Acta agriculturae Slovenica, 88(december 2006)2, 83–89.

http://aas.bf.uni-lj.si

Agris category codes: L10 COBISS Code 1.01

KAPPA CASEIN GEN (CSN3) IN HORSE: GENETIC VARIABILITY IN EXON 1 AND 4

Sebastijan HOBOR [a], Tanja KUNEJ [b], Tina LENASI [c] and Peter DOVČ [d]

[a] Univ. of Ljubljana, Biotechnical Fac., Zootechnical Dept., Groblje3, SI-1230, Slovenia, email: shobor@bfro.uni-lj.si.
[b] Same address as [a], Ph.D.
[c] Departments of Medicine, Microbiology and Immunology, Rosalind Russell Medical Research Center, University of California at San Francisco, 533 Parnassus Avenue, San Francisco, CA 94143-0703, USA, Ph.D.
[d] Same address [a], Prof., Ph.D.

Received July 27, 2006, accepted December 03, 2006.
Delo je prispelo 27. julija 2006, sprejeto 03. decembra 2006.

ABSTRACT

Kappa casein (κ-CN) is milk protein that determines the size and specific function of the casein micelles, and its clevage by chymosine is responsible for milk coagulation. Any variation in gene promoter or coding sequence may change the expression of the gene or amino acid sequence, effecting functional properties of the protein. The mature κ-CN is encoded by part of the exon 3 and the entire exon 4. Since exon 3 has 33 bp and exon 4 is 497 bp long, the major part of the protein is encoded by exon 4. In this study we identified two SNPs in exon 1 and two in exon 4 of the horse kappa casein gene (*CSN3*) and genotyped them in three horse breeds. The nucleotide sequence of the first exon was included in this study due to its possible role in the regulation of the CSN3 expression. Because these polymorphisms were analysed for the first time, we used a reference method (RFLP) or at least two other complementig methods (Bi-PASA/PIRA and ASA-PCR/PIRA), for molecular genetic analysis of above mentioned SNPs. The highest variation in genotype frequencies was present in Slovenian cold blood breed. SNPs in exon 4 cause amino acid (AA) change in the mature product, and may very well render chemical/functional properties of the protein. Analysis of the consequences caused by changes in AA sequence, by online avaible program tools, comfirmed our hypotesis.

Key words: horses / molecular genetics / kappa casein / *CSN3* / nucleotides / polyorphism / genetic variation / amino acids / sequence

KAPA KAZEINSKI GEN (CSN3) PRI KONJU: GENETSKA VARIABILNOST V EKSONU 1 IN 4

IZVLEČEK

Kapa kazein (κ-CN) je mlečni protein, ki določa velikost in specifično funkcijo kazeinskih micel, njegova razgradnja s kimozinom pa je odgovorna za koagulacijo mleka. Sprememba v promotorju ali kodirajočem področju gena lahko vpliva na njegovo izražanje, oziroma spremeni aminokislinsko zaporedje, s tem pa vpliva na funkcionalnost proteina. Zrel κ-CN protein je deloma kodiran z eksonom 3 in s celotnim eksonom 4. Zaporedje eksona 1 smo vključili v raziskavo, ker je možno, da ima regulatorno vlogo. Ker je ekson 3 dolg 33 baznih parov, ekson 4 pa ima 497 baznih parov, je pretežni del proteina kodiran z eksonom 4. V tej študiji smo izvedli genetsko analizo dveh nukleotidnih zamenjav v eksonu 1 in dveh v eksonu 4 v genu za kapa kazein (*CSN3*) pri konju in jih genotipizirali pri treh pasmah. Ker ti polimorfizmi še nikoli niso bili analizirani, smo za genetsko molekularno analizo omenjenih polimorfizmov, uporabili referenčno metodo (RFLP) ali vsaj dve drugi dopolnjujoči metodi (Bi-PASA/PIRA in ASA-

PCR/PIRA). Največja raznolikost genotipov je bila prisotna pri Slovenski hladnokrvni pasmi. Nukleotidni zamenjavi v eksonu 4 povzročita zamenjavo aminokislin v končnem produktu, kar pa lahko spremeni kemične/fukcionalne lastnosti proteina. Analiza posledic sprememb aminokislinske sekvence, s programi na internetu, je potrdila našo hipotezo.

Ključne besede: konji / molekularna genetika / kapa kazein / *CSN3* / nukleotidi / polimorfizem / genetska variabilnost / aminokisline / zaporedje

INTRODUCTION

Kappa casein (κ-CN) is milk protein that determines the size and specific function of the milk micelles, and its clevage by chymosine is responsible for milk coagulation (Yahyaoui *et al*., 2003). K-CN differs from other caseins in its solubility over a broad range of calcium ion concentrations and containes a hydrophilic C-terminal region (Yahyaoui *et al*., 2003). Mature κ-CN protein has a labile peptide bond whose cleavage by chymosin produces a soluble hydrophilic glycopeptide (caseino-macropeptide) as well as insoluble peptide or para-κ-CN. The consequence of the secession of caseino-macropeptide from κ-CN in milk coagulation. The study of genetic polymorphisms of the caseins is of interest, since some variants could be more beneficial from the point of view of human nutrition or be associated with milk quality, composition and technological characteristics (Yahyaoui *et al*., 2003). In case of cows milk, B allelic variant of the *CSN3* gene is preferable and animals with genotype BB produce milk with better cheese making ability and shorter rennet clothing time.In sheep two κ-CN variants were found (Ceriotti *et al*., 2004), bovine κ-CN has 11 variants with A and B being the most common (Farrell *et al*., 2004) and in goat 16 κ-CN variants were identified (Prinzenberg *et al.* 2005). Since equine milk is already used for human nutrition it may be important to identify the most suitable equine κ-CN variants for various milk products. Bovine milk products are used as milk substitutes for human babies. Since bovine milk proteins sometimes trigger alergic reactions, it is important to look for alternative sources for substitution of mothers milk.

The aim of this work was to search for SNPs (single nucleotide polymorphysims) in equine *CSN3* gene, exon 1 and exon 4 and to asses the consequent changes in aminoacid sequence of the mature protein.

MATERIALS AND METHODS

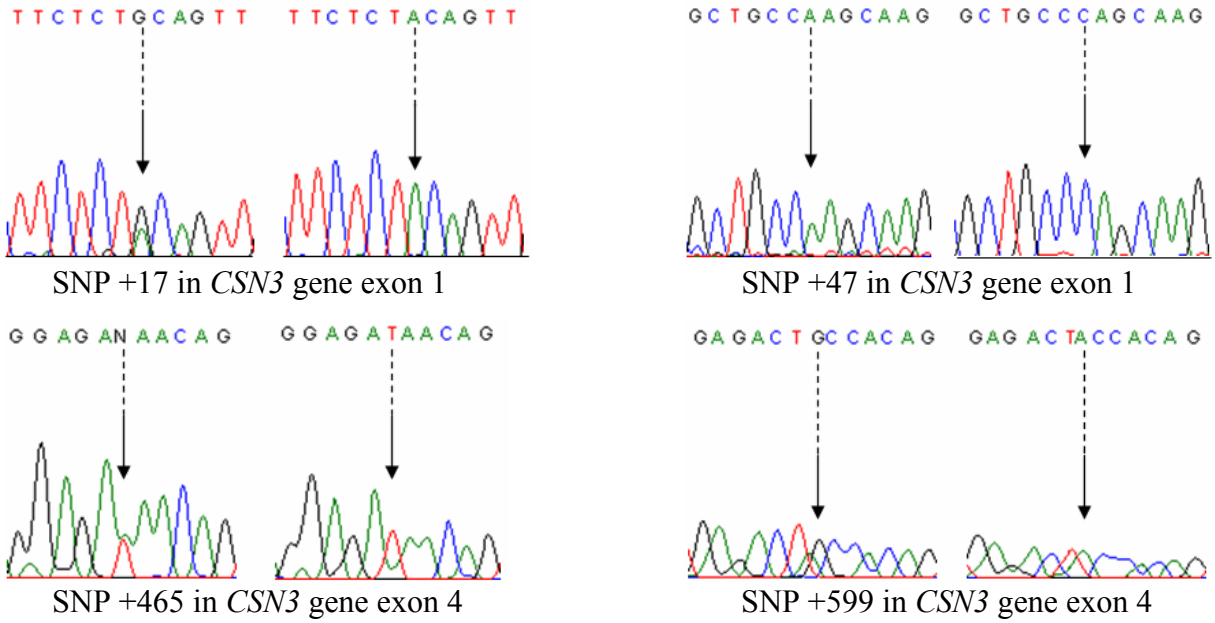
Detection of SNPs in exon 1 and exon 4

Blod samples of horses were received in our laboratory as part of the breeding kontrol program for various horse breeds in Slovenia. DNA was extracted from blood samples using a standard procedure (Miller, 1988). We analysed 51 DNA samples, 17 from each breed (Slovenian coldblood, trotter and haflinger)

Based on the known sequence of *CSN3* gene (acc. No. AY579426) we designed primers HKPR2-F and KEX1-R (Table 1) for amplification of 237 bp region that harboured the end part of the promoter and exon 1. 20 µl PCR reaction included 50 ng of genomic DNA, 1x PCR buffer, 1.5 mM $MgCl_2$, 150 µM dNTPs, 0.25 U *Taq* DNA polymerase (Fermentas, Vilnius, Litvania) and 0.25 pmol of each primer. The conditions of the PCR reaction were: initial denaturation at 94 °C for 5 min, followed by 30 cycles of denaturation at 94 °C (1 min), annealing at 53 °C (30 sec), extension at 72 °C (1 min) with final extension step at 72 °C for 7 min.

Primers for sequencig of exon 4 were designed on the basis of the known cDNA sequence (Lenasi *et al*., 2003) (acc. No. AY040863). PCR conditons were the same as described

previously. PCR products were sequenced by Macrogen (Seul, South Korea) (http://www.macrogen.com/eng/macrogen/macrogen_main.jsp). SNPs were detected by inspecting a DNA sequence manualy (Figure 1).



Slika 1. Prikaz SNP-jev v eksonu 1 in 4 v *CSN3* genu.
Figure 1. Illustrations of SNPs in exons 1 and 4 of *CSN3* gene.

Genotyping of SNPs in exon 1

The same PCR product used for sequencing of exon 1, was also used for genotyping of detected polymorphisms. SNPs +17 (A/G) and +47 (C/A) were genotyped by restriction fragment length polymorphism (RFLP) method using *Pst*I (Fermentas; Vilnius, Lithuania) and *Bse*YI (NEB; Ibswich, MA, USA) restriction endonucleases, respectively. Restrictions were performed in a volumen of 10 µl with 2 U of the encime and 1x restriction buffer at 37 °C for 3 hours. Restriction products were examined by electrophoresis on 3% agarose gels with 0.5x TBE buffer and stained with ethidium bromide.

Genotyping of SNPs in exon 4

SNP +465 (A/T) in exon 4 was genotyped by bidirectional PCR amplification of specific alleles (Bi-PASA) method (Liu *et al*., 1997 ; Jiang and Gibson, 1999) using four primers: left primer KEX4-F, rigt primer KEX4-R, inner T nucleotide-specific primer K-bi-pasa-T and inner A nucleotide-specific primer K-bi-pasa-A (Table 1). Annealing temperature was 58 °C. PCR products were examined on 2% agarose gels.

For confirmation of genotypes obtained with Bi-PASA we used primer-introduced restriction analysis (PIRA) (Ke *et al*., 2001). Primers for PIRA are in table 1, annealing temperature was 56.3 °C. Restriction of 142 bp PCR product was performed by 4 U of encime *Eco*RV (Fermentas, Vilnius, Lithuania) in 10 µl reaction, at 37 °C, over night. Restriction fragments were visualised on 4% agarose gel, stained with ethidium bromide.

SNP +599 (A/G) in exon 4 was genotyped by allele specific amplification-PCR (ASA-PCR) method (Hézard *et al*., 1997), specificity of the reaction was confirmed by

PIRA (Ke *et al*., 2001). In ASA-PCR reaction we used one common forward primer KEX4PO-F and two 3' allele-specific reverse primers: A-specific primer KEX4ASA-As-R, G-specific primer KEX4ASA-Gs-R (Table 1). Annealing temperature was 65 °C. Because one primer was prolonged with 18 bp G+C-rich 5' tail, the SNP specific use of the primer in reaction could be visualised on 3% agarose gel.

PCR for PIRA was performed with primers: KEX4-INTR-F and Pira+599-R (Table 1). Annealing temperature was 54 °C. Depending on the presence of the SNP at the site of introduced mismatch, endonuclease recognised the specific sequence. Restriction of 227 bp fragment was performed using the encime *Hpy*CH4V (NEB; Ibswich, MA, USA). Restriction products were examined on 4% agarose gels.

Table 1. Conditious for PCR-based identification of SNPs
Preglednica 1. Pogoji za detekcijo SNP z verižno reakcijo s polimerazo

SNP	Method, enzyme	primer name	sequence	Ta, °C
+17 (A/G)	RFLP *Pst*I	HKPR2.F KEX1.R	5'-GATGACAACTCTATTTCGCCCT-3' 5'-TTTGCAGGTCAGGTCTTGCT-3'	53
+47 (C/A)	RFLP *Bse*YI	HKPR2.F KEX1.R	5'-GATGACAACTCTATTTCGCCCT-3' 5'-TTTGCAGGTCAGGTCTTGCT-3'	53
+465 (A/T)	Bi-PASA	KEX4.F KEX4.R K-bi-pasa-T K-bi-pasa-A	5'-GATGAAAGGTTTTTCGATCTG-3' 5'-TGGACCACAGGTGAAGTAACTG-3' 5'-GGGGCGGGGAGAAACTTCAGGAGAT-3' 5'-GGGCGGGCTAGGGATGACTGTTT-3'	58
	PIRA *EcoRV*	kCN3cDNA-F PIRAEX4+465-R	5'-TGTTCCAAATTCCTCAATGGC-3' 5'-AATAGTATTGATCTTAGGGATGACTGAT-3'	56.3
+599 (A/G)	ASA-PCR	KEX4PO.F KEX4ASA-As.R KEX4ASA-Gs.R	5'-TCCTACCCCTGAACCAACAG-3' 5'-GAAGTAACTGGGACTGTGGT-3' 5'-GGGCGGGGCGGGGCGGGGGAAGTAACTGGGA CTGTGGC-3'	65
	PIRA *Hpy*CH4V	KEX4-INTR.F Pira+599.R	5'-CTGTGGTACGTCATCCATGC-3' 5'-AGGTGAAGTAACTGGGACTGT-3'	54

Amino acid exchange analysis

Consequent changes in amino acid (AA) sequence of the mature protein were analysed by three tools avaible online: ProtParam (http://www.expasy.org/tools/protparam.html), NetPhos 2.0 Server (http://www.cbs.dtu.dk/services/NetPhos/) and ProtScale (http://www.expasy.org/tools/protscale.html).

RESULTS

SNPs in exon 1 and 4 of horse *CSN3* gene and genotype frequencies

Genetic polymorphisms in kappa casein gene (*CSN3*) in 60 animals from three horse breeds, Slovenian Coldblood, Trotter and Haflinger were investigated. We identified four SNPs: +17 (A/G) and +47 (A/C) in exon 1 and +465 (A/T) and +599 (A/G) in exon 4 (acc. No AY040863). Genotype frequencies of four SNPs in Slovenian Cold blood (n = 20), Trotter (n = 20) and Haflinger (n = 20) breeds are listed in Table 2. By analysis of possible haplotypes we concluded that the most frequent haplotypes were ACTA and AATA (Table 3).

Table 2. Genotype frequences of the *CSN3* gene SNPs in three horse breeds
Preglednica 2. Frekvence genotipov nukleotidnih zamenjav v genu *CSN3* pri treh pasmah konj

		Exon 1						Exon 4					
SNP		+17			+47			+465			+599		
Genotypes		AA	AG	GG	CC	AC	AA	TT	AT	AA	GG	AG	AA
Coldblood	(n)	5	10	2	6	8	3	14	3	0	2	9	6
	(%)	0.29	0.58	0.12	0.35	0.47	0.17	0.82	0.17	0	0.12	0.53	0.35
Trotter	(n)	17	0	0	10	6	1	17	0	0	0	0	17
	(%)	1.00	0	0	0.58	0.35	0.06	1.00	0	0	0	0	1.00
Haflinger	(n)	17	0	0	5	10	2	17	0	0	1	0	16
	(%)	1.00	0	0	0.29	0.58	0.12	1.00	0	0	0.06	0	0.94

Table 3. Haplotye frequencies according to analysis of three horse breeds
Preglednica 3. Frekvence haplotipov glede na analizo treh pasem konj

	Exon 1		Exon 4		Freq.
Position	+17	+47	+465	+599	
Haplotype	A	C	T	A	0.82
	A	A	T	A	0.56
	G	A	T	G	0.21
	G	A	A	G	0.06
	G	C	A	G	0.04
	A	C	A	G	0.04

Change in amino acid sequence of κ-CN

Nucleotide substitutions in exon 4 cause potentialy significant coresponding AA change, that could effect structure in function of the mature protein. Substituted amino acids are different in charge and polarity. SNP +465 (A/T) causes substitution of Isoleucin (hydrophobic/nonpolar) for Lysine (hydrophilic/polar) and SNP +599 (A/G) causes substitution of Alanine (hydrophobic/nonpolar) for Threonine (uncharged/polar). ProtParam results for changes in molecular weight (Mw) and isoelectric point (pI) are in Table 4. We predicted loss of one highly potential phosphorilation site at position 153 With program NetPhos 2.0 (Figure 2) and with ProtScale, we predicted a change in polarity profile of the whole *CSN3* protein as a result of Isoleucin/Lysine exchange at position 108 (Figure 3).

Table 4. ProtParam predictions of theoretical Mw and pI of whole protein, for alternatively included amino acids
Preglednica 4. Napoved teoretične Mw in pI celotnega proteina, s programom ProtParam, za alternativno vključene aminokisline

Position	108	153	Theoretical Mw	Theoretical pI
Amino acid	Isoleucine	Threonine	18844.7	8.03
	Isoleucine	Alanine	18814.6	8.03
	Lysine	Alanine	18829.6	8.55

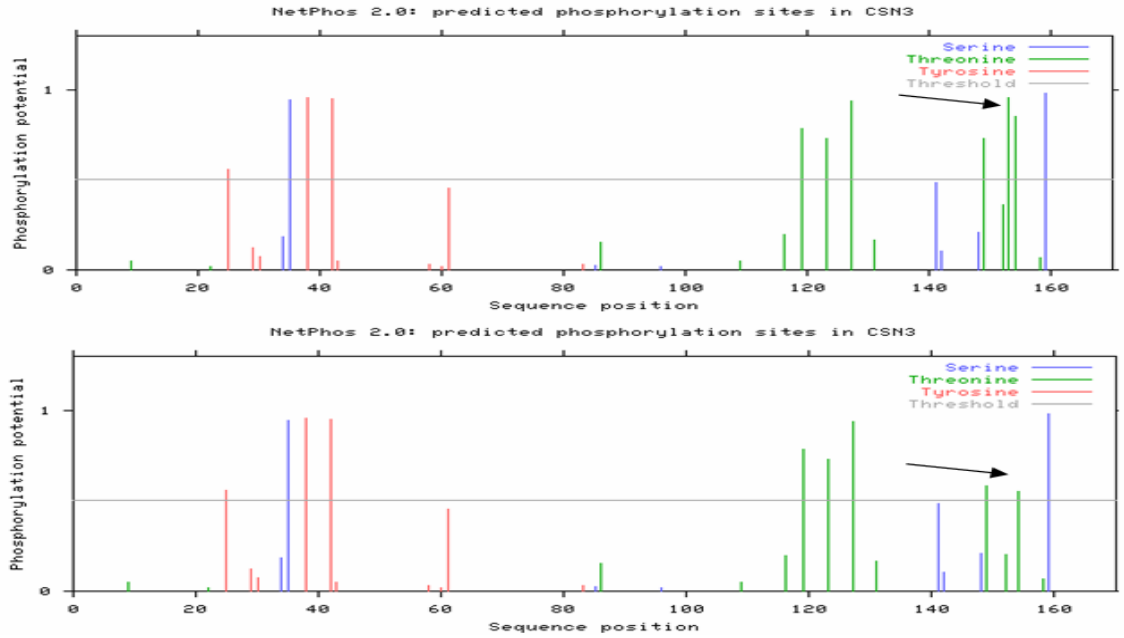


Figure 2. NetPhos 2.0 prediction of a consequent loos of one phosphorilation site at position 153.

Slika 2. Napoved izgube fosforilaciskega mesta na poziciji 153, s programom NetPhos 2.0.

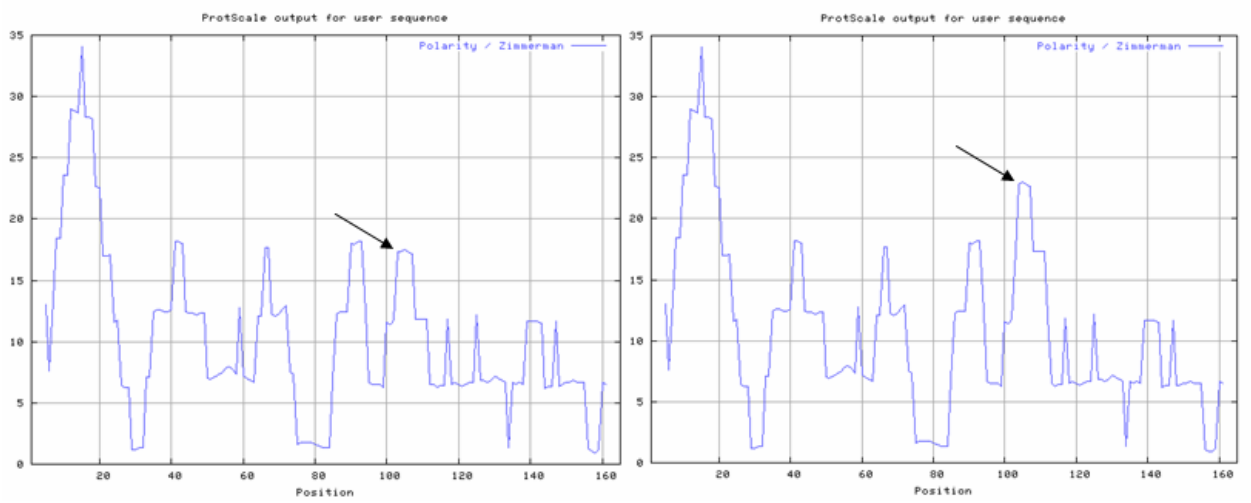


Figure 3. ProtScale prediction of different polarity profiles because of AA change at position 108.

Slika 3. Napoved drugačnega profila polarnosti zaradi zamenjave AK na poziciji 108, s programom ProtScale.

DISCUSSION

The study of four SNPs (+17, +47 in exon 1 and +465, +599 in exon 4) in three horse breeds revealed that the highest variation in genotype frequencies is present in Slovenian cold blood breed. In trotter and haflinger breed, variation was present only for SNP+47. Exon 1 is not part of the coding sequence, therefore SNPs in exon 1 do not cause AA substitutions. However, taking into account, that exon 1 is possibly involved in the regulation of expression of the gene,

functional testing of the detected SNPs could be promissing orientation for the future research. Nucleotide substitutions in exon 4 cause potentialy significant coresponding AA change, that could effect structure in function of the mature protein. Since substituted amino acids differ in charge and polarity, this might be an indication of two allelic variants of equine caseins and a starting point for the first nomenclature of these proteins in horse. To follow an astablished model for detecting new allelic variants of casein genes in cow (Farrell *et al*., 2004), further research at the proteomic level is needed. Two protein analysis methods, isoelectric focusing (IEF) and two dimensional electroforesis (2D-EF) should provide the final answer.

REFERENCES

Ceriotti, G./ Chessa, S./ Bolla, P./ Budelli, E./ Bianchi, L./ Duranti, E./ Caroli, A. Single nucleotide polymorphisms in the ovine casein genes detected by polymerase chain reaction-single strand conformation polymorphism. J Dairy Sci., 87(2004), 2606–2613.

Farrell, H.M. Jr./ Jimenez-Flores, R./ Bleck, G.T./ Brown, E.M./ Butler, J.E./ Creamer, L.K./ Hicks, C.L./ Hollar, C.M./ Ng-Kwai-Hang, K.F./ Swaisgood, H.E. Nomenclature of the Proteins of Cows' Milk-Sixth Revision. J Dairy Sci., 87(2004), 1641–1674.

Hézard, N./ Cornillet, P./ Droullé, C./ Gillot, L./ Potron, G./ Nguyen, P. Factor V Leiden: Detection in whole blood by ASA PCR using an additional mismatch in antepenultimate position. Tromb Res., 88(1997), 59–66.

Jiang, Z.H./ Gibson, J.P. Bi-PASA genotyping of a new polymorphism in the APOB gene shows no evidence for an association with fatness in pigs. Anim Genet., 30(1999), 54–56.

Ke, X./ Collins, A./ Ye, S. PIRA PCR designer for restriction analysis of single nucleotide polymorphisms. Bioinformatics, 17(2001), 838–839.

Lenasi, T./ Rogelj, I./ Dovc, P. Characterization of equine cDNA sequences for αs1, β and κ-casein. J Dairy Res., 70(2003), 29–36.

Liu, Q./ Thorland, E.C./ Heit, J.A./ Sommer, S.S. Overlapping PCR for bidirectional PCR amplification of specific alleles: A rapid one-tube method for simultaneously differentiating homozygotes and heterozygotes. Genome Res., 7(1997), 389–398.

Miller, S.A./ Dykes, D.D./ Polesky, H.F. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res., 16(1988), 1215–1216.

Prinzenberg, E.M./ Gutscher, K./ Chessa, S./ Caroli, A./ Erhardt, G. Caprine kappa-casein (CSN3) polymorphism: new developments in molecular knowledge. J Dairy Sci., 88(2005), 1490–1498.

Yahyaoui, M.H./ Angiolillo, A./ Pilla, F./ Sanchez, A./ Folch, J.M. Characterization and genotyping of the caprine κ-casein variants. J Dairy Sci., 86(2003), 2715–2720.