Ann. Anim. Sci., Vol. 12, No. 4 (2012) 483–494, DOI: 10.2478/v10220-012-0040-4

GENETIC STRUCTURE OF HUCUL AND ANGLO-ARABIAN HORSES AT THE *TERT* LOCUS*

Tomasz Ząbek[1], Paweł Czapla[2,3], Monika Bugno-Poniewierska[1,2,3], Maciej Wnuk[2,3], Anna Lewińska[4], Bernadetta Oklejewicz[2,3], Grzegorz Bartosz[4], Ewa Słota[2,3]

[1]Laboratory of Genomics, National Research Institute of Animal Production, 32-083 Balice n. Kraków, Poland
[2]Department of Genetics, University of Rzeszów, Rejtana 16C, 35-959 Rzeszów, Poland
[3]Centre of Applied Biotechnology and Basic Sciences, University of Rzeszów, Werynia 502, 36-100 Kolbuszowa, Poland
[4]Department of Biochemistry and Cell Biology, University of Rzeszów, Rejtana 16C, 35-959 Rzeszów, Poland

Abstract

The objective of the study was to identify single nucleotide polymorphism (SNP) genetic markers in the equine *TERT* gene sequence, which were used to assess the degree of differentiation between Anglo-Arabian and Hucul horses. Polymorphisms were identified by sequencing 30 amplification products representing 18000 bp of *TERT* sequences. Twenty-seven SNP markers were investigated, which were at genetic equilibrium. Haplotypes and genotypes were determined, and usefulness of polymorphisms for genetic studies was assessed based on minor allele frequency (MAF). Alleles characteristic of both horse breeds were identified. SNP markers with MAF > 0.18 were considered suitable for genetic analyses concerning association studies and parentage testing. In total 26 haplotypes were identified, of which three were common to the investigated horse populations. Twelve haplotypes were found only in Anglo-Arabians and 11 in Hucul horses. Identified polymorphism of *TERT* gene might be useful in the search for genetic basis of aging in the *Equus caballus* species.

Key words: polymorphism, *TERT*, genetic structure, horse

The most common type of DNA sequence variants across genome are single nucleotide polymorphisms (SNPs). The determination of the complete genomic sequence for several species of mammals and the introduction of novel DNA genotyping and sequencing techniques on a genomic scale made it possible to identify from

* Author's project of the State Committee for Scientific Research, project no. N N311 068937.

several dozen thousand to one million informative SNP markers. The identification of numerous SNP polymorphisms and the availability of new technology in the form of DNA microarrays enabled developing new methods, suitable for genome-wide genetic studies in humans and farm animal species. A significant role in these studies is played by the identification of SNP haplotype blocks (a set of alleles representing linked loci), associated with particular phenotypic traits (Corbin et al., 2010). During selection of informative SNP markers based on population studies, account is taken of the frequency of the less frequent allele (minor allele frequency, MAF) in a given population (Matukumalli et al., 2009).

The polymorphism of genetic markers in farm animal populations provides valuable information on the amount of their genetic variation, which may be reduced as a result of breeding work. Large differences observed in the frequency of certain alleles between the populations of several cattle and horse breeds under comparison reflect the differences in the gene pool characteristic of populations that represent the commercial breeding of these species and of the native breeds represented by small populations covered by conservation programmes (Lubieniecka et al., 2001; Grzybowski and Prusak, 2004; Ząbek et al., 2005).

The aim of this study was to analyse the degree of sequence differentiation at the *TERT* gene locus of Hucul and Anglo-Arabian horses. The human *hTERT* gene spans 41881 base pairs and encodes one of the proteins of the telomerase ribonucleoprotein (RNP) complex (Nakamura et al., 1997). This enzyme is one of the factors regulating telomere length. Telomeres are tandemly repeated sequences located at chromosome ends, which stabilize the chromosome structure and serve as the cell's biological clock (Kipling, 1995). In the human *hTERT* gene region as many as 1218 polymorphic sites were identified, from which 29 are of probable functional significance (http://www.ncbi.nlm.nih.gov/gene?term=7015[uid]).

The equine *TERT* gene sequence (ENSECAG00000001347), deposited in the ENSEMBL database, is the chromosome ECA21 fragment of 17000 bp, in which predicted transcript sequence spans 1806 bp and consists of 14 exons ranging in length from 62 to 197 bp. Like in humans the equine *TERT* gene presents a potential source for extensive polymorphism needful for genetic structure analysis of differentiated horse populations and association studies concerning age related changes of the telomere length within equidae species.

Hucul and Anglo-Arabian horses differ considerably in type, conformation traits, environmental conditions they are exposed to, and nutritional requirements. Hucul horses (hc) are a local breed of primitive horses, included in the FAO/UNDP Global Strategy for the Management of Farm Animal Genetic Resources (Wężyk et al., 1995). Anglo-Arabians (aa), which arose from crosses between pure Arabian (oo) and a Thoroughbred (xx), are a specialized breed of warmblood horse, used as a sport horse worldwide. Differences between the breeds under comparison may translate into polymorphic variants of the *TERT* gene sequence, characteristic of the populations of horses representing different production types, and show polymorphisms useful for genetic analysis of the horse.

Material and methods

Samples of peripheral blood were collected from 50 animals representing two breeds of horses from the Studs in Walewice (25 Anglo-Arabian horses) and Odrzechowa (25 Hucul horses). Genomic DNA was isolated using Promega Wizard Genomic DNA Purification Kit (Promega). To amplify the *TERT* gene sequences, primers were designed for the amplification of 30 DNA fragments covering a DNA region of total length of 18000 bp. The reference sequence was obtained from the ENSEMBL database. Genomic DNA was amplified with a Geneamp 9700 thermocycler (Applied Biosystems), and the PCR products were sequenced from both strands with a BigDye Terminator Cycle Sequencing kit 3.1 (Applied Biosystems). The sequencing reaction was performed in a Veriti thermal cycler using a thermal programme recommended by the manufacturer (Applied Biosystems). Sequencing products were purified from unused reaction components using ethanol or DyeX Terminator chemistry (Applied Biosystems) and electrophoresed using POP-7 polymer on an ABI3130xl genetic analyser (Applied Biosystems).

The DNA sequences obtained were compared with the horse reference sequence to identify polymorphic SNP sites, using Genedoc program (Nicholas et al., 1997). The identified polymorphisms were verified using Sequence Scanner software (Applied Biosystems). Data from the Genedoc program, confirmed by analysis of sequencing reads, served to determine genotypes in individual SNP loci, which provided a basis for statistical calculations.

Allele and genotype frequencies were calculated and the test for genetic equilibrium (Guo and Thompson, 1992) was performed using Arlequin (Excoffier, 2010) and Genepop software (Rousset, 2008). Minor allele frequency (MAF), genotype distribution and heterozygosity were analysed for every SNP marker to determine the efficiency of SNP loci as markers useful for genetic tests in general. The observed heterozygosity (Ho) and expected heterozygosity (He) was calculated to determine the amount of genetic variation in the populations studied (Nei, 1978). To assess the degree of differentiation between the two horse populations, pairwise Wright's fixation index values (Fst) were calculated, which were used as a measure of genetic distance (Wright, 1969). In estimating the genetic distance values, we additionally used the Fst coefficient modified by Reynolds et al. (1983) and Slatkin (1995), which accounts for hypothetical time in which any given populations diverged.

To compare the genetic structure of both horse populations at the *TERT* sequence locus, haplotypes were analysed with the Expectation-Maximization (EM) algorithm using Arlequin software (Dempster et al., 1977; Excoffier and Slatkin, 1995). This is an iterative procedure that yields maximum-likelihood estimates of haplotype frequencies from multilocus genotype data with unknown linkage phase.

Results

Comparison of DNA sequences of the analysed horses with the reference sequence of the ECA21 chromosome in the *TERT* gene region resulted in the identifi-

cation of 65 polymorphic SNP sites (data not presented), from which a group of 27 markers with complete genotype in 50 analysed horses were selected. The results of the statistical test for 27 markers showed no deviations from Hardy-Weinberg equilibrium (Table 1). Because two loci were monomorphic in each of the horse breeds studied, the hypothesis of the population's genetic equilibrium was not tested for these markers (Table 1).

Table 1. Localization of 27 SNP polymorphisms and exact test results in Hucul horses (hc) and Anglo-Arabian horses (aa)

SNP no.	Name of marker	Genomic localization of SNP in the sequence of the ECA21 chromosome (5' - 3' orientation)	*TERT* gene region	Exact test result (P value)	
				hc	aa
1	H*	NC_009164:g.56790058G>T	intron 4	-	1
2	Zb	NC_009164:g.56900975C>T	intron 12	1	1
3	Ka	NC_009164:g.56791676C>A	intron 4	1	1
4	Kb	NC_009164:g.56791744A>G	intron 4	1	0.68763
5	Qa	NC_009164:g.56795179A>G	intron 8	0.54071	1
6	Qb	NC_009164:g.56795397C>T	intron 8	1	1
7	Qd*	NC_009164:g.56795668C>T	intron 8	1	-
8	Ra	NC_009164:g.56796093C>T	intron 8	1	1
9	S	NC_009164:g.56796351A>G	intron 9	1	0.68772
10	Sa	NC_009164:g.56796443A>G	intron 9	0.55786	1
11	V*	NC_009164:g.56798471C>G	intron 10	-	1
12	Va	NC_009164:g.56798652T>G	intron 10	0.18694	1
13	Wb	NC_009164:g.56799016A>G	intron 11	0.54191	1
14	Wc	NC_009164:g.56799083C>T	intron 11	1	1
15	Wd	NC_009164:g.56799055C>T	intron 11	1	0.68779
16	X	NC_009164:g.56799352A>G	intron 11	1	1
17	Xa	NC_009164:g.56799512C>T	intron 11	1	0.68866
18	Xb*	NC_009164:g.56799482A>G	intron 11	1	-
19	Xc	NC_009164:g.56799587C>G	intron 11	1	1
20	Oc	NC_009164:g.56794244C>T	intron 7	0.30107	1
21	Ob.	NC_009164:g.56794242C>T	intron 7	1	1
22	Oa	NC_009164:g.56794234A>G	intron 7	1	0.68842
23	Y	NC_009164:g.56799867A>G	intron 11	1	1
24	Ya	NC_009164:g.56800383C>T	intron 11	0.68849	1
25	Yb	NC_009164:g.56800428A>G	intron 11	0.68882	1
26	Yc	NC_009164:g.56800512A>G	intron 11	1	1
27	Yd	NC_009164:g.56800509A>T	intron 11	1	0.5421

* – monomorphic locus – exact test not performed.

Most *TERT* sequence polymorphisms were located in intron 11 (twelve polymorphisms), followed by intron 8 (four polymorphisms), introns 4 and 7 (three each), and introns 9 and 10 (two each). One polymorphic site was located in intron 12 (Table 1). Analysis of the genetic structure of the horse breeds studied revealed the presence of monomorphic loci in Anglo-Arabian (H and V loci) and Hucul horses (Qd and Xb

loci). The H_T and V_C alleles are only found in aa horses, and the Qd_C and Xb_A alleles only in hc horses. In the group of aa horses, the H_T and V_C alleles have low frequency (f<0.1) (Table 2). At the loci of 27 SNP markers, no differences were observed in minor allele frequency (MAF) variants between the horse populations studied. MAF greater than 0.18 was found at 10 SNP loci in aa horses (Kb, Qb, S, Wd, X, Xa, Oa, Y, Yc and Yd) and at 17 SNP loci in hc horses (Kb, Qa, Qb, Qd, S, Sa, Va, Wb, Wd, Xa, Xb, Oc, Oa, Y, Ya, Yb and Yc). The highest MAF (>0.4) was found at the loci of 6 markers in the group of aa horses (Kb, S, Wd, Xa, Oa and Yc) and at 3 SNP loci in the hc population (Qb, Ya and Yb). An equal proportion of both alleles in the gene pool of the hc population was found at the Ya and Yb loci. The lowest MAF (<0.1) was observed in the aa horse population at 11 SNP loci (H, Zb, Ka, Qa, Qd, Sa, V, Va, Wb, Xb, Oc, Ya and Yb), and in the group of hc horses at 6 loci (H, Ra, V, Wc, X, Xc, Ob and Y) (Table 2).

Table 2. Allele frequency (f) of individual loci in Hucul horses (hc) and Anglo-Arabian horses (aa)

Allele	aa	hc	Allele	aa	hc
H_T	0.02	0	Wd_C	0.46	0.78
H_G	0.98	1	Wd_T	0.54	0.22
Zb_C	0.04	0.1	X_A	0.18	0.08
Zb_T	0.96	0.9	X_G	0.82	0.92
Ka_C	0.98	0.88	Xa_C	0.54	0.22
Ka_A	0.02	0.12	Xa_T	0.46	0.78
Kb_A	0.54	0.22	Xb_A	0	0.24
Kb_G	0.46	0.78	Xb_G	1	0.76
Qa_A	0.08	0.2	Xc_C	0.9	0.92
Qa_G	0.92	0.8	Xc_G	0.1	0.08
Qb_C	0.22	0.58	Oc_C	0.92	0.78
Qb_T	0.78	0.42	Oc_T	0.08	0.22
Qd_C	0	0.24	Ob_C	0.9	0.92
Qd_T	1	0.76	Ob_T	0.1	0.08
Ra_C	0.9	0.92	Oa_A	0.46	0.78
Ra_T	0.1	0.08	Oa_G	0.54	0.22
S_A	0.54	0.22	Y_A	0.18	0.32
S_G	0.46	0.78	Y_G	0.82	0.68
Sa_A	0.08	0.18	Ya_C	0.96	0.5
Sa_G	0.92	0.82	Ya_T	0.04	0.5
V_C	0.06	0	Yb_A	0.04	0.5
V_G	0.94	1	Yb_G	0.96	0.5
Va_T	0.92	0.64	Yc_A	0.56	0.22
Va_G	0.08	0.36	Yc_G	0.44	0.78
Wb_A	0.08	0.2	Yd_A	0.8	0.94
Wb_G	0.92	0.8	Yd_T	0.2	0.06
Wc_C	0.9	0.92			
Wc_T	0.1	0.08			

aa –Anglo-Arabian horses, hc – Hucul horses.

Table 3. Frequency of genotypes in the horse populations studied in Hucul horses (hc) and Anglo-Arabian horses (aa)

SNP	Locus and genotype	Frequency		SNP	Locus and genotype	Frequency	
		aa	hc			aa	hc
H	TT	0	0	Qd	CC	0	0.04
	GT	0.04	0		TC	0	0.4
	GG	0.96	1		TT	1	0.56
Zb	CC	0	0	Ra	CC	0.8	0.84
	TC	0.08	0.2		TC	0.2	0.16
	TT	0.92	0.8		TT	0	0
Ka	CC	0.96	0.76	S	AA	0.32	0.04
	AC	0.04	0.24		GA	0.44	0.36
	AA	0	0		GG	0.24	0.6
Kb	AA	0.32	0.04	Sa	AA	0	0
	GA	0.44	0.36		GA	0.16	0.36
	GG	0.24	0.6		GG	0.84	0.64
Qa	AA	0	0	V	CC	0	0
	GA	0.16	0.4		GC	0.12	0
	GG	0.84	0.6		GG	0.88	1
Qb	CC	0.04	0.32	Va	TT	0.84	0.48
	TC	0.36	0.52		GT	0.16	0.32
	TT	0.6	0.16		GG	0	0.2
Wb	AA	0	0	Ob	CC	0.8	0.84
	GA	0.16	0.4		TC	0.2	0.16
	GG	0.84	0.6		TT	0	0
Wc	CC	0.8	0.84	Oa	AA	0.24	0.6
	TC	0.2	0.16		GA	0.44	0.36
	TT	0	0		GG	0.32	0.04
Wd	CC	0.24	0.6	Y	AA	0.04	0.08
	TC	0.44	0.36		GA	0.28	0.48
	TT	0.32	0.04		GG	0.68	0.44
X	AA	0.04	0	Ya	CC	0.92	0.28
	GA	0.28	0.16		TC	0.08	0.44
	GG	0.68	0.84		TT	0	0.28
Xa	CC	0.32	0.04	Yb	AA	0	0.28
	TC	0.44	0.36		GA	0.08	0.44
	TT	0.24	0.6		GG	0.92	0.28
Xb	AA	0	0.04	Yc	AA	0.32	0.04
	GA	0	0.4		GA	0.48	0.36
	GG	1	0.56		GG	0.2	0.6
Xc	CC	0.8	0.84	Yd	AA	0.6	0.88
	GC	0.2	0.16		TA	0.4	0.12
	GG	0	0		TT	0	0
Oc	CC	0.84	0.56				
	TC	0.16	0.44				
	TT	0	0				

Frequency distribution of homo- and heterozygous genotypes, close to the Mendelian proportions of 1:2:1, was found at the loci of 6 SNP markers in the group of aa horses (Kb, S, Wd, Xa, Oa and Yc) and at the 3 SNP loci of the hc population (Qb, Ya and Yb) (Table 3). In the population of aa horses, no homozygous genotype for the minor allele was established in 15 SNP loci (no homozygotes with the T allele at the H, Ra, Wc, Oc, Ob and Ya loci; no homozygotes with the A allele at the Ka, Qa, Sa and Wb loci; no homozygotes with the C allele at the Zb and V loci; and no homozygotes with the G allele at the Va and Xc loci) (Table 3).

In the hc horse population, no homozygotes for the minor allele were found at 7 SNP loci (no homozygotes with the T allele at the Ra, Wc, Ob and Yd loci; no homozygotes with the C allele at the Zb locus; no homozygotes with the A allele at the X locus; and no homozygotes with the G allele at the Xc locus) (Table 3).

Table 4. Values of observed (H obs.) and expected heterozygosity (H exp.) for 27 SNP markers in the horse populations studied

SNP	aa		hc	
	H obs.	H exp.	H obs.	H exp.
H	0.04	0.04	0	0
Zb	0.08	0.078	0.2	0.184
Ka	0.04	0.04	0.24	0.216
Kb	0.44	0.507	0.36	0.35
Qa	0.16	0.15	0.4	0.327
Qb	0.36	0.35	0.52	0.497
Qd	0	0	0.4	0.372
Ra	0.2	0.184	0.16	0.15
S	0.44	0.507	0.36	0.35
Sa	0.16	0.15	0.36	0.301
V	0.12	0.115	0	0
Va	0.16	0.15	0.32	0.47
Wb	0.16	0.15	0.4	0.327
Wc	0.2	0.184	0.16	0.15
Wd	0.44	0.507	0.36	0.35
X	0.28	0.301	0.16	0.15
Xa	0.44	0.507	0.36	0.35
Xb	0	0	0.4	0.372
Xc	0.2	0.184	0.16	0.15
Oc	0.16	0.15	0.44	0.35
Ob	0.2	0.184	0.16	0.15
Oa	0.44	0.507	0.36	0.35
Y	0.28	0.301	0.48	0.444
Ya	0.08	0.078	0.44	0.51
Yb	0.08	0.078	0.44	0.51
Yc	0.48	0.503	0.36	0.35
Yd	0.4	0.327	0.12	0.115
Mean	0.242	0.249	0.325	0.314

The analysis of observed and expected heterozygosity in the population of Anglo-Arabian horses revealed the presence of 10 SNP loci with heterozygosity values close to or greater than 0.3 (Kb, Qb, S, Wd, X, Xa, Oa, Y, Yc, Yd), 10 loci with heterozygosity values ranging from 0.1 to 0.2 (Qa, Ra, Sa, V, Va, Wb, Wc, Xc, Oc, Ob) and 5 loci with heterozygosity values below 0.1 (H, Zb, Ka, Oa, Ya, Yb). In the case of 7 SNP loci, at least 40% of the aa population is made up of horses with the heterozygous genotype (Kb, S, Wd, Xa, Oa, Yc, Yd). The greatest proportion of heterozygotes in the aa population is found at the Yc marker locus (0.48%) (Table 4). In the aa horse population, lower observed than expected heterozygosity values are found at the loci of Kb, S, Wd and Xa markers. At the Yd locus, aa heterozygotes were in excess of their expectation under Hardy-Weinberg equilibrium (Table 4).

Table 5. Haplotypes in both horse populations

Haplotype no.	Haplotype sequence	Frequency	
		aa	hc
1	GCAGGCTCGGGTGCCGTGCCCAGTAGA	0	0.1
2	GCCGGCTCGGGTGCCATGCCCAACGAT	0.02	0
3	GCCGGCTTGGGGGTCATGGCTAACGGA	0.02	0
4	GTAGGCTCGGGTGCCGTGCCCAGTAG	0	0.02
5	GTAGGCTCGGGTGCCGTGCCCAGTAGT	0.02	0
6	GTCAGTTCAGGTGCTGCGCCCGGCGAA	0.488103	0.22
7	GTCAGTTCAGGTGCTGCGCCCGGCGAT	0.051897	0
8	GTCGATTCGAGTACCGTGCTCAGCGGA	0.067138	0.18
9	GTCGATTCGAGTACCGTGCTCAGCGGT	0.012862	0
10	GTCGATTCGGGTACCGTGCTCAGCGGA	0	0.02
11	GTCGGCCCGGGGGCCGTACCCAATAGA	0	0.14
12	GTCGGCCCGGGGGCCGTACCCAATAGT	0	0.02
13	GTCGGCCCGGGGGCCGTACTCAATAGA	0	0.02
14	GTCGGCCCGGGTGCCGTACCCAATAGA	0	0.06
15	GTCGGCTCGGGGGCCGTGCCCAGTAGA	0	0.12
16	GTCGGCTCGGGGGCCGTGCCCAGTAGT	0.02	0.02
17	GTCGGCTCGGGTGCCATGCCCAACGGA	0.044759	0
18	GTCGGCTCGGGTGCCATGCCCAACGGT	0.015241	0
19	GTCGGCTTGGCTGTCATGGCTAACGGA	0.04	0
20	GTCGGCTTGGGGGTCATGGCTAACGGA	0	0.04
21	GTCGGCTTGGGGGTCATGGCTAACGGT	0.04	0
22	GTCGGCTTGGGTGTCATGGCTAACGGA	0	0.02
23	GTCGGCTTGGGTGTCATGGCTAACGGT	0	0.02
24	GTCGGTTCGGCTGCCGTGCCCAGCGGT	0.02	0
25	GTCGGTTCGGGTGCCGTGCCCAGCGGA	0.12	0
26	TTCGGTTCGGGTGCCGTGCCCAGCGGA	0.02	0

In the population of Hucul horses, we identified 17 loci with heterozygosity equal to or greater than 0.3 (Kb, Qa, Qb, Qd, S, Sa, Va, Wb, Wd, Xa, Xb, Oc, Oa, Y, Ya, Yb, Yc), 2 loci with heterozygosity equal to or greater than 0.2 (Zb, Ka) and 6 loci with heterozygosity ranging from 0.1 to 0.2 (Ra, Wc, X, Xc, Ob, Yd) (Table 4). At

9 SNP loci (Qa, Qd, Wb, Xb, Oc, Ya, Yb, Y, Qb) heterozygotes constitute approximately 40% of the Hucul horse population studied. At the Qb marker locus, heterozygotes form over 50% of genotypes in the hc horse population studied. At the Va, Ya and Yb loci, observed heterozygosity values were lower than expected values. At the Wb and Oc loci, hc heterozygotes were in excess of their expectation under Hardy-Weinberg equilibrium (Table 4).

Twenty-six haplotypes were identified, of which three (6, 8 and 16) occur in both Anglo-Arabian and Hucul horses. The most frequent haplotype in both populations is no. 6. Of all the haplotypes identified, five in Hucul horses (no. 1, 6, 8, 11 and 15) and two in Anglo-Arabian horses (no. 6 and 25) have a frequency greater than 0.1. The frequency of the other haplotypes ranges from 0.02 to 0.07 (Table 5).

Analysis of genetic distance was performed using Wright's fixation index (Fst) and its modifications by Slatkin and Reynolds. Pairwise Fst distance between the two populations was 0.1619, Fst distance proposed by Reynolds was 0.17662, and the distance based on Wright's fixation index modified by Slatkin was 0.19318.

Discussion

This study investigated *TERT* gene sequence polymorphism in *Equus caballus*, represented by two populations of horses differing in external conformation and purpose of breeding. All the *TERT* gene polymorphisms were localized in non-coding sequences, namely the intron regions with usually a greater number of polymorphic DNA sites in relation to coding regions. As many as 12 polymorphic SNP sites were identified in the intron 11 sequence. Such a high accumulation of natural mutations in intron sequences may suggest that factors that disrupt genetic equilibrium have a smaller effect in these genomic areas. This is because DNA repair processes (repair of errors during replication) are more often disturbed in non-coding sequences, resulting in the generation of spontaneous mutations arising during the course of evolution (Kunkel and Bebenek, 2000).

Analysis of genetic structure of the horse breeds studied at the loci of 27 SNP markers revealed alleles characteristic of aa and hc horses, which may indicate that Anglo-Arabian and Hucul horses are genetically distinct in the *TERT* gene sequence. Alleles at the Qd and Xb loci, which do not occur in the group of Anglo-Arabian horses, can be considered characteristic of hc horses. Likewise, alleles at the H and V loci are only characteristic of Anglo-Arabians, but because of their low frequency (H_T and V_C frequency <0.1) they can be regarded as artifacts due to polymerase errors in the sequencing reaction.

In the Hucul horse population, higher levels of genetic diversity are observed at the loci of 27 SNP markers. This is evidenced by a greater number of SNP loci with MAF greater than 0.18 in hc horses, which is a reflection of high heterozygosity. Smaller discrepancies in the distribution of allele frequency were observed in the population of Hucul horses. In the group of Anglo-Arabian horses, despite the twofold greater number of markers with the highest MAF (greater than 0.4), as many as

11 SNP loci in which MAF does not exceed one percent were identified. For comparison, in hc horses MAF smaller than 0.1 is characteristic of only 6 marker loci. The differences in allele frequencies between the analysed horse populations resulted in differences in the level of heterozygosity in individual SNP loci. The low MAF at SNP loci is associated with the lack of homozygotes for the minor allele, which is only found in the heterozygous system, thus drastically reducing the informativeness and suitability of the markers for genetic analysis in general. No homozygotes for the minor allele were noted at 15 SNP loci in the Anglo-Arabian population and at 7 SNP loci in the Hucul population. Meanwhile the 6 most informative SNP markers characterized by a high proportion of both alleles was identified in the Anglo-Arabian population and 3 in the Hucul horses. Using MAF >0.18 as a criterion of high heterozygosity, 8 SNP loci are potential markers suitable for genetic analysis concerning genome scanning or parentage verification tests in both horse breeds. By way of comparison, MAF for SNP loci, selected to develop DNA microarray suitable for genome scanning of cattle, ranged between 0.24 and 0.27 (Matukumalli et al., 2009). According to Werner et al. (2004), MAF greater than 0.1 served as a criterion for choosing SNP markers for parentage tests proposed in cattle. Analysis of mean heterozygosity for 27 SNP loci confirms that *TERT* gene sequence variants are more diverse in Hucul compared to Anglo-Arabian horses. This result confirms a number of population studies using microsatellite DNA typing data, which show that the amount of genetic variation is higher in the breeds or types of horses that are not subjected to intensive breeding (Hamanova et al., 2001; Ząbek et al., 2005). Differences in the extent of genetic variation between Anglo-Arabian and Hucul horses are translated into the level of diversity of both populations as a whole. Analysis of Wright's fixation index (Fst) showed that almost 16 to 19% of total genetic variation is formed by the differences between both horse groups. The Fst value, considered as a measure of genetic distance between the horse populations studied, is close to the Fst value between local populations of Spanish horses and Thoroughbreds (TH), a specialized breed subjected to targeted selection for racing performance traits. Genetic variation resulting from the differences between TH horses and the compared groups of Spanish horses (Asturcon, Gallego, Losin, Jaca Navarra) ranged from 12 to 16% (Canon et al., 2000). Considerable genetic differences in the gene pool of Anglo-Arabian and Hucul horses are also supported by the presence of haplotypes specific to both horse breeds. Out of 26 haplotypes identified, only 3 are shared by both horse populations. The presence of considerable differences in haplotypes results from large differences in the frequency of individual alleles and the presence of characteristic allelic variants in the group of Anglo-Arabian and Hucul horses. The presence of a greater number of haplotypes with frequency greater than 0.1 in Huculs further confirms that the gene pool of this breed shows greater variation compared to Anglo-Arabian horses.

The present study is novel in that it characterizes polymorphic sites in the *TERT* gene sequence of *Equus caballus*. Observed differences in the genetic variants of the *TERT* sequence between Hucul and Anglo-Arabian horses are not only related to their different ancestry but also may be the result of different breeding trends and environmental factors influencing their formation. This may affect also the lifespan

being different for primitive and modern horse breeds. As the *hTERT* gene polymorphism was found to be crucial for age related changes in humans, detected polymorphism in its equine homologue might facilitate finding genetic basis of aging in horses. Thus SNP typing data obtained for equine *TERT* gene in this work may be a basis for genetic association studies in horse breeds, whose longevity in general should be different.

References

Canon J., Checa M.L., Carleos C., Vega-Pla J.L., Vallejo M., Dunner S. (2000). The genetic structure of Spanish Celtic horse breeds inferred from microsatellite data. Anim. Genet., 31: 39–48.

Corbin L.J., Blott S.C., Swinburne J.E., Vaudin M., Bishop S.C., Woolliams J.A. (2010). Linkage disequilibrium and historical effective population size in the Thoroughbred horse. Anim. Genet., 41, Suppl. 2: 8–15.

Dempster A., Laird N., Rubin D. (1977). Maximum likelihood estimation from incomplete data via the EM algorithm. J. Roy. Stat.. Soc. B Met., 39 (1): 1–38.

Excoffier L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour., 10: 564–567.

Excoffier L., Slatkin M. (1995). Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. Mol. Biol. Evol., 12: 921–927.

Grzybowski G., Prusak B. (2004). Genetic variation in nine European cattle breeds as determined on the basis of microsatellite markers. III. Genetic integrity of Polish Red cattle included in the breeds preservation programme. Anim. Sci. Pap. Rep., 22: 45–56.

Guo S.W., Thompson E.A. (1992). Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics, 48: 361–372.

Hamanova K., Majzlik I., Glasnak V., Schröffelova D. (2001). Characterization and comparison of some horse breeds in Czech Republic as based on microsatellite markers polymorphism. Anim. Sci. Pap. Rep., 19: 219–230.

Kipling D. (1995). Telomerase: immortality enzyme or oncogene? Nat. Genet., 9: 104–106.

Kunkel T., Bebenek K. (2000). DNA replication fidelity. Annu. Rev. Biochem., 69: 497–529.

Lubieniecka J., Grzybowski G., Lubieniecki K. (2001). Genetic diversity of European cattle population (in Polish). Proc. XIV Conferrence of Polish Society of Genetics (Book of Abstracts), Poznań, pp. 251.

Matukumalli L.K., Lawley C.T., Schnabel R.D., Taylor J.F., Allan M.F., Heaton M.P., O'Connell J., Moore S.S., Smith T.P., Sonstegard T.S., Van Tassell C.P. (2009). Development and characterization of a high density SNP genotyping assay for cattle. PLoS One, 4: e5350.

Nakamura T.M., Morin G.B., Chapman K.B., Weinrich S.L., Andrews W.H., Lingner J., Harley C.B., Cech T.R. (1997). Telomerase catalytic subunit homologs from fission yeast and human. Science, 277: 955–959.

Nei M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics, 89: 583–590.

Nicholas K.B., Nicholas H.B.J., Deerfield D.W. (1997). GeneDoc: Analysis and visualization of genetic variation. EMBnet. News, 4: 1–4.

Reynolds J., Weir B.S., Cockerham C.C. (1983). Estimation of the coancestry coefficient: basis for a short-term genetic distance. Genetics, 105: 767–779.

Rousset F. (2008). Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. Mol. Ecol. Resources, 8: 103–106.

Slatkin M. (1995). A measure of population subdivision based on microsatellite allele frequencies. Genetics, 139: 457–462.

Werner F., Durstewitz G., Habermann F., Thaller G., Krämer W., Kollers S.,

Buitkamp J., Georges M., Brem G., Mosner J., Fries R. (2004). Detection and characterization of SNPs useful for identity control and parentage testing in major European dairy breeds. Anim. Genet., 35 (1): 44–49.
Wężyk S., Reklewski Z., Świtoński M. (1995). Scientific advances in genetics and animal husbandry in the years 1991–1993 (in Polish). Prz. Hod., 1: 21–24.
Wright S. (1969). The theory of gene frequencies: evolution and the genetics of populations. Chicago University Press, 2, pp. 291.
Ząbek T., Nogaj A., Radko A., Nogaj J., Slota E. (2005). Genetic variation of Polish endangered Bilgoraj horses and two common horse breeds in microsatellite loci. J. Appl. Genet., 46: 299–305.

Accepted for printing 4 IV 2012

TOMASZ ZĄBEK, PAWEŁ CZAPLA, MONIKA BUGNO-PONIEWIERSKA, MACIEJ WNUK, ANNA LEWIŃSKA, BERNADETTA OKLEJEWICZ, GRZEGORZ BARTOSZ, EWA SŁOTA

Struktura genetyczna koni huculskich i anglo-arabskich w locus *TERT*

STRESZCZENIE

Celem pracy była identyfikacja markerów genetycznych typu podstawień pojedynczego nukleotydu (SNP) w sekwencji genu *TERT* konia, które wykorzystano do oceny stopnia zróżnicowania genetycznego dwóch populacji koni różniących się eksterierem – 25 koni angloarabskich oraz 25 koni huculskich. Polimorfizmy zidentyfikowano metodą sekwencjonowania 30 produktów amplifikacji pokrywających łącznie 18000 pz sekwencji genu *TERT*. Zidentyfikowano 27 markerów SNP przydatnych do analiz statystycznych, które charakteryzował stan zachowania równowagi genetycznej. Określono haplotypy i genotypy oraz oceniono przydatność polimorfizmów do badań genetycznych na podstawie współczynnika MAF. Zidentyfikowano allele charakterystyczne dla obu ras koni. Za przydatne do analiz genetycznych z zakresu badań asocjacyjnych i kontroli rodowodów uznano markery SNP z MAF >0,18. Zidentyfikowano 26 haplotypów, spośród których 3 były wspólne dla badanych populacji koni. 12 haplotypów było charakterystycznych dla koni aa, a 11 dla koni hc. Zanalizowany polimorfizm w sekwencji genu *TERT* u badanych ras koni może być przydatny w poszukiwaniu genetycznych uwarunkowań procesu starzenia u gatunku *Equus caballus*.