Comparative characteristics of DNA polymorphisms of κ-casein gene (CSN3) in the horse and donkey

M. Selvaggi[1], A.G. D'Alessandro[2] and C. Dario[1]

[1]Department DETO - Section of Veterinary Science and Animal Production,
University of Bari 'Aldo Moro', Valenzano, Italy
[2]Department of Agro-Environmental and Territorial Sciences,
University of Bari 'Aldo Moro', Bari, Italy

Corresponding author: M. Selvaggi
E-mail: maria.selvaggi@uniba.it

ABSTRACT. The aims of this study were to assess the genetic variability in the exon 1 of the κ-casein gene in four Italian horse populations (Italian Saddle horse, Italian Trotter, Italian Heavy Draught horse, and Murgese horse) and in a sample of Martina Franca donkey by estimating genotype, allele and haplotype frequencies, as well as several population genetic indices. Genotyping of the selected polymorphisms was performed using the PCR-RFLP technique with two restriction enzymes: *Pst*I and *Bse*YI aimed to discover the presence of c.-66A>G and c.-36C>A polymorphism, respectively. Both these loci were found to be polymorphic in horses with some differences depending on the breed. No genetic variability was observed in Martina Franca donkey breed. In the equine species no selective pressure for milk purpose was performed, therefore the polymorphisms at milk protein loci were mainly considered as result of natural selection or as indirect consequence of selection oriented to increase body size or to improve conformation. From this point of view

these two single nucleotide polymorphisms and particularly the c.-36C>A one could be useful instruments for population studies.

Key words: CSN3 gene; Polymorphism; Horse; Donkey

INTRODUCTION

In the past, several authors reported the absence of κ-casein in mare's milk (Visser et al., 1982; Ono et al., 1989; Ochirkhuyag et al., 2000). Conversely, other researchers showed its presence, albeit at a low concentration (Kotts and Jenness, 1976; Malacarne et al., 2000; Iametti et al., 2001, 2002; Egito et al., 2002). The primary structure of equine κ-casein has been derived (Iametti et al., 2001; Lenasi et al., 2003; Miranda et al., 2004); it contains 165 amino acids residues (four less than bovine κ-casein but three more than human κ-casein). Equine κ-casein shows several biochemical properties similar to those of bovine and human κ-casein, such as the presence of carbohydrate moieties and the susceptibility to hydrolysis by chymosin-group II (Egito et al., 2001). κ-casein is the only glycosylated member of the casein family; it is located mainly on the surface of the casein micelles being responsible for their stability (Walstra, 1990). Due to their economical importance, κ-caseins and their genetic polymorphisms have been widely investigated in many ruminant species (Pinders et al., 1991; Prinzenberg et al., 2005; Moioli et al., 2007; Caroli et al., 2009; Selvaggi and Tufarelli, 2012; Selvaggi et al., 2014a,b). The comparison of amino acids sequences of the equine caseins with corresponding camel, pig, human, bovine, ovine and goat counterparts revealed sequences identity between 40 and 67% (Lenasi et al., 2003). Equine κ-casein proves to be the most conserved casein, closely followed by β-casein in agreement with the physiological function of these two proteins in micelles formation and with the role of κ-casein in milk coagulation (Holt, 1992). To date, only four SNPs were detected in exon 1 and exon 4 of the equine κ-casein gene (CSN3) (Hobor et al., 2006, 2008). The SNPs located in exon 1 were investigated with PCR-RFLP analysis using two restriction enzymes (*Pst*I and *Bse*YI). The polymorphism studied using *Pst*I is a transition A→G (c.-66A>G, GenBank AY579426). The second polymorphism is a transversion C→A (c.-36C>A, GenBank AY579426). Exon 1 is not part of the coding sequence of the gene, therefore these SNPs do not cause an amino acid substitution. However, due to the closeness between the CSN3 promoter region and the exon 1, this exon may be involved in the regulation of the expression of the gene. The coding region of the CSN3 gene comprises exon 3 and a majority of the exon 4. Two SNPs were localized in exon 4 of the equine CSN3 gene at position c.383A>T and c.517A>G causing two amino acid substitutions. A comparative analysis of gene sequences among horse, donkey and zebra was done only on a 400bp long fragment belonging to exon 4 of CSN3 gene (Hobor et al., 2008).

In *Equidae*, only few studies examined the *CSN3* gene so far. The aims of the present study were to investigate the two SNPs in exon 1 of *CSN3* gene in four Italian horse populations (Italian Saddle horse, Italian Trotter, Italian Heavy Draught horse and Murgese horse) and in a sample of Martina Franca donkey breed by estimating genotype, allele and haplotype frequencies, as well as several population genetic indices.

The Italian Saddle horse is a population of recent creation initially obtained by crossing many Italian breeds (mainly Maremmano and Sardinian) with English thoroughbreds, Arabian and Anglo-Arabian horses. To date, it is possible to use stallions belonging to different European stud books to have horses with a noble appearance, the robustness typical to the descendants of English thoroughbreds and the reliability of the Italian breeds. This breed is excellent for sportive

competitions. The Italian Trotter is a harness racing horse developed during the 19th century when trotting races came into popularity worldwide by crossing thoroughbred stallions with French (Norman) trotters, Russian Trotters and American Standardbred mares. These horses are strong and balanced and are considered among the best trotters in the world. The Italian Heavy Draught horse breed dates back to 1860 when the horse production branch of the Department of War began crossbreeding Breton stallions to local derived Hackney, Percheron and Breton mares. In 1926, a stud book was formed and until the mid-20th century this breed was very popular; in fact, these horses were not just powerful, but fast as well. Initially developed for agricultural and draught uses as well for artillery transport by the Italian army and for mules production, nowadays it is mainly used for meat production and heavy draught works. In 1976, a breed association was formed in Italy to preserve and promote the Italian Heavy Draught horse. The Murgese horse breed is very ancient and it has Spanish, Neapolitan, Berber and Arabian ancestors; its origins date back to the 15th century. This breed was fully developed in the 1920s; it originated from local horses found in the arid and rocky Italian hills of Apulia. The modern Murgese is amenable to equestrian tourism and as parade horse. Effective selection for the improvement and conservation of a pure breed has ensured these valuable aptitudes and desirable traits in the Murgese, which is now receiving rapid and well-deserved notoriety (Selvaggi et al., 2010). Martina Franca donkey derives from Catalan donkey breed, brought to Puglia at the time of the Spanish rule. It is tall and well built and it has good temperament. This donkey was used to carry cargoes by alpine troops, because it adapted itself to the territory difficulties. This breed has been greatly appreciated in the past for its elevated stature compared to other donkeys and it is considered useful for the production of hybrids. The typical breeding area is on Mediterranean woody scrubland at altitudes of more than 400 m s.l., with cold and rainy weather during winter and humid conditions in summer (Selvaggi and Dario, 2011). At present, asses are bred for amateur interest using a system of extensive or semi-extensive breeding (D'Alessandro et al., 2007).

MATERIAL AND METHODS

DNA samples

Genomic DNA samples were obtained from unrelated animals: one-hundred and ninety-three horses and seventy-five donkeys reared in southern Italy. The horses belonged to four different breeds: Murgese (N = 63), Italian Heavy Draught horse (N = 42), Italian Saddle horse (N = 48), Italian Trotter (N = 40). All donkeys belong to Martina Franca breed. DNA samples were extracted from individual whole blood samples collected on K3-EDTA tubes and stored at -25°C using the ZR Genomic DNA II Kit (Zymo Research).

PCR conditions and genotyping by PCR-RFLP

After DNA isolation, the samples were genotyped for the two SNPs in exon 1. Both these SNPs were determined as described by Hobor et al. (2008). The sequences of primers used for the amplification of a part of the promoter region and a part of the exon 1 of equine *CSN3* gene were as follows: CSN3F (forward): 5'-GATGACAACTCTATTTCGCCCT-3'; CSN3R (reverse): 5'-TTTGCAGGTCAGGTCTTGCT-3'. The 237 bp gene fragment harboring both the SNPs was amplified using thirty-six amplification cycles (94°C = 1 min, 57°C = 30 s, and 72°C = 45 s). The considered SNPs were genotyped by RFLP using two enzymes: *Pst*I (Fermentas; 2 U = 10 mL)

and *Bse*YI (NEB; 2 U = 20 mL). After digestion (37°C, 4 hours), the restriction fragments were analyzed using electrophoresis on a 2% agarose gel. The *Pst*I cuts the 237 bp product into 189 and 48 bp fragments for G allele, while A allele remains uncut. The following restriction fragments were expected: 189 and 48 bp (GG genotype); 237, 189, and 48 bp (AG); and 237 bp (AA). The *Bse*YI cuts the amplicon into 215 and 22 bp fragments for the C allele while allele A remains uncut. The possible genotype patterns were: 215 and 22 bp (CC); 237, 215, and 22 bp (CA); and 237 bp (AA).

The primers used for donkey samples were designed on the basis of the GenBank DNA sequence (EU429803) to amplify a part of the promoter region and a part of the exon 1 of CSN3 gene (Selvaggi and Dario, 2011). The sequences of primers were as follow: CSN3F (forward): 5'-GATGACAACTCTATTTCCCCCT-3'; CSN3R (reverse): 5'-CCAGGGTCAGGTCTTGCT-3'. The 235 bp gene fragment, harboring both the SNPs, was amplified using thirty-six amplification cycles (94°C = 1 min, 55°C = 30 s and 72°C = 45 s). The two considered SNPs were genotyped by RFLP with the same restriction enzymes and conditions used for equine samples. The *Pst*I potentially cuts the 235 bp product into 189 and 46 bp fragments for allele G, while allele A remains uncut. As a consequence the following restriction fragments were expected: 189 and 46 bp (GG genotype); 235, 189 and 46 bp (AG) and 235 bp (AA). The *Bse*YI cuts the amplicon into 215 and 20 bp fragments for the C allele, while allele A remains uncut. The possible genotype patterns were: 215 and 20 bp (CC); 235, 215 and 20 bp (CA) and 235 bp (AA).

Statistical analysis

The allele frequencies were calculated by simple allele counting. Polymorphisms were tested for deviation from Hardy-Weinberg equilibrium using the χ^2 test. Differences for both allelic and genotypic frequencies, for the two different loci, among and between the considered populations were analyzed using a χ^2 test. Population genetic indices, namely, gene heterozygosity (H_E), gene homozygosity (H_O), effective allele numbers (N_E), and fixation index (FIS) were performed by POPGENE32 software version 1.32 (Yeh et al., 2000). Moreover, PIC (polymorphism information content) was calculated according to Botstein et al. (1980). Haplotype estimations were performed by ARLEQUIN software version 3.11 (Excoffier et al., 2005) a program which estimates the frequency of haplotypes by maximum likelihood methods (Excoffier and Slatkin, 1995).

RESULTS AND DISCUSSION

In relation to c.-66A>G polymorphism, only two out of three possible genotypes were found in the considered populations: in particular, in the Murgese breed the frequencies of the AA and AG genotypes were 60.32 and 39.68%, respectively; in Italian Heavy Draught horse and Italian Saddle horse these frequencies were similar: 92.86 and 91.67% for AA genotypes and 7.14 and 8.33% for AG, respectively (see also Table 1). Italian Trotter and Martina Franca Donkey groups were found monomorphic: all the animals were genotyped as AA and as a consequence, no test for Hardy-Weinberg equilibrium was done. The distribution of genotypes was in Hardy-Weinberg equilibrium in Murgese sample (χ^2 = 3.683, df = 1, P = 0.055), Italian Heavy Draught horse (χ^2 = 0.038, df = 1, P = 0.846) and Italian Saddle horse (χ^2 = 0.067, df = 1, P = 0.795). As reported in Table 1 the A allele was the most frequent, its frequency ranging from 0.802 in Murgese to 1.000 in Martina Franca Donkey and Italian Trotter.

Table 1. Genotypic and allelic frequencies of both SNPs.

Breed	Genotype frequencies						Allele frequencies			
	c.-66A>G			c.-36C>A			c.-66A>G		c.-36C>A	
	AA	AG	GG	CC	CA	AA	A	G	C	A
Murgese	38 (60.32%)	25 (39.68%)	0	31 (49.20%)	32 (50.80%)	0	0.802	0.198	0.746	0.254
Italian Heavy Draught Horse	39 (92.86%)	3 (7.14%)	0	3 (7.14%)	0	39 (92.86%)	0.964	0.036	0.071	0.929
Italian Saddle Horse	44 (91.67%)	4 (8.33%)	0	28 (58.33%)	0	20 (41.67%)	0.958	0.042	0.583	0.417
Italian Trotter	40 (100%)	0	0	24 (60.00%)	4 (10.00%)	12 (30.00%)	1.000	0.000	0.650	0.350
Martina Franca Donkey	75 (100%)	0	0	0	0	75 (100%)	1.000	0.000	0.000	1.000

Considering all the populations as one, all the three possible polymorphic patterns (CC, CA and AA) were detected at c.-36C>A locus. In the sample of Murgese horses the observed frequencies of CC and CA genotypes at this locus were 49.20 and 50.80% and no AA animals were found. Conversely, the AA genotype was the most frequent in Italian Heavy Draught horse breed (92.86%) followed by CC one (7.14%), no heterozygous individuals were observed in this breed. In Italian Saddle horse population, the frequencies of the CC and AA genotypes were 58.33 and 41.67%, respectively. In Italian Trotter sample the observed frequencies were 60.00% (CC), 10.00% (CA) and 30.00% (AA). Martina Franca Donkey showed no polymorphism at this *locus*. The χ^2 test performed to verify the Hardy-Weinberg equilibrium revealed the absence of equilibrium in all horse breeds with a lack of heterozygosity in Italian Heavy Draught horse, Italian Saddle horse and Italian Trotter. As showed in Table 1, the frequency of C allele ranged from 0.000 in donkey breed to 0.746 in Murgese horse breed.

Genotypic frequencies for the two different polymorphisms at the c.-66A>G *locus* were found to be significantly different among the four horse breeds ($\chi^2 = 37.33$, df = 3, P < 0.001). The same results were found by comparing the genotypic frequencies observed at the c.-36C>A *locus* among the four breeds ($\chi^2 = 130.85$, df = 6, P < 0.001). Significant differences for allelic frequencies among equine populations were also revealed ($\chi^2 = 33.95$, df = 3, P < 0.001 and $\chi^2 = 100.36$, df = 3, P < 0.001 at c.-66A>G and c.-36C>A loci respectively).

In Tables 2 and 3 the differences for genotypic and allelic frequencies between breeds at the considered loci were showed. In relation to c.-66A>G polymorphism, no significant difference was found between the populations except for Murgese horse whose frequencies were different compared to those observed in Italian Heavy Draught horse, Italian Saddle horse and Italian Trotter (P < 0.01). The same result was obtained considering the differences between breeds related to allelic frequencies at the same *locus*.

Table 2. Statistical analyses of genotypic and allelic frequencies in the various equine breeds at the c.-66A>G locus.

	Murgese	Italian Heavy Draught Horse	Italian Saddle Horse	Italian Trotter
Murgese		$\chi^2 = 13.64$ (P < 0.01)	$\chi^2 = 13.87$ (P < 0.01)	$\chi^2 = 20.96$ (P < 0.01)
Italian Heavy Draught Horse	$\chi^2 = 11.55$ (P < 0.01)		$\chi^2 = 0.04$ (P = 0.83)	$\chi^2 = 2.97$ (P = 0.09)
Italian Saddle Horse	$\chi^2 = 11.79$ (P < 0.01)	$\chi^2 = 0.04$ (P = 0.84)		$\chi^2 = 3.49$ (P = 0.06)
Italian Trotter	$\chi^2 = 18.07$ (P < 0.01)	$\chi^2 = 2.91$ (P = 0.09)	$\chi^2 = 3.41$ (P = 0.06)	

χ^2 (P value) represented for differences of genotypic frequencies between two breeds in the up-triangle of above table; χ^2 (P value) represented for differences of allelic frequencies between two breeds in the down-triangle of above table. For all analyses, d.f. = 1.

Table 3. Statistical analyses of genotypic and allelic frequencies in the various equine breeds at the c.-36C>A locus.

	Murgese	Italian Heavy Draught Horse	Italian Saddle Horse	Italian Trotter
Murgese		χ^2 = 93.60 (P < 0.01)	χ^2 = 51.06 (P < 0.01)	χ^2 = 31.08 (P < 0.01)
Italian Heavy Draught Horse	χ^2 = 91.95 (P < 0.01)		χ^2 = 26.00 (P < 0.01) d.f. = 1	χ^2 = 34.60 (P < 0.01)
Italian Saddle Horse	χ^2 = 6.58 (P < 0.05)	χ^2 = 51.99 (P < 0.01)		χ^2 = 5.63 (P = 0.06)
Italian Trotter	χ^2 = 2.19 (P = 0.14)	χ^2 = 60.01 (P < 0.01)	χ^2 = 0.82 (P = 0.37)	

χ^2 (P value) represented for differences of genotypic frequencies between two breeds in the up-triangle of above table; χ^2 (P value) represented for differences of allelic frequencies between two breeds in the down-triangle of above table. For genotypic analyses, d.f. = 2; for allelic analyses, d.f. = 1.

As reported in Table 3, there were significant differences for genotypic and allelic frequencies at c.-36C>A *locus* except for Italian Trotter and Italian Saddle horse breed whose genotypic and allelic frequencies were similar P = 0.06 and P = 0.37 respectively. It was obvious that the breed factor significantly affected the distribution of genotypic and allelic patterns at c.-36C>A *locus*.

In the present populations, gene homozygosity (H_O), gene heterozygosity (H_E), effective allele numbers (N_E), polymorphism information content (PIC) and fixation index (F_{IS}) for both loci are shown in Table 4. Comparison of genetic diversity at c.-66A>G *locus* among the considered breeds demonstrates that Italian Heavy Draught horse and Italian Saddle horse had the highest homozygosity (0.929 and 0.917 respectively) and the lowest PIC (0.067 and 0.077 respectively), while Murgese had the lowest homozygosity (0.603) and the highest PIC (0.267). A slight excess of heterozygosity (F_{IS} = -0.248) was found in Murgese sample; this result was also supported by the values of N_E (1.467), in fact the higher the value of N_E the less homozygous are the alleles studied. F_{IS} is a measure of the deviation of genotypic frequencies from panmictic frequencies in terms of heterozygous deficiency or excess. Negative F_{IS} values indicate heterozygote excess and positive values indicate heterozygote deficiency compared with Hardy-Weinberg equilibrium expectations. The polymorphism information content (PIC) is a parameter indicative of the degree of informativeness of a marker. The PIC value may range from 0 to 1. In the studied populations PIC ranged from 0.000 to 0.267 for the c.-66A>G *locus*. According to the classification of PIC (low polymorphism if PIC value < 0.25, median if 0.25 < PIC value < 0.50, and high if PIC value > 0.50), this *locus* possessed low genetic diversity. Italian Trotter and Martina Franca Donkey populations showed the absence of polymorphism at c.-66A>G *locus* with the fixation of A allele.

Table 4. Genetic indices at the two polymorphic loci.

Locus		Murgese	Italian Heavy Draught Horse	Italian Saddle Horse	Italian Trotter	Martina Franca Donkey
c.-66A>G	Gene homozygosity (H_O)	0.603	0.929	1.000	1.000	1.000
	Gene heterozygosity (H_E)	0.397	0.071	0.000	0.000	0.000
	Effective allele number (N_E)	1.467	1.074	1.000	1.000	1.000
	Polymorphism information content (PIC)	0.267	0.067	0.000	0.000	0.000
	Fixation index (F_{IS})	-0.248	-0.037	-	-	-
c.-36C>A	Gene homozygosity (H_O)	0.492	1.000	1.000	0.900	1.000
	Gene heterozygosity (H_E)	0.508	0.000	0.000	0.100	0.000
	Effective allele number (N_E)	1.610	1.153	1.000	1.835	1.000
	Polymorphism information content (PIC)	0.307	0.123	0.000	0.351	0.000
	Fixation index (F_{IS})	-0.340	1.000	-	0.780	-

In relation to c.-36C>A *locus*, Murgese breed had the lowest homozygosity (0.492), similarly to as observed at c.-66A>G *locus*. Gene heterozygosity varied from 0.000 in Martina Franca Donkey, Italian Heavy Draught horse and Italian Saddle horse to 0.508 in Murgese breed and N_E ranged from 1.000 in Martina Franca donkey to 1.946 in Italian Saddle horse. As a consequence, in this population the highest value of PIC (0.368) was found. According to the classification of PIC, three populations out of five possessed middle genetic diversity at c.-36C>A *locus*. The value of F_{IS} obtained in Italian Saddle horse and in Italian Heavy Draught horse is due to the absolute absence of heterozygotes in these groups, indicating a high level of inbreeding. As for c.-66A>G *locus*, a slight excess of heterozygosity (F_{IS} = -0.340) was found in Murgese sample at c.-36C>A *locus*.

In Table 5 the combined genotypic frequencies and the possible haplotype frequencies were illustrated. Considering the five populations as one, five out of nine possible genotypic combinations were found. In consistence with the allelic frequencies, the most frequent combined genotype was AACC in Murgese, Italian Saddle horse and Italian Trotter groups, while the frequency of AAAA was the highest one in Italian Heavy Draught horse breed. It is important to underline that AACC genotype was found in all the considered horse breeds. AGCA combined genotype was found only in Murgese sample with a frequency of 39.68%. AGAA genotype was observed only in Italian Saddle horse and Italian Heavy Draught horse. On the basis of the analysis of the possible haplotypes, the most frequent haplotype in three out of four horse populations was AC; conversely, AA haplotype was the most frequent (0.893) in Italian Heavy Draught horse group. In table 6, the allele frequencies of both SNPs as previously observed by other authors in different populations were also reported (Hobor et al., 2008; Selvaggi et al., 2010; Bohaczyk and Mroczkowski, 2011; Selvaggi and Dario, 2011). The allelic frequencies found in Murgese breed at c.-66A>G *locus* were similar to those reported by Hobor et al. (2008) in Posavina horse and Slovenian warmblood and by Bohaczyk and Mroczkowski (2011). Moreover, Hobor et al. (2008) reported the fixation of the A allele at this *locus* in Slovenian Haflinger, Lippizan and Ljutomer trotter, similarly to as found in Italian Trotter in the present study. On the other side, the frequencies of G allele observed in Italian Heavy Draught horse and Italian saddle horse were really close to zero, without presence of individuals genotyped as GG. In the second restriction place, aimed to discover the presence of c.-36C>A polymorphism, the allelic frequencies observed in the present study and those reported by other authors showed a wide variability with different participation of both variants depending on horse breed, making this *locus* more informative than the first one. In fact, the C allele showed the highest frequencies in Murgese, Slovenian warm-blood, Ljutomer and Italian Trotter and Lippizan. The occurrence of alleles with similarly equalized frequencies was recorded in Italian Saddle horse as well as in Slovenian cold-blood, Slovenian Haflinger, Posavina horse and Polish cold-blood. Slovenian warm-blood and Murgese showed similar allelic frequencies even if it considers the two SNPs simultaneously as for Italian trotter and Slovenian Haflinger.

In the equine species no selective pressure for milk purpose was performed, therefore the polymorphisms at milk protein loci were mainly considered as result of natural selection or as indirect consequence of selection oriented to increase body size or to improve conformation, as suggested by Bohaczyk and Mroczkowski (2011) also considering that the investigated SNPs were located in a non-coding region. From this point of view these two SNPs and particularly the c.-36C>A polymorphism, could be useful instruments for population studies. Moreover, considering the increasing interest in mare's milk consumption and the strong association between κ-casein gene polymorphisms and production traits, well documented in other species, it will be interesting to study the possible relationship among these SNPs and some milk performance traits in horses.

Table 5. Frequencies of the combined genotypes (%) and haplotypes of the tested breeds.

	Martina Franca Donkey	Murgese	Italian Heavy Draught Horse	Italian Saddle Horse	Italian Trotter
Combined genotype					
AACC	-	49.21	7.14	58.33	60.00
AACA	-	11.11	-	-	10.00
AAAA	100.00	-	85.72	33.34	30.00
AGCC	-	-	-	-	-
AGCA	-	39.68	-	-	-
AGAA	-	-	7.14	8.33	-
GGCC	-	-	-	-	-
GGCA	-	-	-	-	-
GGAA	-	-	-	-	-
Haplotype (c.-66A>G; c.-36C>A)					
[A;A]	1.000	0.056	0.893	0.375	0.350
[A;C]	-	0.746	0.071	0.583	0.650
[G;A]	-	0.198	0.036	0.042	-
[G;C]	-	-	-	-	-

Table 6. Previously reported allele frequencies of the SNPs used here.

Breed	c.-66A>G		c.-36C>A		Reference
	A	G	C	A	
Slovenian cold-blood	0.600	0.400	0.520	0.470	Hobor et al., 2008
Ljutomer Trotter	1.000	-	0.800	0.200	Hobor et al., 2008
Slovenian Haflinger	1.000	-	0.570	0.430	Hobor et al., 2008
Posavina horse	0.820	0.180	0.550	0.450	Hobor et al., 2008
Slovenian warm-blood	0.800	0.200	0.800	0.200	Hobor et al., 2008
Lippizan	1.000	-	0.970	0.030	Hobor et al., 2008
Murgese horse	0.800	0.200	0.740	0.260	Selvaggi et al., 2010
Polish cold-blood	0.740	0.260	0.450	0.550	Bohaczyk and Mroczkowski, 2011
Martina Franca donkey	1.000	-	-	1.000	Selvaggi and Dario, 2011

ACKNOWLEDGMENTS

REFERENCES

Bohaczyk M and Mroczkowski S (2011). Polymorphism of LEP, PKM2 and CSN3 genes in Polish cold-blooded horse population. *Rocz. Nauk. PTZ.* 7: 9-19.

Botstein D, White RL, Skalnick MH and Davies RW (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphism. *Am. J. Hum. Genet.* 32: 314-331.

Caroli AM, Chessa S and Erhardt GJ (2009). Milk protein polymorphisms in cattle: Effect on animal breeding and human nutrition. *J. Dairy Sci.* 92: 5335-5352.

D'Alessandro AG, Martemucci G and Casamassima D (2007). Behavioural profile during suckling period in Martina Franca breed jennies reared under semi-extensive conditions. *Appl. Anim. Behav. Sci.* 106: 107-124.

Egito AS, Girardet JM, Miclo L, Mollé D, et al. (2001). Susceptibility of equine κ- and β-caseins to hydrolysis by chymosin. *Int. Dairy J.* 11: 885-893.

Egito AS, Miclo L, Lopez C, Adam A, et al. (2002). Separation and characterization of mare's milk αs1, β-, κ-caseins, γ-casein-like, and proteose peptone components 5-like peptides. *J. Dairy Sci.* 85: 697-706.

Excoffier L, Laval G and Schneider S (2005). Arlequin ver. 3.0 An integrated software package for population genetics data analysis. *Evol. Bioinform. Online*. 1: 47-50.

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

Hobor S, Kunej T and Dovc P (2008). Polymorphisms in the kappa casein (CSN3) gene in horse and comparative analysis of its promoter and coding region. *Anim. Genet*. 39: 520-530.

Hobor S, Kunej T, Lenasi T and Dovc P. (2006). Kappa Casein gene (CSN3) in horse. *Acta Agric. Slov.* 88: 83-89.

Holt C (1992). Structure and stability of bovine casein micelles. *Adv. Protein Chem*. 43: 63-151.

Iametti S, Tedeschi G, Oungre E and Bonomi F (2001). Primary structure of κ-casein isolated from mare's milk. *J. Dairy Res*. 68: 53-61.

Kotts C and Jenness R (1976). Isolation of κ-Casein-Like Proteins from Milks of Various Species. *J. Dairy Sci*. 59: 816-822.

Lenasi T, Rogelj I and Dovc P (2003). Characterization of equine cDNA sequences for αS1-, β-and κ-casein. *J. Dairy Res.* 70: 29-36.

Malacarne M, Summer A, Formaggioni P and Mariani P (2000). Observations on percentage distribution of the main mare milk caseins separated by reversed-phase HPLC. *Annali Facol. di Med. Vet. Univ. di Parma*. 20: 143-152.

Miranda G, Mahe´ MF, Leroux C and Martin P (2004). Proteomic tools to characterize the protein fractions of Equidae milk. *Proteomics.* 4: 2496-2509.

Moioli B, D'Andrea M and Pilla F (2007). Candidate genes affecting sheep and goat milk quality. *Small Ruminant Res*. 68: 179-192.

Ochirkhuyag B, Chobert JM, Dalgalarrondo M and Haertlé T (2009). Characterization of mare caseins. Identification of αs1 and αs2 caseins. *Le Lait*. 80: 223-235.

Ono T, Kohno H, Odagiri S and Takagi T (1989). Subunit components of casein micelles from bovine, ovine, caprine and equine milks. *J. Dairy Res*. 56: 61-66.

Pinders J, Perryc BN, Skidmoraend J and Sawa D (1991). Analysis of polymorphism in the bovine casein genes by use of the polymerase chain reaction. *Anim. Genet*. 22: 11-20.

Prinzenberg EM, Gutscher K, Chessa S, Caroli A et al. (2005). Caprine κ-casein (CSN3) polymorphisms: new developments in molecular knowledge. *J. Dairy Sci*. 88: 1490-1498.

Selvaggi M and Dario C (2011). Analysis of two single-nucleotide polymorphisms (SNPs) located in exon 1 of kappa-casein gene (CSN3) in Martina Franca donkey breed. *Afr. J. Biotechnol*. 10: 5118-5120.

Selvaggi M and Tufarelli V (2012). Caseins of goat and sheep milk: analytical and technological aspects. In: *Casein: Production, Uses and Health Effects*, Nova Science Publisher 1-26.

Selvaggi M, Pesce Delfino AR, and Dario C (2010). Exon 1 polymorphisms in the equine CSN3 gene: SNPs distribution analysis in Murgese horse breed. *Anim. Biotechnol*. 21: 252-256.

Selvaggi M, Laudadio V, Dario C, and Tufarelli V. (2014a). Major proteins in goat milk: an updated overview on genetic variability. *Mol. Biol. Rep*. 41: 1035-1048.

Selvaggi M, Laudadio V, Dario C, and Tufarelli V (2014b). Investigating the genetic polymorphism of sheep milk proteins: an useful tool for dairy production. *J. Sci. Food. Agric*. 94: 3069-3306.

Visser S, Jenness R and Mullin RJ (1982). Isolation and characterization of beta-and gamma-caseins from horse milk. *Biochem. J.* 203: 131-139.

Walstra P (1990). On the stability of casein micelles. *J. Dairy Sci*. 73: 1965-1979.

Yeh FC, Yang R, Boyle TJ, Ye Z, et al. (2000). PopGene32, Microsoft Windows-based freeware for population genetic analysis, version 1.32. Molecular Biology and Biotechnology Centre, University of Alberta, Edmonton, Alberta, Canada.