ABSTRACT

Currently, research interest is increasing in horse milk composition and its effect on human health. Despite previously published studies describing the presence of intra- and interbreed variability of equine milk components, no investigations have focused on the genetic background of this variation. Among horse caseins and the genes encoding them, least is known about the structure and expression of the α-S2 casein gene, CSN1S2. Herein, based on direct sequencing of the equine CSN1S2 coding sequence, we describe the presence of 51-bp insertion-deletion (in/del) polymorphism, which significantly changes the protein sequence (lack or presence of 17-amino acid serine-rich peptide). Bioinformatic analysis revealed that the observed in/del polymorphism spanned exactly 2 exons; therefore, we hypothesized that we were observing different CSN1S2 splicing isoforms. However, further investigation indicated that the detected sequence variation was caused by a large (1.3-kb) deletion in the genomic DNA. We found that the polymorphic forms (A, longer; B, shorter; KP658381 and KP658382 GenBank records, respectively) were unevenly distributed among different horse breeds (the highest frequency of variant B was observed in coldblood horses and Haflingers). We propose that the analyzed polymorphism is associated with CSN1S2 expression level (the highest expression was recorded for individuals carrying the BB genotype), which was much more pronounced for milk CSN1S2 protein content than for relative transcript abundance (measured in milk somatic cells). Our results provide insight into the equine CSN1S2 structure and lay a foundation for further functional analyses regarding, for example, allergenicity or physiochemical properties of the observed CSN1S2 variants.

Key words: horse milk, casein polymorphism, gene expression

INTRODUCTION

Interest has increased recently in the use of horse milk as a valuable product with numerous health-promoting attributes. Thus, investigations regarding mare milk composition and its effect on human health are increasing (Salimei and Fantuz, 2012). Even though the concentration of caseins in equine milk is significantly lower than that in bovine milk (~55 vs. 80% of total protein, respectively) (Uniacke-Lowe et al., 2010), it is still worth investigating, especially within the context of the potential allergenicity of mare milk.

Numerous genetic variants of the 4 caseins—α-S1-CN (CSN1S1), α-S2-CN (CSN1S2), β-CN (CSN2), and κ-casein (CSN3)—have been described in ruminants (Ramunno et al., 2005; Feligini et al., 2005; Caroli et al., 2009; Cosenza et al., 2009). Most have also been characterized at the DNA level. Association studies reveal that some of the discovered polymorphisms are related to variability in expression levels of casein gene and milk composition traits. Moreover, experiments on ruminants have shown that the occurrence of different variants of caseins might modulate the allergenicity of milk (Ballabio et al., 2011). This is important because milk protein hypersensitivity is one of the most frequent food allergies affecting both infants and adults (El-Agamy, 2007). From a molecular point of view, due to their multi-exon structure, the genes encoding caseins are considered a good model for alternative splicing process studies (Lenasi et al., 2006). Compared with ruminant species, knowledge about the genetic variants of the equine caseins (and the genes encoding them) is very limited. Only a few studies have examined the distribution of CSN2 and
CSN3 gene polymorphisms among different horse breeds (Hobor et al., 2008; Selvaggi et al., 2010). Additionally, alternative splicing variants of the equine CSN2 and CSN1S1 mRNA have also been described (Milenkovic et al., 2002; Lenasi et al., 2003, 2006). The least information is available for the CSN1S2 gene and protein. As shown in the GenBank database (accession number: NC_009146), the sequence of the equine CSN1S2 consists of 17 exons and spans about 16.5 kb. The predicted protein sequence (NP_001164238) contains 214 amino acids. There is no information about genetic variants of the horse CSN1S2 gene and protein. Among the Equidae family, more data are available for CSN1S2 gene and protein structure of the donkey. Cosenza et al. (2010) reported that donkey CSN1S2 has 2 forms (CSN1S2 I and II) that differ in structure and encoded protein sequence. The major form (CSN1S2 I) consists of 19 exons and encodes a 221-amino acid protein, whereas the CSN1S2 II form is shorter (16 exons and 168 encoded amino acids). Moreover, the existence of different splicing isoforms has been suggested for donkey CSN1S2 (Cosenza et al., 2010; Saletti et al., 2012). Considering that the structural variation of caseins may influence their allergenicity (Ballabio et al., 2012), and to assess the distribution of genetic variants across different horse breeds. Moreover, we evaluated the potential effect of the detected genetic variants on CSN1S2 expression variability at the mRNA (measured in milk somatic cells) and protein levels.

**MATERIALS AND METHODS**

**RNA Isolation and cDNA Synthesis**

Mares representing 3 horse breeds—Polish Primitive Horse (PPH, n = 20); Polish Coldblood Horse (PCH, n = 27); and Polish Warmblood Horse (PWH, n = 27)—originated from 4 Polish national studs (Kobylniki, Sieraków, Racot, and Nowe Jankowice). Investigated animals (mean age: 8.8 ± 3.3; mean lactation number: 4.33 ± 2.6) were kept in similar environmental conditions. The animals’ diet was consisted mainly of hay (~1.5 kg/100 kg of BW) and oat grains (0.5 to 1 kg/100 kg of BW) administered twice a day; water was supplied ad libitum. To closely monitor milk composition changes during lactation, mares were milked 3 times: at wk 5, 10, and 15 postpartum. Milking began after a short period (~1 min) of foal suckling. A total of 222 milk samples were collected. Samples were stored in liquid nitrogen temperature (for gene expression studies) and at −20°C (for milk composition analyses). Extraction of RNA was performed on 5-mL milk samples frozen in liquid nitrogen. After slow thawing, the milk sample was centrifuged (10 min, 4,000 × g), the supernatant was discarded, and the pellet was dissolved in 1 mL of TriPure Isolation Reagent (Roche, Indianapolis, IN). After a 5-min incubation of the pellet in TriPure, 200 μL of chloroform was added and shaken vigorously (30 s). Then, samples were incubated at room temperature (10 min) and centrifuged (15 min, 12,000 × g). The upper phase was removed by pipette, added to 0.5 mL of isopropanol, and incubated for 8 min at room temperature. Then, the obtained RNA was purified by 75% ethanol wash, dried at 40°C, and resuspended in water. Synthesis of cDNA [with the use of random hexamer and oligo(dt) primers mix] was conducted using the Transcriptor High Fidelity cDNA Synthesis Kit (Roche), according to the manufacturer’s instructions. Obtained cDNA was stored at −20°C until further analyses. The study was approved by the National Commission for Ethics of Animal Experimentation, Local Ethics Committee for Animal Research (Poznan, Poland; permission number: 39/2012). During milk collection, mares remained under veterinary control and did not manifest any disease symptoms.

**Screening for Polymorphisms and Animal Genotyping**

Screening for polymorphisms in the coding sequence of the equine CSN1S2 was performed for 24 randomly selected cDNA samples (PPH, n = 8; PCH, n = 8; PWH, n = 8) using the direct sequencing method (ABI Prism 3130 sequencer, Applied Biosystems, Foster City, CA). The PCR primers were designed using the Primer3 software (Koressaar and Remm, 2007) based on the NM_001170767 GenBank sequence and synthesized by Sigma-Aldrich (Steinheim, Germany). Amplification was conducted in a T-300 thermocycler (Bioread, Hercules, CA) applying the following conditions: initial denaturation (95°C, 5 min); 35 cycles of denaturation (95°C, 1 min), primer annealing (60°C, 1 min), and DNA synthesis (72°C, 1 min); and final synthesis (72°C, 10 min).

Genotyping of horses was performed by PCR amplification and electrophoresis (120 V, 45 min) of PCR products in 1.5% agarose gel (Supplemental Figure S1; http://dx.doi.org/10.3168/jds.2015-9807), using a combination of 3 primers (CSN1S1F, CSN1S1R, CSN1S2R) in a single reaction (final concentration of each primer was 0.3 μM). General PCR conditions were similar to those described above for cDNA amplification, except for primer annealing (63°C, 1 min) and DNA synthesis.
(72°C, 2 min). To assess the interbreed distribution of genotypes we used the genomic DNA of 869 horses representing 10 breeds (PWH, n = 86; PPH, n = 94; Arabian, n = 33; PCH, n = 152; Hallinger, n = 63; Quarter Horse, n = 20; Thoroughbred, n = 158; Belgian Draft Horse, n = 77; Welsh Pony, n = 87; and Shetland Pony, n = 99). All genomic DNA samples used for animal genotyping were provided by the Horse Genetic Markers Laboratory (Poznan University of Life Sciences, Poznan, Poland) from the routine parentage control analyses. Primer sequences and characteristics are reported in Supplemental Table S1 (http://dx.doi.org/10.3168/jds.2015-9807).

Assessment of Amino Acid Sequence Changes Using Mass Spectrometry

To confirm putative differences in the milk CSN1S2 AA sequence between 2 individuals, each one carrying a different homozygous genotype, the mass spectrometry technique was applied. Milk samples (100 μg) were fractionated by reverse-phase (RP) chromatography using a C8 column (J. T. Baker/Mallinckrodt, Center Valley, PA) and eluted sequentially using 20, 30, 40, 50, and 60% acetonitrile (ACN; Merck, Darmstadt, Germany). All fractions were then evaporated on a SpeedVac (Thermo Scientific, Waltham, MA). Ten-microgram aliquots of milk protein fractions as well as raw milk samples without fractionation were reduced, alkylated, and digested with 0.2 μg of sequencing-grade trypsin (Promega, Madison, WI) overnight at 37°C. The obtained peptide mixtures were analyzed in a quadrupole Orbitrap (QExactive, Thermo Scientific) spectrometer coupled with a Dionex UltiMate 3000 Nano LC system (Thermo Scientific). The effluent from the nanoLC column (RP C18 pre-column connected to a 25-cm-long, 75-μm-i.d. RP C18 column) was directly introduced into the Orbitrap in the positive electrospray ionization (ESI) mode. Samples were eluted from the column with a gradient of 5 to 60% ACN within 230 min (solvent A, 0.1% formic acid; solvent B, 99.9% ACN and 0.1% formic acid) at a flow rate of 0.3 μL/min. Acquisitions were run under the control of the Thermo QExactive Tune and Xcalibur software (Thermo Scientific). Protein database searches were performed via a Proteome Discoverer (Thermo Scientific). Proteins and peptides were identified using the Mascot (Matrix Science, London, UK) program against the SwissProt database (http://www.uniprot.org/). The false discovery rate for peptide identification was 0.05 in all analyses. The protein search was performed using the following search parameters: precursor-ion mass tolerance, ±10 ppm; fragment-ion mass tolerance, ±0.8 Da, and cysteine-treated with iodoacetamide to form carbamidomethyl-cysteine and optionally oxidized methionine residues. Trypsin was set as the enzyme, with a maximum of one missed cleavage.

Analysis of CSN1S2 Relative Transcript Level

Assessment of the CSN1S2 relative transcript level in equine milk somatic cells was performed in a LightCycler 480 instrument (Roche), based on TaqMan probe signal detection. Primers and probes were designed and synthesized by TIB Molbiol (Berlin, Germany); nucleotide sequences are shown in Supplemental Table S1 (http://dx.doi.org/10.3168/jds.2015-9807). Analysis of each sample was conducted in duplicate using the following cycling conditions: initial denaturation (95°C, 5 min), 45 cycles of denaturation (95°C, 10 s); primers and probe annealing (60°C, 30 s); and DNA synthesis (72°C, 1 s), and then samples were cooled (40°C, 30 s). Obtained results were normalized to the geometric mean of 4 internal reference genes: β-actin (ACTB), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), topoisomerase (DNA) II β (TOP2B); housekeeping genes) and keratin 8 (KRT8; marker of mammary epithelial cell number). Selection of the reference genes was conducted by the application of 2 bioinformatic algorithms: NormFinder and geNorm (Vandesompele et al., 2002; Andersen et al., 2004).

Analysis of CSN1S2 Concentration in Mare Milk

Five hundred microliters of milk was frozen at −20°C. A solution containing 0.1 M Bis-Tris buffer, 6 M guanidine hydrochloride, 5.37 mM sodium citrate, and 19.5 mM dithiothreitol was added directly to frozen aliquots at a 1:1 ratio (vol:vol) at room temperature. After thawing, each sample was shaken for 30 s, incubated for 1 h at room temperature, and centrifuged for 15 min at 14,000 × g in a microcentrifuge, and then the fat layer was removed. The remaining solubilized sample (3 mL) was diluted at 1:3 (vol:vol) with a solution containing 4.5 M guanidine hydrochloride and solvent A. Then, all solutions were filtered through a nylon filter.

The concentration of CSN1S2 in mare milk samples was established using an Agilent 1100 Series RP-HPLC (Agilent Technologies, Waldbronn, Germany) according to methodology similar to that described by Puppel et al. (2014) for whey proteins. The identification of CSN1S2 peaks was confirmed by comparison with the bovine protein standard (lot no. 075k7425, Sigma-Aldrich, St. Louis, MO). Separation was performed at ambient temperature using solvent gradient on a Jupiter column 5μ C18 300A (Phenomenex, Torrance, CA). The chromatographic conditions were as follows. Solvent A was ACN (Merck), water (Sigma-Aldrich),
and trifluoroacetic acid (Sigma-Aldrich) at a ratio of 100:900:1 (vol:vol:vol). Solvent B was ACN, water, and trifluoroacetic acid at a ratio of 900:100:1 (vol:vol:vol). The total run time was 16 min, the flow rate was 1.2 mL/min, and the detection wavelength was 220 nm. Injection volume of the final solution was 25 μL.

Statistical Analysis

A general mixed model, based on REML, to estimate all unknown variance-covariance parameters was applied to test the hypotheses that the mRNA and protein levels depend on genotype and lactation stage. The fixed effect of breed was also included in the model. The repeated-measure factor was sampling time (wk 5, 10, and 15 of lactation). To verify the hypotheses we used the F-test and the multiple comparison procedure based on least significant differences (LSD) with the Tukey-Kramer adjustment. All calculations were performed using SAS 9.3 (SAS Institute Inc., Cary, NC).

RESULTS AND DISCUSSION

Sequencing the CSN1S2 Gene

Direct sequencing of the equine CSN1S2 coding region (cDNA) of the 3 horse breeds revealed the presence of one insertion/deletion polymorphism spanning 51 nucleotides (Supplemental Figure S2; http://dx.doi.org/10.3168/jds.2015-9807). This polymorphism was initially found only in the PCH breed, with the following distribution of genotypes: AA (n = 2), AB (n = 5), and BB (n = 1). All horses representing the PPH (n = 8) and PWH (n = 8) breeds were AA genotype carriers. The detailed bioinformatic analysis showed that this polymorphism might be interesting because it spans 2 coding exons and changes the protein sequence by deleting 17 amino acids. Considering that we found different genotypes among the sequenced individuals (with or without 2 exons), we hypothesized 2 alternative splicing isoforms of the CSN1S2, probably related to the occurrence of another mutation present in some individuals (e.g., affecting a splicing acceptor or donor site). Starting from the 2 deleted exons, we sequenced a part of the CSN1S2 moving in both directions (to the 5’ and 3’ ends of the gene) based on the genomic DNA of the same horses for which the coding sequence (cDNA) had been sequenced previously. Although we found some polymorphisms close to the region of interest, none was associated with the pattern of the putative alternative splicing. When we checked the quality of the obtained PCR products using agarose gel electrophoresis, we found we were not able to amplify the CSN1S2 fragment in several individuals, independently of primer pairs used. Comparing these results with previously obtained cDNA sequences, we discovered that this problem with amplification was observed only for individuals with the detected deletion of 51 nucleotides (in the cDNA sequence). Therefore, we hypothesized that the observed variation in the coding sequence of the CSN1S2 was not an effect of alternative splicing but rather due to a large deletion occurring in the genomic DNA. To verify this hypothesis, we designed a few more primer pairs (moving to the 5’ and 3’ ends of the gene) to find the sequences for which we were able to amplify the expected fragments in all individuals (independently of the genotype observed in cDNA). Then, we designed a primer pair spanning the whole potentially deleted region. Amplification of DNA from several individuals representing the PCH breed and visualization of the results by agarose gel electrophoresis revealed the presence of 3 potential genotypes. Direct sequencing and a comparison of results with the CSN1S2 GenBank sequence confirmed the presence of a 1,339-bp deletion in the analyzed region. This fact confirmed our hypothesis that the observed variants of the equine CSN1S2 coding sequence are not related to alternative splicing but are the result of a large deletion in genomic DNA, spanning 2 coding exons, without shifting the open reading frame (ORF). Comparison of the 2 discovered CSN1S2 variants genomic sequences is shown in Supplemental Figure S3 (http://dx.doi.org/10.3168/jds.2015-9807).

Sequencing the entire coding region of the equine CSN1S2 as well as about 2,500 bp of its genomic sequence (spanning exons 7–9, according to the numbering shown in the NC_009146 GenBank sequence) and a comparison of obtained results with previously deposited GenBank resources (NC_009146 and NM_001170767) revealed some discrepancies. First, in its predominant form (variant A), equine CSN1S2 consists of 19 exons, not 17 (Figure 1). Interestingly, in the NC_009146 GenBank sequence, exons 9 and 10 are present but they are not assigned as exons. On the other hand, in the mRNA record (NM_001170767), the sequence corresponding to the shorter form (variant B) was described. Thus, according to the database, the CSN1S2 protein consists of 214 amino acids (NP_001164238), whereas results of this study and the almost-perfect homology with the previously published CSN1S2 protein sequence in the donkey (EMBL FM946022) strongly indicate that the main form of the equine CSN1S2 is the longer variant (231 amino acids, which corresponds to the polymorphic variant A).

Unfortunately, because of the very complicated molecular context observed in the region of interest (local-
ization of exons 8 and 10 in the duplicated fragment of 309 bp), we were not able to pinpoint out the exact localization of the detected 1,339-bp deletion (Figure 1). It is irrefutable that this deletion spans exon 9, but we cannot be sure which of the identical exons 8 and 10 is also harbored. Based on a detailed analysis of sequencing results, the most probable scenario is that the deletion spans exons 9 and 10. Our hypothesis is related to the fact that in some individuals, a SNP (c.4448A>G, according to NC_009146 GenBank sequence) was found in the first of the 309-bp duplicated fragments, about 190 bp upstream of exon 8 (Supplemental Figure S4; http://dx.doi.org/10.3168/jds.2015-9807). This polymorphism is also found in homozygous BB individuals (the shorter variant), which suggests that the fragment upstream of exon 8 remains in the sequence despite the presence of the 1,339-bp deletion. On the other hand, because of the “symmetrical” location of the beginning and the end of the deletion in duplicated fragments, it is possible that the exact location of the investigated deletion is shifted by an unknown number of nucleotides, within the range of 309-bp duplications (Figure 1). However, it should be noted that despite the multiple potential locations of the 1.3-kb deletion, its effect on the equine CSN1S2 genomic DNA, cDNA, and amino acid sequence is always identical.

Confirming Amino Acid Sequence Changes by Mass Spectrometry

Analysis performed for individuals carrying different homozygous genotypes confirmed the predicted effect of studied polymorphism on the equine CSN1S2 amino acid sequence. Because this analysis was conducted on the level of full milk samples (without previous extraction of CSN1S2 protein), about 80% of sequence coverage was attained. Fortunately, the region of interest (with or without additional amino acids) was read in all the analyzed samples, and the obtained results seem to be irrefutable. In the case of variant A, the existence of an additional serine-rich peptide (17 amino acids: ESAKFPTEVYSSSSSSE) was observed between positions 66 and 67 (if compared with the NP_001164238 sequence). However, if we compare it with the previously published sequence of donkey CSN1S2 (EMBL FM946022), we can see that the longer variant A should be treated as wild type (transformed in some individuals into variant B by the 1.3-kb deletion in CSN1S2). Because we cannot pinpoint the exact location of the deletion, we cannot exclude the alternative sequence of the deleted 17-amino acid serine-rich peptide VYSSSSSSEESAKFPTE (sequence encoded by exons 9 and 10). In the variant B (shorter sequence), none of the above-mentioned amino acids were observed. Sequences of both genetic variants were deposited in the GenBank database (KP658381 and KP658382 for A and B variants, respectively).

The investigated polymorphism of equine CSN1S2 is interesting because it significantly alters the amino acid sequence of CSN1S2. A comparison of the CSN1S2 protein sequence in different species (horse, donkey, cattle, sheep, and pig) showed the highest percentage identity between the horse and donkey (96.3%), which is partly related to the fact that the existence of the 2 identical exons (8 and 10) was described only in these equids. The percentage identity of the equine CSN1S2 protein sequence with other species varies around 60% (data not shown). Interestingly, previously published studies reported the presence of several CSN1S2 isoforms in donkey milk (Cosenza et al., 2010; Saletti et al., 2012) but none seem to correspond to the genetic variants described here.

One of the major factors influencing physicochemical properties of caseins (including their ability to form

![Figure 1](https://example.com/figure1.png)  
**Figure 1.** Updated structure of the equine αS2-CN (CSN1S2) based on NC_009146 sequence and the present study. Dup. = duplicated region of 309 bp; Del. = the most probable location of the 1,339-bp deletion; dashed line = alternate location of the 1,339-bp deletion; Ex = exon. All other possible deletion locations must be located (see Results) within the region marked by asterisks (****).
milk micelles and bind calcium) is the level of protein phosphorylation (Bijl et al., 2014). Thus, if more than 27% of serine residues (6 of 22) are removed from the CSN1S2 protein sequence because of the 1.3-kb deletion in the genomic DNA, it is likely that important differences exist in phosphorylation status between horse CSN1S2 genetic variants A and B. Such differences were previously described for splicing variants in horse CSN1S1 (Matéos et al., 2009). Thus, further studies are needed to determine whether the recorded differences in protein sequence between individuals carrying different variants (A or B) of the equine CSN1S2 have functional implications.

Although mare milk is generally considered less allergenic than ruminant milk, its tolerability is not identical for each person. Businco et al. (2000) reported that in about 8% of children with cow milk allergy, signs of hypersensitivity were also observed after using mare milk in a skin test. Some evidence suggests that the use of products (e.g., pills, cosmetics) containing horse milk may cause strong allergic reactions in adults (Doyen et al., 2013). Because the α-caseins are considered among the most allergenic milk proteins (Restani et al., 2009), the genetic variants of the equine CSN1S2 described in this study should be tested to determine whether they have different allergenicities. On the other hand, in vitro investigations of goat milk showed that the effect of CSN1S2 on overall casein fraction allergenicity is low (Marletta et al., 2004).

Numerous previously published reports have described caseins as an important source of bioactive peptides, which are activated after enzymatic proteolysis of the parent protein. These peptides may manifest various activities in the organism, such as regulation of metabolism and immunological reactions, antimicrobial activity, and so on (Phelan et al., 2009). Such peptides presenting antimicrobial activity are described also for bovine CSN1S2 (Zucht et al., 1995; Alvarez-Ordóñez et al., 2013). It seems clear that the function of casein-derived bioactive peptides will depend on the parent protein sequence and may be different for various casein genetic variants. Therefore, this study could lay the foundation for further analyses focused on potential equine milk-derived bioactive peptides.

**Distribution of CSN1S2 Genetic Variants A and B Among Different Horse Breeds**

Analysis performed for horses of 10 breeds revealed the presence of the investigated insertion/deletion (in/del) polymorphism in all breeds except for Arabian horses (in which only variant A was recorded). How-

![Figure 2](image_url)
ever, we observed an uneven distribution of genotypes between the analyzed breeds (Figure 2). The highest level of polymorphism was recorded for the Haflinger [minor allele frequency (MAF) = 0.44; Belgian Draft horse, MAF = 0.40; and PCH, MAF = 0.40]. A very low frequency of variant B was observed for the PPH breed (only a few individuals with the heterozygous genotype were recorded). In the Welsh pony, genotype frequencies revealed a significant ($P < 0.01$) deviation from Hardy-Weinberg equilibrium.

Assessment of the distribution of $CSN1S2$ genetic variants A and B among various horse breeds showed that the studied in/del polymorphism was present in almost all analyzed breeds but at different frequencies. This may suggest that the $CSN1S2$ sequence variation is phylogenetically old. On the other hand, we did not detect the 1.3-kb deletion in horses of the Arabian breed, which is considered one of the oldest breeds in the world and is an important component of many modern horse breeds. However, the group of Arabian

![Figure 3](image)

**Figure 3.** Comparison of mean $\alpha_{\text{s2}}$-CN ($CSN1S2$) expression levels between 3 time points of lactation. Values with different lowercase (uppercase) letters differ significantly at $P < 0.05$ ($P < 0.01$). Error bars indicate SEM.
horses used in the present study (n = 33) was too small to definitively exclude the presence of the studied polymorphism in this breed.

**Association of CSN1S2 A and B Variants with Gene Expression Level**

A comparison of the CSN1S2 mean expression level between 3 time points of lactation showed similar trends for mRNA and milk protein contents, with the lowest expression noted in wk 10 postpartum (Figure 3).

Based on previous investigations in ruminants, showing that some polymorphisms of caseins are related to milk protein content variability (Ramunno et al., 2001a,b), we decided to verify whether genetic variants (A and B) of the equine CSN1S2 were associated with its expression variability (measured at relative transcript abundance and milk CSN1S2 levels). Results showed a significantly elevated ($P < 0.05$) protein content for horses carrying the BB genotype (Table 1). This difference was pronounced ($P < 0.01$) if only the PCH breed was considered. A separate analysis for the PPH and PWH breeds was impossible because of the low frequency of the investigated polymorphism. Despite a similar tendency recorded for the relative transcript level (increased expression in the BB homozygous genotype), the overall effect of the genotype on mRNA abundance was not statistically significant ($P = 0.13$). These results should be verified in a larger animal group and if similar tendencies are noticed, some functional analyses regarding the role of the 1.3-kb deletion in equine CSN1S2 expression would be an interesting extension of the present study.

**CONCLUSIONS**

Our study indicates the presence of variation in the number of exons in equine CSN1S2, which is related to a large (1.3-kb) deletion in the genomic DNA. It is likely that this polymorphism affects gene expression and is thus associated with CSN1S2 content in mare milk. Because of the significant differences observed in the protein sequence between the 2 CSN1S2 variants (lack or presence of a 17-amino acid, serine-rich peptide), our study may be a promising introduction to further investigations of the physicochemical properties of horse CSN1S2 variants (e.g., protein phosphorylation status), as well as an assessment of their potential allergenicity.

**ACKNOWLEDGMENTS**

The study was funded by the National Science Centre (Poland; grant no. 2011/03/D/NZ9/05337).

**REFERENCES**


