

A new view on dam lines in Polish Arabian horses based on mtDNA analysis

Iwona GŁAŻEWSKA^{a*}, Anna WYSOCKA^a, Barbara GRALAK^b
Renata PRUS^a, Jerzy SELL^a

^a Department of Genetics and Cytology, University of Gdańsk, Kładki 24,
80-822 Gdańsk, Poland

^b Polish Academy of Sciences Institute of Genetics and Animal Breeding, Jastrzębiec,
05-552 Wólka Kosowska, Poland

(Received 5 December 2006; accepted 9 March 2007)

Abstract – Polish Arabian horses are one of the oldest and the most important Arab populations in the world. The Polish Arabian Stud Book and the Genealogical Charts by Skorkowski are the main sources of information on the ancestors of Polish Arabs. Both publications were viewed as credible sources of information until the 1990s when the data regarding one of the dam lines was questioned. The aim of the current study was to check the accuracy of the pedigree data of Polish dam lines using mtDNA analysis. The analyses of a 458 bp mtDNA D-loop fragment from representatives of 15 Polish Arabian dam lines revealed 14 distinct haplotypes. The results were inconsistent with pedigree data in the case of two lines. A detailed analysis of the historical sources was performed to explain these discrepancies. Our study revealed that representatives of different lines shared the same haplotypes. We also noted a genetic identity between some lines founded by Polish mares of unknown origin and lines established by desert-bred mares.

mitochondrial DNA / horse / pedigree / phylogeny

1. INTRODUCTION

Polish Arabian horses are one of the oldest and most important Arab populations in the world. The first stud described in historical sources was established in 1778 [16]. Nineteenth century studs were located in the southeast region of historical Poland, which corresponds now to the Ukrainian territory. The foundation breeding material was of a different origin. The Biała Cerkiew and Sławuta stocks are traced back to local mares, and five dam lines originating from these studs are currently present in Polish breeding (Tab. I).

* Corresponding author: glazew@biotech.ug.gda.pl

Table I. Dam lines in Arabian horses in Poland: founders and main branches of the lines.

Source stud	Founder of the line*	The most important representatives**		N of horses, (generation)
Biała Cerkiew	Szamrajówka ~1810 (5.3)	Piewica 1953 SU, imp. PL (12)		3 (16,17,17)
Sławuta	Szweykowska ~1803 (6.1)	Lezginka 1895 (8)	Kalina 1909 (9)	1 (17)
			Elstera 1913 (9)	1 (18)
	Wołoszka ~1810 (7.7)	Rusałka 1911 (9)	Karima 1924 (10)	1 (15)
			Niespodzianka 1931 (10)	1 (17)
	Ukrainka ~1815 (4.6)	Forta 1943 (12)		3 (14,17,19)
	Milordka ~1816 (15.0)	Delfina 1825 (1)	Anielka 1909 A, imp. PL (8)	2 (16,16)
		Zaira 1826 (1)	Siglavi-Bagdady 1908 A, imp. PL (8)	3 (17,18,19)
Koalicja 1918 A, imp. PL (9)			1 (18)	
Jarczowce	Gazella db imp. 1845 (21.1)	Łania (2)	Gazella II 1914 (6)	4 (13,14,14,17)
		Gazella II (2)	Fryga 1914 (5)	3 (12,14,16)
	Mlecha db imp. 1845 (13.2)	Mlecha (1)	Łysa 1915 (6)	3 (13,13,14)
		Rozmaita I (1)	Pomponia II 1924 (7)	2 (15,15)
	Sahara db imp. 1845 (9.2)	Zulejma 1914 (6)	Dziwa 1922 (7)	1 (15)
			Huryska 1926 (7)	1 (14)
Crabbet Park (Great Britain)	Selma db 1865 imp. 1894 (5.2)	Sardhana 1924 imp. PL (2)	1 (8)	
		Star of the Hills 1927 (2) imp. SU	Potencja 1952 imp. PL (6)	1 (11)
			Tiwiriada 1966 imp. PL (7)	1 (9)
	Rodania db 1869 imp. 1881 (2.9)	Ridaa 1892 (2)	Noma 1953 SU imp. PL (7)	1 (12)
		Risala 1900 (3)	Pienoczka 1952 SU imp. PL (7)	1 (12)
Tiaret (France)	Cherifa db 1869 imp. 1870 (1.3)	Bad 1929 imp. PL (6)		3 (12,13,15)
Babolna (Hungary)	Adjuze db 1876 imp. 1885 (0.15)	Bryła 1939, imp. PL (4)		2 (9,11)
	Bent-El-Arab db 1880 imp. 1885 (0.3)	Balía 1936, imp. PL (5)		1 (10)
	Scherife db 1896 imp. 1902 (3.7)	Bulwa 1937, imp. PL (3)		1 (7)
	Semrie db 1896 imp. 1902 (2.6)	112 Koheilan Raschid 1903 (1)	Baza 1934, imp. PL (5)	1 (10)
		6 Koheilan IV 1912 (1)	Brda 1934, imp. PL (3)	1 (8)

* In brackets: percentage of representatives of the line of brood mares registered in 2003; db: desert-bred mares.

** In brackets: generation number with the founder generation considered as 0; abbreviations: A: Austro-Hungarian Empire, Radautz stud; PL: Poland; SU-USSR: Tersk stud.

The second group of founders consisted of horses imported from the Near East, and included the mares Gazella, Mlecha, and Sahara, the founders of very valuable dam lines.

World War I and the October Revolution resulted in the total destruction of most of the Polish studs [7]. Extensive reconstruction of Polish Arabian breeding was undertaken in the subsequent years, and mares representing the new dam lines of Rodania, Selma, and Cherifa were imported to Polish studs. World War II inflicted another ruinous blow to Polish breeding, during which most of the studs were destroyed, and the horses were either plundered or were missing [13, 14, 16]. Post-war breeding started with 59 surviving brood mares and 14 mares from the Babolna stud [7]. The representatives of four new dam lines were among Babolna's mares (Tab. I), and the Szamrajówka and Rodania lines were revived by importing representatives from the Tersk stud (former USSR).

While Polish Arabian breeding started at the end of the 18th century, the history of the Polish studbook dates back to 1926; only a few reports containing incomplete pedigree data were published beforehand [16]. More or less accurate breeding registers were kept in some studs; unfortunately, they were completely destroyed in 1914–1918. Efforts to create an official studbook were undertaken in the 1920s and the first studbook, edited by Dr. Edward Skorkowski was published in 1926, which was then revised and published in 1932 as the Polish Arabian Stud Book (PASB), Vol. I. In 1938, Skorkowski published the Genealogical Charts of Polish Arabian horses. Due to the scrupulous work performed by the editor in preparing these charts, both publications were considered for years to be credible sources of information on the ancestors of Polish Arabs. It was not until the 1990s that one of the female lines was called into doubt. According to Kwiatkowski [11], one of two branches of the Milordka line, in fact, originates from the mare Malikarda.

Research methods based on mtDNA analysis can permit the verification of Kwiatkowski's hypothesis. The verification of dam lines, the history of which was re-created based on incomplete sources, was the second aim of the current study. The third aim of the work was to evaluate the genetic similarity between the founders of the oldest Polish lines, the ancestors of which are unknown, and the founders purchased in the Near East and admitted as purebred Arabian mares.

2. MATERIAL AND METHODS

The material analysed comprised peripheral blood samples from 44 randomly chosen representatives of 15 dam lines and their most important branches in Poland (Tab. I). The samples came from the collection for routine testing of parentage controls carried out at the Institute of Genetics and Animal Breeding in Jastrzębiec. In 2002, these lines were represented by 98.4% of the total 722 registered brood mares [14]. In addition to the lines analysed in the current study, single representatives of different dam lines imported in recent years are also present in Polish studs.

The samples were stored at -20°C . Total DNA was extracted using the Genomic Mini kit (A&A Biotechnology) according to the manufacturer's guidelines. The mtDNA D-loop hypervariable region was amplified using the following: forward primer 5'-AACGTTTCCTCCCAAGGACT-3' [4] and reverse primer IRD 800 5'-ATGGCCCTGAAGAAAGAACC-3' [9]. These primers amplify a 458 bp fragment between sites 15 400 and 15 857. Amplification of the mitochondrial D-loop region was carried out in a total volume of 20 μL polymerase chain reaction mixture with 1.6 μL dNTP (10 mM), 2 μL 10 X PCR buffer for Taq DNA Polymerase (Invitrogen), 1 μL MgCl_2 (50 mM), 0.6 μL 5 μM of each primer, about 50 ng of template and 0.2 μL Taq DNA Polymerase (5U \cdot μL^{-1} , Invitrogen). PCR conditions were as follows: 5 min at 94°C ; 35 cycles of the following: 1 min at 94°C , 45 s at 58.5°C , 1 min at 72°C ; final extension at 72°C for 10 min. The amplified DNA was electrophoresed in 1% agarose gel containing ethidium bromide to check fragment length. PCR products were sequenced directly in both directions, using the same primers then for amplification. To check the quality of sequencing, BLAST [1] searches were performed. The sequences were aligned using BioEdit [8] and the GenBank X79547 reference sequence [21] was included. Finally, the haplotypes of Polish Arabs were compared to those of Arabian horses in the USA [4] based on the equine mtDNA control region between 15 427 and 15 740 bp.

3. RESULTS

The analyses of a 458 bp mtDNA D-loop fragment in representatives of 15 Polish Arabian dam lines identified 14 distinct haplotypes. They differ from the reference sequence (GenBank X79547) by four to ten sites, and altogether by 32 sites as base substitutions (all transitions; Tab. I). Differences between the haplotypes ranged from one to 14 with an average of 9.07 nucleotides.

Table II. Variable site positions in the mtDNA control region for 14 Polish Arabian horse haplotypes defined by sequence analysis. Dots indicate concordance with the X79547 GenBank sequence; the DNA sequences of all haplotypes are available in GenBank under Acc. No. EF 437553-EF 437566.

Haplotypes and lines/branches	Variable site positions																																	
	15 494	15 495	15 496	15 534	15 538	15 542	15 585	15 597	15 602	15 603	15 604	15 617	15 635	15 649	15 650	15 659	15 666	15 667	15 703	15 709	15 720	15 726	15 740	15 770	15 771	15 777	15 806	15 809	15 810	15 811	15 826	15 827		
X79547	T	T	A	C	A	C	G	A	C	T	G	T	C	A	A	T	G	A	T	C	G	G	A	C	A	C	A	C	A	A	C	A	A	
A Szamrajówka	C	C	G	T	T	C	.	.	.	G	A	.	.	.	T
B Wołoszka	.	C	T	G	A	G	.
C Szwejkowska	.	C	T	.	.	C	.	.	.	C	A	.	.	.	T	.	.	T	G	.
D Miłordka/Zaira	.	C	A	G	.	A	.	.	.	A	G	.	.	G	.
E Miłordko/Delfina	.	C	G	T	.	A	C	.	A	.	G	.	T	G	T	.	.
F Gazella/Gazella II	.	C	.	.	G	.	A	.	T	G	T	A	.	.	.	T	G	.
G Mlecha; Ukrainka	C	C	G	T	.	.	A	.	T	C	.	.	.	G	A	.	.	.	T
H Sahara; Gazella/Fryga	.	C	A	.	T	.	A	C	.	A	.	G	.	T	G	T	.	.
I Selma	.	C	G	T	.	A	C	.	A	A	G	.	T	T	.	.
K Rodania; Cherifa	.	C	G	T	.	A	G	C	.	A	.	.	.	T	G	.	G
L Adjuze	C	C	G	T	.	.	A	.	G	C	.	.	.	G	A	.	.	.	T
M Bent-El-Arab	.	C	.	.	.	T	.	G	T	.	.	.	T	.	G	.	A	.	C	.	A
N Scherife	.	C	A	.	.	T	G	.	.
O Semrie	.	C	.	.	G	.	A	G	T	G	T	A	.	.	.	T	G	.

Eight haplotypes were each represented in a single dam line. All Babolna's dam lines, Szamrajówka, Selma, and two of four of Sławuta's lines (Szweykowska and Wołoszka) were concerned. Two haplotypes were each shared by representatives of two lines. The G haplotype was found in the Ukrainka and Mlecha lines, and the K haplotype was registered in the Rodania and Cherifa lines. More than one haplotype was found in two lines. The D and E haplotypes, differing from each other by 13 nucleotides, were found in the Milordka line. The D haplotype was identified in the branch established by Zaira, a daughter of Milordka. The E haplotype was found in the branch comprised of approximately 28% of the representatives of the line [14] and founded by Milordka's second daughter, Delfina, whose origin has been questioned by Kwiatkowski [11].

Two haplotypes were also found in the Gazella line. The F haplotype was identified in the branch continued by Gazella II (Tab. I). The second haplotype, which is identical to the H haplotype of the Sahara line, was found in the second branch continued by Fryga. This branch is represented by approximately 35% of the mares from the line.

Figure 1 shows the relationships between the haplotypes of Polish and American lines [4]. The analysis yielded four groups of lines with differences not greater than four nucleotides. The largest group includes 16 lines, of which are those founded by Polish mares of unknown origin: Wołoszka, Szweykowska, and Milordka's daughter Zaira. The group of the highest similarity proved to be that of Adjuze, Mlecha, Ukrainka, Szamrajówka and six foreign lines, with differences from one to three nucleotides.

Haplotypes identical to those found in American Arabian horses analysed by Bowling *et al.* were present in eight lines (Fig. 1). Particular attention was given to the genetic identity between the line founded by Szamrajówka, a mare of unknown origin born about 1810, and the line of Balkis, a desert-bred (d.b.) mare imported to France in 1880. The genetic identity between the largest Polish line of Gazella d.b. and the line of Murana I d.b., imported in 1816 to the Weil stud (Württemberg, Germany) is also worth noting.

4. DISCUSSION

The current results were inconsistent with pedigree data recorded in the Genealogical Charts by Skorkowski [18] in two cases. A detailed analysis of historical sources was performed to explain these discrepancies.

In the case of the Milordka line, the results indicate a possible error in the record of the line, which was suggested by Kwiatkowski [11]. Haplotype D

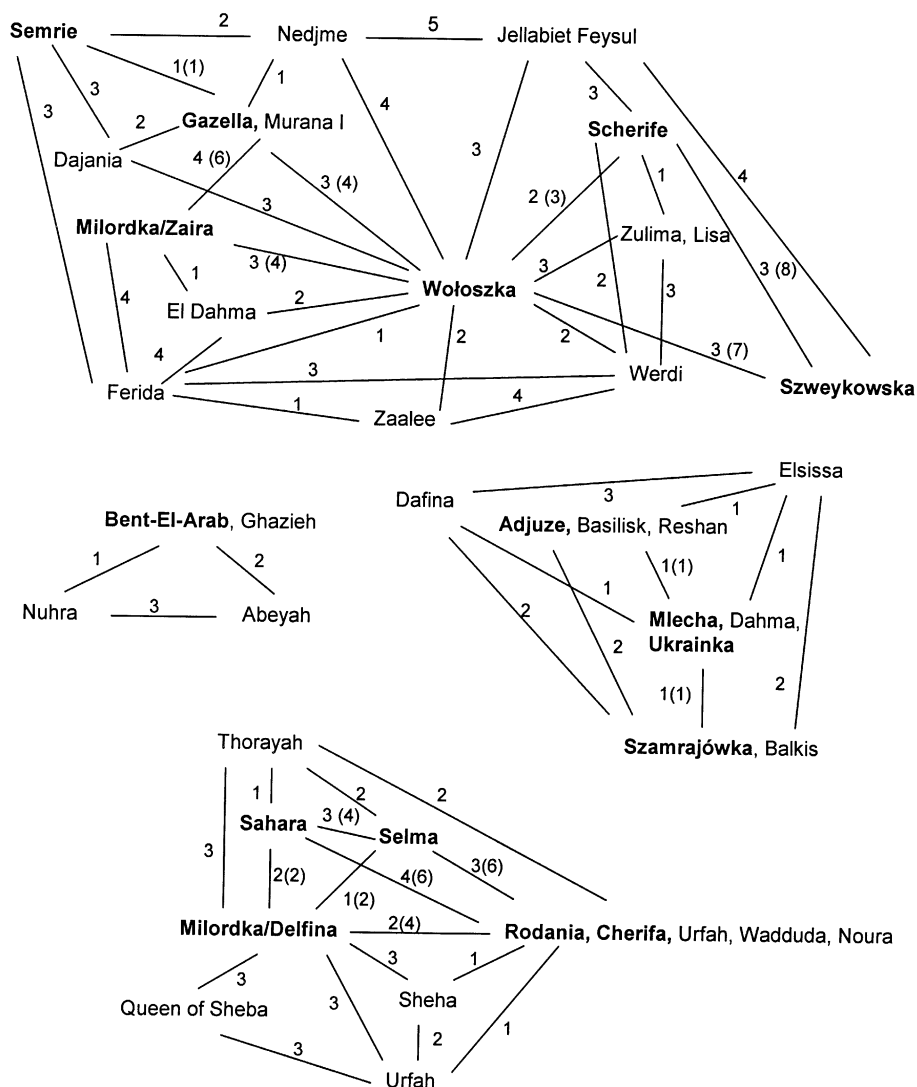


Figure 1. Relationship among Arabian dam lines based on sequence data of the 15 427–15 740 bp mtDNA region. The number of nucleotide substitutions is given; respective values in brackets correspond to the 15 400–15 857 bp fragment.

was found in the branch established by Milordka's daughter Zaira 1826 (Tab. I). This haplotype was identical (within the region analysed by Bowling *et al.* [4]) to the A25 haplotype found in representatives of the Milordka line in the USA. A second haplotype (E) was found in the horses from the branch that originated from Milordka's second daughter, Delfina 1825. The difference

between the haplotypes is significant and refers to 13 sites, which excludes the common origin of both branches. Considering the accuracy in stud registration and the clear system of foal marking at the Sławuta stud [2], errors regarding the identity of some mares from the line are rather improbable. Errors in reading the hand-written data of the studbook, however, cannot be excluded. In some of the documents from the Sławuta stud, copies of which are in Kwiatkowski's possession, the mare identified by Skorkowski [18] as Milordka's daughter Delfina, is registered as Delfinka 1825, the daughter of Malikarda ~1810. The last record is described in the documents with the annotation "z dawnego". This term can be translated as "from the old" (stud) and can be interpreted to be synonymous with the designation of Malikarda as a founder mare (Kwiatkowski, personal communication). The current findings provide strong evidence for the validity of the hypothesis that there are two independent lines in the present Milordka line.

Two haplotypes were also found in the Gazella line. The F haplotype was found in the branch continued by Gazella II born in 1914 in the Jezupol stud and her four daughters. The haplotype is identical, within the region compared, to the A19 haplotype found in the American representatives of the line [4]. The second haplotype (H) was identified in the branch continued by Fryga born in 1914 in the Pełkinie stud and her two daughters. This haplotype is identical to the haplotype found in the Sahara line in both Polish and American (A04 haplotype) horses. The presence of the H haplotype in the Fryga branch undermines the official pedigree data. The analysis of historical sources indicates some possible reasons of the error. Firstly, a proper stud register was not maintained in the Jarczowce stud [16], which might have led to the erroneous reconstruction of the Fryga pedigree. The faulty identification of one of the mare's ancestors is equally probable; after the death of the first owner in 1884, the horses from the Jarczowce stud were divided between his heirs into five studs, including Pełkinie and Jezupol [2]. However, an error in the identification of Fryga during World War I appears to be most probable. The Russian army confiscated almost all the horses from the Pełkinie stud in 1914, and only a few foals were left in the stud, together with Fryga [6]. Without stud data it is impossible to determine precisely the identity of the mare. Nevertheless, on the basis of information from three sources [3, 5, 12], it can be speculated that Fryga was actually a descendant of Sahara's great-granddaughter Mołodycia, whose daughter and two granddaughters were in the stud in 1906.

The pedigree errors found in the Milordka and Gazella lines concern 11% of the brood mares registered in 2002 in PASB. This number is comparable to the results of the Lipizzan horses analysis, where discrepancies between pedigree

data and mtDNA haplotypes were found in 18 of the 56 lines analysed, which also covered about 11% of the population [10].

The occurrence of the same haplotype in two lines is another issue worth consideration. The K haplotype was found in the Rodania and Cherifa lines. There is no reason to question the pedigree data of both lines. The occurrence of this haplotype in different lines was also confirmed by a study of American Arabs [4]. Haplotype A1, the 397 nucleotide-long fragment of the K haplotype, was noted in the Rodania, Wadduda, Urfah, and Noura lines. The same haplotype occurring in five lines may be recognised as evidence of their common origin. The existence of similar haplotypes in different dam lines has also been demonstrated in Lipizzan [10] and Thoroughbred [9] horses.

A common haplotype (G) was also found in the Mlecha and Ukrainka lines. The accuracy of the pedigree data of the Mlecha line is beyond all question. The main branches of the line were established by two of Mlecha's daughters (Tab. I). The G haplotype is identical to the A13 haplotype (within the region analysed) found in the American descendants of Mlecha [4]. The Ukrainka line was bred in Sławuta and then in the Gumniska (from the fifth generation) studs until 1944. There were no representatives of the Mlecha line in these studs [14, 15], which precludes potential errors in horse identification during stud activity. There is also no convincing evidence to suggest that the only continuator of the line (Forta 1943) was misidentified during wartime ([13, 14, 20], Pankiewicz, personal communication). Furthermore, it should be pointed out that identical haplotypes (within the region compared) were also found in American representatives of the Dahma line [4].

The presence of identical haplotypes in lines established by mares recognised as desert-bred Arabs and in those of unknown origin may be admitted as evidence of the Arabian origin of the latter. Different opinions on the ancestry of foundation stocks in the oldest Polish studs are presented in the historical sources. According to Skorkowski [17], the stocks were comprised of local breeding Tarpan-type mares (so-called Polish mares) that were repeatedly mated by oriental stallions. Bojanowski and Tarnowski [2, 19], in turn, reported that the foundation stock of the Biała Cerkiew stud in 1778 was composed of Polish, Persian, and Cherkesian mares. The results of the current study shed new light on the origin of the ancestors of the oldest Polish Arabian dam lines. This issue will be considered fully in the authors' next study.

It should be emphasised that any size difference in the mtDNA region analysed may influence the accuracy of the evaluation of genetic similarity. As shown in Figure 1, there are differences in the number of nucleotides that differentiate particular lines depending on the length of the mtDNA control region

analysed. For example, the Szweykowska and Wołoszka haplotypes differ by three (397 bp region) or seven sites (458 bp region), while the Szweykowska and Scherife haplotypes vary by three or eight sites, respectively.

The current analysis of Polish Arabian dam lines, similarly to those by Hill *et al.* [9] and Kavar *et al.* [10], demonstrates the usefulness of mtDNA analysis in the verification of historical pedigree records. As Hill *et al.* justly point out, the findings of such analyses may be of great importance for horse breeders since their decisions often depend on pedigree information.

ACKNOWLEDGEMENTS

This paper is dedicated to memory of Dr. Edward Skorkowski (1899–1985), the preeminent authority on Polish Arabian horse pedigrees, the first Secretary of the Arab Horse Breeding Society of Poland, the first editor of Polish Arabian Stud Book, and the compiler of Genealogical Charts of Polish Arabian horses.

REFERENCES

- [1] Altshul S.F., Gish W., Miller W., Myers E.W., Lipman D.J., Basic local alignment search tool, *J. Mol. Biol.* 215 (1990) 403–410.
- [2] Bojanowski S., Wycieczka po konie, *Jeździec i Myśliwy* 6, 13, 20 (1902).
- [3] Bojanowski S., *Sylwetki koni orientalnych i ich hodowców*, Kraków, 1906.
- [4] 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.
- [5] Czartoryski W., *Stado ks. W. Czartoryskiego w Pełkiniach, Jeździec i Myśliwy* 13 (1902).
- [6] Czartoryski W., *Stado Pełkińskie, Jeździec i Hodowca* 7 (1933) 110–112.
- [7] Głazewska I., Jezierski T., Pedigree analysis of Polish Arabian horses based on founder contributions, *Livest. Prod. Sci.* 90 (2004) 293–298.
- [8] Hall T.A., BioEdit: a user - friendly biological sequence alignment editor and analysis program for Windows 95/98/NT (1999), <http://www.mbio.ncsu.edu/BioEdit/bioedit.html>
- [9] Hill E.W., Bradley D.G., Al-Barody M., Ertugrul O., Splan R.K., Zakahrov I., Cunningham E.P., History and integrity of thoroughbred dam lines revealed in equine mtDNA variation, *Anim. Genet.* 33 (2002) 287–294.
- [10] Kavar T., Brem G., Habe F., Sölkner J., Dovč P., History of Lipizzan horse maternal lines as revealed by mtDNA analysis, *Genet. Sel. Evol.* 34 (2002) 635–648.
- [11] Kwiatkowski W., *A guide to the complete pedigrees of Arabian horses from Poland, Kawalkada Brwinów*, 1993.
- [12] Ostaszewski Ostoja K., *Oficjalna Księga Stad koni orientalnych dla Galicji i Bukowiny*, Kraków, 1899.

- [13] Pankiewicz R., Polska hodowla koni czystej krwi arabskiej 1918–1939, Ed. Petroniusz Frejlich, 2002.
- [14] PASB (Polish Arabian Stud Book) 1932 (vol. I) – 2003 (vol. XIV, supp. 2), Present Ed. Polski Klub Wyścigów Konnych Warszawa.
- [15] Pruski W., Stado arabskie w Gumniskach, *Jeździec i Hodowca* 12 (1930) 225–228.
- [16] Pruski W., Dwa wieki polskiej hodowli koni arabskich (1778–1978) i jej sukcesy na świecie, PWRiL Warszawa, 1983.
- [17] Skorkowski E., Koń arabski w Polsce, *Jeździec i Hodowca* 31 (1927) 351–353.
- [18] Skorkowski E., Tablice genealogiczne polskich koni arabskich czystej krwi, Instytut Zootechniki Kraków, 1960.
- [19] Tarnowski M., Stada hrabiów Branickich, *Jeździec i Hodowca* 30 (1934) 665–667.
- [20] Vetulani T., Wrażenia z wycieczki do Racotu, *Hodowca Koni* 4 (1946) 12–15.
- [21] Xu X., Aranson U., The complete mitochondrial DNA sequence of the horse *Equus caballus*: extensive heteroplasmy of the control region, *Gene* 148 (1994) 357–362.