 7 families were represented by two (1, 5, 6, 9, 12, 16 and 19) and even three haplotypes.

It is important to note that the appearance of the ancestors of the families of this breed dates from the end of the XVII - the beginning of the XVIII century. 40 or more generations separate them from modern descendants. Breeding horses of this breed in this country has a long tradition. 287 stallions and 366 mares were recorded in 1 volume of the Russian pedigree book of thoroughbred horses which was published in 1836.

In 1894 B. Low proposed to classify the female lines of thoroughbred riding horses. The classification was compiled taking into account the number and winnings of the progeny of the founding mares. In the mid-twentieth century the number of main dam families of thoroughbred horse breed increased from 47 to 174, taking into account the peculiarities of the genealogical structure of populations of different countries [13].

A study of the characteristics of the mitochondrial genome in 251 horses of a thoroughbred Arabian breed from different countries showed [11] that they belong to 13 mtDNA haplogroups according to the classification of Achilli et al. [7]. The largest number of mtDNA haplotypes and haplogroups were detected in a comparative
assessment of the populations of different breeding zones in Arabian horses of the Syrian population that is in the region where this breed was created. Arabian horses had a much wider range of mitochondrial genome haplotypes compared to a purebred horse breed. It should be noted that high HapL frequency was recorded in all populations.

The objective of our research was to study the variability of the mtDNA D-loop site in representatives of different dam families of thoroughbred horse breed in order to assess the level of genetic differentiation of its genealogical structure according to maternal lines and the formation of a database of haplotypes of the mitochondrial genome of this breed.

2. Material and Methods

Sixty four thoroughbred horse mares with a proven origin from the 20 most numerous dam families representing the most promising branches were selected for analysis of the non-coding region of the mtDNA D-loop. DNA was isolated from hair follicles using the kit “ExtraGene DNA Prep 200” manufactured by LLC Laboratory “Isogen” (Moscow), according to the manufacturer’s instructions. Primers for amplification of the studied region of the mtDNA D-loop were selected taking into account the reference sequence of the fossil Swedish horse X79547 [14].

During polymerase chain reaction (PCR), the composition of the reaction mixture included 0.2 mM dNTP, 0.5 μM of each primer, 2.5 mM MgCl2, 1xPCR buffer, 1 unit Taq polymerase (PE Applied Biosystems) and 1 unit AmpliTaqGold Polymerase (PE Applied Biosystems), 50 ng DNA. The reaction mixture was heated at 95 °C for 5 min. The next 30 cycles included 40 s of denaturation at 94 °C, annealing of the primers for 45 s at 55 °C and elongation at 72 °C for 45 s. The final elongation was carried out for 2 minutes at a temperature of 72 °C. Sequencing was performed using the BigDyeTM Terminator Cycle Sequencing Kit (PE Applied Biosystems) on an ABI 3130xl genetic analyzer (PE Applied Biosystems).

To identify the obtained nucleotide sequences, we used GeneBank data for 17 known haplogroups located under the access number (n = 83): JN398377-JN398457, EF597513-EF597514. The analysis of the mtDNA D-loop region from 15471 to 16000 bases was performed using the MEGA4 program.

3. Results

The study of polymorphism of the hypervariable region of the 530 bp mtDNA D-loop in 64 mares of thoroughbred horse breed belonging to the 20 main families showed the presence of 23 haplotypes corresponding to 8 haplogroups: A, B, G, H, I, L, M and N according to the classification of Achilli et al. [7]. At the same time, 16 mares from 20 brood families were identified as belonging to only one haplogroup (Table 1). 2 haplogroups were identified in 4 dam families of this breed (4, 5, 9, 19) due mainly to transitions (single nucleotide substitutions) at different breeding periods of thoroughbred riding horses. The historical reconstruction of the mitochondrial genome of the main female lines of the thoroughbred horse breed conducted by E.W. Hill et al. [10] revealed the presence of relatively recent changes in the structure of the mtDNA D-loop of modern mares belonging to families 1, 12, 16 and 19, as well as significant divergences in the mares genomes of 11th and 12th families in the early stages of their formation.
Table 1. Distribution of thoroughbred horse mares of different families by mtDNA haplogroups (Achilli et al., 2012) [7]

| Family number | Family ancestor                     | Main branches | Number of mares | mtDNA haplogroup |
|---------------|------------------------------------|---------------|-----------------|------------------|
| 1             | Tregonwell’s Natural Barb Mare      | 1-c, 1-d, 1-t | 6               | N                |
| 2             | Barton Barb Mare                   | 2-b, 2-n      | 2               | G                |
| 3             | Byerley Turk Mare                  | 3-h, 3-i, 3-j | 7               | L                |
| 4             | Layton Barb Mare                   | 4-b, 4-c, 4-h, 4-m, 4-o | 7* + 3 | I*, L           |
| 5             | The Massey Mare                    | 5-b, 5-e      | 3               | B, L             |
| 6             | Old Bald Peg                       | 6             | 2               | L                |
| 7             | Blackleges Royal Mare              | 7-b           | 1               | L                |
| 9             | Vintner Mare                       | 9-b, 9-e, 9-f | 4               | G*, M            |
| 10            | Childers Mare                      | 10, 10-a      | 4               | L                |
| 11            | Sudbury Royal Mare                 | 11-d          | 1               | I                |
| 13            | Royal Mare                         | 13-c          | 2               | I                |
| 14            | The Oldfield Mare                  | 14, 14-a, 14-f | 4              | L                |
| 16            | Hutton’s Spot Mare                 | 16-c, 16-g    | 4               | L                |
| 17            | Byerley Turk Mare                  | 17-b, 17-d    | 3               | L                |
| 18            | Old Woodcock Mare                  | 18-a          | 2               | A                |
| 19            | Davill’s Old Woodcock Mare         | 19            | 2               | H, L             |
| 21            | Queen Anne’s Moonah Barb Mare      | 21-a          | 2               | B                |
| 23            | Piping’s Peg’s Dam                 | 23            | 2               | M                |
| 27            | Spanker Mare                       | 27, 27-a      | 2               | L                |
| 43            | Natural Barb Mare                  | 43            | 1               | B                |

Note: * haplogroup of the family ancestor.

The genetic structure of a thoroughbred horse breed by haplogroups of the mtDNA D-loop is generally typical of European populations. Haplogroups such as B, G, I, and L are most widely used [11, 15, 16]. It can be considered logical that the mitochondrial genome of thoroughbred riding horses is represented by only 8 of 18 haplogroups, since the gene pool of this breed was formed by a limited number of horses recorded in Volume I of the General Studentbook, published in England in 1791.
Conducted studies have shown that most of the dam families of the domestic population of thoroughbred horse breed belong to haplogroups I (4, 11, 13) or L (3, 4, 5, 6, 10, 16, 17, 27), which is clearly demonstrated by the dendrogram in Figure 1. The largest family in the world ranking such as 1 Tregonwell’s Natural Barb Mare family was represented by 6 mares of three different branches in our study. All mares represented a rare haplogroup of mtDNA N with a high degree of bootstrap support (96). Two variants of group N haplotypes were identified in mares of this family, while all tested mares of branches 1-c and 1-d had almost the same sequence, slightly different from that in branch 1-t.

The 4th largest family of Layton Barb Mare is the leader in the number of mares of thoroughbred horse breed in this country. It was represented by 10 mares belonging to five different branches (4-b, 4-c, 4-h, 4-m, 4-o). Moreover, the genetic testing showed that all the mares of the first four branches belong to haplogroup I, while the other three mares, representing independent branches of the 4-o branch, belong to haplogroup L with a 98% correspondence level (Figure 2). A genealogical analysis of the development of this branch showed that all three tested mares date back to the mare of Waternymph born in 1860, representing the 20th generation from the ancestor of this family. Obviously, there was a historical divergence of the maternal line in the 4th family, affecting the 4th branch, which includes the modern families of Veltash, Omoriki and Flambonet.

Haplogroup H is rare in thoroughbred riding horses. It was determined only in the representative of the 19th family. Its haplotype is shown in Table 2. Haplogroup G turned out to be typical for mares of the 2nd and 9th families, each of which was characterized by peculiar haplotypes of the mtDNA D-loop. The same haplotype of haplogroup M was tested in two mares of 23 families and one mare from 9th families, in which, as it was established [10], genetic divergence of the mitochondrial genome occurred.

![Figure 1. The distribution of thoroughbred dam mtDNA haplogroups](image-url)
Figure 2. Dendrogram of haplogroups D-loops of mtDNA of thoroughbred horses according to the classification of Achilli et al. (2012) [7]

Table 2. The nucleotide sequence variability in the 362 bp fragment of the mtDNA D-loop in the dam families of thoroughbred horse breed

| Haplogroup | Family number | 35465 | 35478 | 35494 | 35526 | 35534 | 35540 | 35542 | 35545 | 35551 | 35557 | 35558 | 35561 | 35571 |
|------------|---------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| A 18-a     |               | C     | C     | T     | A     | G     | T     | C     | A     | C     | G     | A     | T     | C     |
| B 43       |               |       |       | A     |       | G     | A     |       |       |       |       |       |       |       |
| G 2-n      |               |       |       |       |       |       |       |       |       |       |       |       |       |       |
| H 14-f     |               |       |       |       |       |       |       |       |       |       |       |       |       |       |
| I 2-a      |               |       |       |       |       |       |       |       |       |       |       |       |       |       |
| M 23       |               |       |       |       |       |       |       |       |       |       |       |       |       |       |
| N 1-d      |               |       |       |       |       |       |       |       |       |       |       |       |       |       |
| L1 3-j     |               |       |       |       |       |       |       |       |       |       |       |       |       |       |
| L2 4-o     |               |       |       |       |       |       |       |       |       |       |       |       |       |       |
| L3 17-b    |               |       |       |       |       |       |       |       |       |       |       |       |       |       |
An analysis of the phylogenetic relationships of maternal lines in a thoroughbred breed of horses shows that a significant part of the ancestors of the main families, almost 8 out of 20 (40%), have haplogroup L and possibly common female ancestors in more distant generations. At the same time, 8 different haplotypes typical of modern dam families were recorded within this predominant haplogroup of mtDNA.

4. Conclusions

The genetic analysis of the 530 bp mtDNA D-loop in 20 main dam lines of thoroughbred horse breed indicates a rather high level of haplotype diversity and a clear differentiation of female families in the domestic population of this breed. The data obtained make it possible to supplement important information on the genetic characteristics of the existing genealogical structure of a thoroughbred horse breed.

Acknowledgments

The research was supported by the Russian Science Foundation (Project No. 19-7620058).

References

(1) The Genetics of the Horse.; Bowling, A. T., Ruvinsky, A., Eds.; CABI: Wallingford, 2000. https://doi.org/10.1079/9780851994291.0000.

(2) Jansen, T.; Forster, P.; Levine, M. A.; Oelke, H.; Hurles, M.; Renfrew, C.; Weber, J.; Olek, K. Mitochondrial DNA and the Origins of the Domestic Horse. *Proceedings of the National Academy of Sciences* **2002**, 99 (16), 10905–10910. https://doi.org/10.1073/pnas.152330099.

(3) Cardinali, I.; Lancioni, H.; Giontella, A.; Capodiferro, M. R.; Capomaccio, S.; Buttazzoni, L.; Biggio, G. P.; Cherchi, R.; Albertini, E.; Olivieri, A.; et al. An Overview of Ten Italian Horse Breeds through Mitochondrial DNA. *PLoS ONE* **2016**, 11 (4), e0153004. https://doi.org/10.1371/journal.pone.0153004.

(4) McGahern, A.; Bower, M. A. M.; Edwards, C. J.; Brophy, P. O.; Sulimova, G.; Zakharov, I.; Vizuete-Forster, M.; Levine, M.; Li, S.; MacHugh, D. E.; et al. Evidence for Biogeographic Patterning of Mitochondrial DNA Sequences in Eastern Horse Populations. *Animal Genetics* **2006**, 37 (5), 494–497. https://doi.org/10.1111/j.1365-2052.2006.01495.x.

(5) Moridi, M.; Masoudi, A. A.; Vaez Torshizi, R.; Hill, E. W. Mitochondrial DNA D-Loop Sequence Variation in Maternal Lineages of Iranian Native Horses. *Animal Genetics* **2013**, 44 (2), 209–213. https://doi.org/10.1111/j.1365-2052.2012.02389.x.

(6) Vilstrup, J. T.; Seguin-Orlando, A.; Stiller, M.; Ginolhoc, A.; Raghavan, M.; Nielsen, S. C. A.; Weinstock, J.; Froese, D.; Vasiliev, S. K.; Ovodov, N. D.; et al. Mitochondrial Phylogenomics of Modern and Ancient Equids. *PLoS ONE* **2013**, 8 (2), e55950. https://doi.org/10.1371/journal.pone.0055950.
(7) Achilli, A.; Olivieri, A.; Soares, P.; Lancioni, H.; Kashani, B. H.; Perego, U. A.; Nergadze, S. G.; Carossa, V.; Santagostino, M.; Capomaccio, S.; et al. Mitochondrial Genomes from Modern Horses Reveal the Major Haplogroups That Underwent Domestication. Proceedings of the National Academy of Sciences 2012, 109 (7), 2449–2454. https://doi.org/10.1073/pnas.1111637109.

(8) Cothran, E. G.; Juras, R.; Macijauskiene, V. Mitochondrial DNA D-Loop Sequence Variation among 5 Maternal Lines of the Zemaitukai Horse Breed. Genetics and Molecular Biology 2005, 28 (4), 677–681. https://doi.org/10.1590/S1415-47572005000500006.

(9) Głażewska, I.; Wysocka, A.; Gralak, B.; Prus, R.; Sell, J. A New View on Dam Lines in Polish Arabian Horses Based on MtDNA Analysis. Genetics Selection Evolution 2007, 39 (5), 609–619. https://doi.org/10.1186/1297-9686-39-5-609.

(10) Hill, E. W.; Bradley, D. G.; Al-Barody, M.; Ertugrul, O.; Splan, R. K.; Zakharov, I.; Cunningham, E. P. History and Integrity of Thoroughbred Dam Lines Revealed in Equine MtDNA Variation. Animal Genetics 2002, 33 (4), 287–294. https://doi.org/10.1046/j.1365-2052.2002.00870.x.

(11) Khanshour, A. M.; Cothran, E. Maternal Phylogenetic Relationships and Genetic Variation among Arabian Horse Populations Using Whole Mitochondrial DNA D-Loop Sequencing. BMC Genetics 2013, 14, 83. https://doi.org/10.1186/1471-2156-14-83.

(12) Sorokin, S. I. Molecular and Genetic Analysis of D-Loop Mitochondrial DNA of Blood Families of Vladimir Breed (in Russian). Konevodstvo i Konny Sport (Horse Breeding and Equestrian Sports) 2015, 6, 27–29.

(13) Bochkarev, K. P. Genealogical Tables the Mares Lines Thoroughbred Horses in USSR (in Russian); Agropromizdat: Moscow, 1989.

(14) Xiufeng, X.; Árnason, Ú. The Complete Mitochondrial DNA Sequence of the Horse, Equus Caballus: Extensive Heteroplasmy of the Control Region. Gene 1994, 148 (2), 357–362. https://doi.org/10.1016/0378-1119(94)90713-7.

(15) Morelli, L.; Useli, A.; Sanna, D.; Barbato, M.; Contu, D.; Pala, M.; Cancedda, M.; Francalacci, P. Mitochondrial DNA Lineages of Italian Giara and Sarcidano Horses. Genetics and Molecular Research 2014, 13 (4), 8241–8257. https://doi.org/10.4238/2014.October.20.1.

(16) Pérez-Gutiérrez, L. M.; De la Peña, A.; Arana, P. Genetic Analysis of the Hispano-Breton Heavy Horse. Animal Genetics 2008, 39 (5), 506–514. https://doi.org/10.1111/j.1365-2052.2008.01762.x.
Information about Authors

Lyudmila Aleksandrovna KHRABROVA: D.Sc. in Agriculture, Professor, Chief Researcher of the All-Russian Research Institute of Horse Breeding; 35 apt. 29 (All-Russian Research Institute of Horse Breeding), Divovo (Settlement), Rybnovsky District, Ryazan Region, 391105, Russia; e-mail: l.khrabrova@yandex.ru; ORCID ID: https://orcid.org/0000-0003-2590-8472.

Alexander Mikhailovich ZAITSEV: Ph.D. in Agriculture, Director of the All-Russian Research Institute of Horse Breeding; All-Russian Research Institute of Horse Breeding, Divovo (Settlement), Rybnovsky District, Ryazan Region, 391105, Russia; e-mail: vniik08@mail.ru; ORCID ID: https://orcid.org/0000-0003-4260-602X.

Larisa Leonidovna VIKULOVA: Ph.D. in Agriculture, Senior Researcher, Genetics Laboratory, All-Russian Research Institute of Horse Breeding; All-Russian Research Institute of Horse Breeding, Divovo (Settlement), Rybnovsky District, Ryazan Region, 391105, Russia; e-mail: larvik@list.ru; ORCID ID: https://orcid.org/0000-0001-8773-7117.

Marina Viktorovna ADAMKOVSKAYA: Ph.D. in Biology, Researcher, All-Russian Research Institute of Horse Breeding; All-Russian Research Institute of Horse Breeding, Divovo (Settlement), Rybnovsky District, Ryazan Region, 391105, Russia; e-mail: admare@mail.ru; ORCID ID: https://orcid.org/0000-0003-1537-7225.

Nina Vasilievna BLOKHINA: Ph.D. in Agriculture, Researcher, Genetics Laboratory, All-Russian Research Institute of Horse Breeding; 26 (All-Russian Research Institute of Horse Breeding), Divovo (Settlement), Rybnovsky District, Ryazan Region, 391105, Russia; e-mail: nbloh16@yandex.ru; ORCID ID: https://orcid.org/0000-0001-7406-6385.

Sergey Ivanovich SOROKIN: Ph.D. in Agriculture, Head of the Laboratory “Horse Gene”; 6/1-128 Sumsky Str., Moscow, 117208, Russia; e-mail: sorokinsergey77@gmail.com; ORCID ID: https://orcid.org/0000-0002-0012-413X.