Association of nucleotide-sequence polymorphism in the 5’-flanking regions of bovine casein genes with casein content in cow’s milk

Małgorzata SZYMANOWSKA, Eulalia SIADKOWSKA, Marek ŁUKASZEWICZ, Lech ZWIERZCHOWSKI*

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

(Received 27 August 2003; accepted 14 October 2004)

Abstract – The 5’-flanking regions (promoters) of the bovine αs1-, αs2- and β-casein genes were analysed for DNA sequence variants using PCR/RFLP in Polish Red (PR) and Black-and-White (BW) cattle. The polymorphic sites occurred at positions –1084 and –186 in the promoter region of the αs2-casein gene and at –728 and –109 in the αs1- and β-casein genes, respectively. These polymorphic sites were located within known potential regulatory sequences, suggesting an influence on the binding of transcription factors and expression of the bovine casein genes. Associations were found between various genotypes in the promoter region of the αs2-casein gene and αs1- and β-casein genotypes, thus showing the existence of intergenic haplotypes within the bovine casein locus. Milk proteins derived from cows of varying genotypes in the casein gene promoters were analysed using SDS-PAGE and HPLC techniques. It was shown that nucleotide sequence polymorphism in the promoter region of the bovine αs2-casein gene was associated with various contents of αs2- and β-casein in the milk.

Keywords: Bovine casein gene / 5’-flanking region / gene polymorphism / milk protein


Keywords: Gène de la caséine bovine / région 5’-non codante / polymorphisme / protéine du lait

* Corresponding author: l.zwierzchowski@ighz.pl
1. INTRODUCTION

In bovine milk, six major milk protein fractions – αs1-, αs2-, β-, κ-caseins, α-lactalbumin and β-lactoglobulin – exist in different allelic forms, which are controlled by codominant autosomal genes according to Mendelian inheritance. All four casein genes are clustered in a region less than 300 kb on bovine chromosome 6 in the order: αs1, β, αs2 and κ [10, 26, 31]. It is assumed that the three genes coding for calcium-sensitive αs1-, αs2- and β-casein evolved from one ancestral gene by exon shuffling [16] and intra- and intergenic duplications [4, 33]. In contrast, the κ-casein encoding gene evolved differently [1]. Lien et al. [21] reported a strong genetic linkage of the casein genes by using a “single sperm typing” method. So, one may suppose that close localisation and linkage of the casein genes might influence a common inheritance of these genes. Moreover, the close proximity of the casein genes supported the hypothesis of common hormonal regulation of the entire complex [27].

Genetic polymorphism of milk proteins in cows, sheep and goats is well documented [13, 20, 24]. The proteins differ in amino acid substitutions or deletions, resulting from the differences in the coding sequences of corresponding genes. Associations have been described between polymorphism in the coding regions of the milk protein genes and their expression levels in cattle [32]. Variable sites have also been identified in the 5’-noncoding sequences of these milk protein genes [3, 6, 12, 17, 28, 29]. At least some of the genetically variable sites are located in putative regulatory sequences, e.g. in transcription factor-binding sites [28, 30], and thus they may influence gene expression. It has been suggested that differential expression of various milk protein alleles is a possible result of linkage between variants of coding and regulatory regions of their genes [32].

The objective of this study was to examine associations between polymorphisms in the 5’-flanking region of the bovine αs1-, αs2- and β-casein genes and the casein content in cow’s milk.

2. MATERIALS AND METHODS

2.1. PCR-RFLP

Samples of blood were taken from Polish Red (PR) cows and heifers, maintained at the Research Station for Preserve Animal Breeding (Popielno); from animals kept by individual owners in the south of Poland (Podhale region); and from Polish Black-and-White (BW) cattle from the experimental station at Jastrzbiec. Numbers of animals used in individual experiments are given in tables and in the text. Genomic DNA was isolated from blood samples and collected on K2EDTA as described previously by Kanai et al. [18]. The PCR-based detection of restriction fragment length polymorphism (RFLP) was carried out using primers and restriction endonucleases given in the paper by Schild and Geldermann [28]. The primer sequences and polymerase-chain-reaction (PCR) conditions are summarised in Table 1. The PCR was performed in a reaction volume of 25 µL, containing ca. 100 ng of bovine genomic DNA, 0.25 µmol–1 of each primer, 160 µmol–1 dNTPs and 2 units of Taq polymerase. PCR-amplified αs2-casein gene fragments –1150/+117 and –1150/–872, were digested at 37 °C for 3 h with 5 units of EcoRV and MaeII, respectively. For αs1- and β-casein, respective gene fragments –1145/+101 or –1049/+87 were amplified and digested with SspI or NlaIII endonucleases. The products of restriction digestion were separated on 2% agarose gels in 1× TRIS-borate-EDTA (TBE) buffer. All PCR reactions were performed using the MJ Research PTC-225 Thermal Cycler.

2.2. Analysis of milk proteins

Samples of milk were collected from PR and BW cows that were known from previous RFLP analyses to carry various genotypes in the αs2 and αs1 gene promoters. Protein separation was carried out using SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and high performance liquid chromatography (HPLC).
Gene polymorphism and casein content in milk 581

2.3. SDS-PAGE

Polyacrylamide gel electrophoresis, performed under denaturating conditions with 0.1% SDS, was used to separate milk proteins. Skimmed milk samples, α-, β- and κ-casein standards [34], and an αs2-casein standard (received as a gift from Dr C.J. Slangen, Netherlands Institute for Dairy Research, NIZO) were denatured by boiling them with 2-mercaptoethanol (Sigma-Aldrich Corporation, St. Louis, MO, USA). All SDS-PAGE protein analyses were carried out as described by Grosclaude et al. [14]. The gels were run overnight at 180 V. After electrophoresis, gels were stained overnight with Coomassie brilliant blue R-250. The stained proteins were scanned in the FX Molecular Imager (Bio-Rad Laboratories Inc., Hercules, CA, USA), and densitometry was performed using the Quantity One program.

2.4. HPLC

High performance liquid chromatography analyses were carried out at the Institute of Animal Breeding, at Warsaw Agricultural University. All chemicals and solvents were purchased from Sigma (Sigma-Aldrich Corporation, St. Louis, MO, USA). Skimmed milk samples were precipitated with 10% CH3COOH at its isoelectric pH (4.6). Casein was dissolved in BIS-TRIS® buffer (pH 7.0) and filtered through 0.45 µm Milipore, Milex-XV filters (Milipore Corporation, Bedford, MA, USA). Proteins were fractionated and quantitatively analysed by HPLC, using a Hewlett Packard Agilent 1050 (Hewlett-Packard GmbH, Waldbronn, Germany). Reversed-phase columns (C4 and C18) were used. Elution was done with acetonitrile-water-trifluoroacetic acid solution. The flow rate ranged from 0.5 to 0.8 mL·min–1, and eluted casein peaks were detected at 214, 220 and 280 nm. Quantitative analyses were done with external standards purchased from Sigma, and with casein standards that were a gift (see 2.3).

2.5. Statistical analysis

Differences in casein content, in milk from cows carrying specific mutations in the casein gene promoters, were statistically evaluated with the least squares method using Harvey software [15].

The model used was:

\[ y_{ij} = \mu + M_i + LS_j + e_{ij} \]

where: \( y_{ij} \): casein content; \( \mu \): overall means; \( M_i \): effect of mutation (\( i = 1, 2 \)); \( LS_j \): combined effect of lactation number and lactation stage (\( j = 1, 11 \)); \( e_{ij} \): random error.
Occurrence of the common genotypes (haplotypes) was analysed by Chi-square testing.

3. RESULTS AND DISCUSSION

Restriction fragment length polymorphisms (RFLPs) of the bovine $\alpha_{s2}$-, $\alpha_{s1}$- and $\beta$-casein gene promoters were investigated at polymorphic sites previously reported by Schild and Geldermann [28]. In the $\alpha_{s2}$-casein gene promoter, polymorphisms occurred at positions –1084 and –186, relative to the transcription start point. Both of them were C/T transitions. In the case of the –1084 polymorphic site, a 278-bp DNA fragment was amplified and digested with MaeII endonuclease. Three genotypes were observed: CC, showing one undigested fragment; TT with two fragments (212 bp and 66 bp), and CT with three fragments – 278 bp, 212 bp and 66 bp (not shown). The analyses covered 195 PR and 158 BW cows. The frequency of allele C was 0.86 for BW, and 0.94 for PR cattle (Tab. II). The TT genotype was not found amongst PR animals, and only one TT animal was found within 60 BW animals (Tab. II). The frequency of allele T was 0.97 and 1.00 in PR and BW, respectively. The low frequency of allele C found in this study confirmed results reported earlier by Schild and Geldermann [28], who identified only two animals with allele C (in combination with T), from one Zebu and one German Simmental.

RFLP–EcoRV analyses of the polymorphic site at position –186, in the bovine $\alpha_{s2}$-casein gene promoter, resulted in two restriction fragments (965 bp and 302 bp) for the TT genotype, whilst the CT genotype showed three fragments 1267 bp, 965 bp and 302 bp (not shown). The TT and CT genotypes were found amongst 188 PR cows but only the TT genotype was found within 60 BW animals (Tab. II). The frequency of allele T was 0.97 and 1.00 in PR and BW, respectively. The low frequency of allele C found in this study confirmed results reported earlier by Schild and Geldermann [28], who identified only two animals with allele C (in combination with T), from one Zebu and one German Simmental.

The mutation in the bovine $\alpha_{s1}$-casein gene at position –728 was a deletion T/–. Digestion with SspI endonuclease showed two genotypes – (–/–) with two restriction fragments of 828 bp and 145 bp, and the (T/–) genotype with three fragments of 972 bp, 828 bp and 145 bp (not shown). The frequency of allele (–) was 0.95 and 0.88 for BW and PR cattle, respectively (Tab. II).

### Table II. Frequency of genotypes in 5’-flanking regions (promoters) of $\alpha_{s1}$- and $\alpha_{s2}$-casein genes in Polish Red and Black-and-White cattle.

<table>
<thead>
<tr>
<th>Locus/position of polymorphic site</th>
<th>Genotype</th>
<th>Black-and-White cattle</th>
<th>Polish Red cattle</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>number of animals</td>
<td>allele frequency</td>
</tr>
<tr>
<td>$\alpha_{s2}$-casein –186</td>
<td>TT</td>
<td>60</td>
<td>T - 1.00</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>0</td>
<td>C - 0.00</td>
</tr>
<tr>
<td></td>
<td>total: 60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_{s2}$-casein –1084</td>
<td>CC</td>
<td>116</td>
<td>C - 0.86</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>41</td>
<td>T - 0.14</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>total: 158</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_{s1}$-casein –728</td>
<td>(–/-)</td>
<td>49</td>
<td>(–) - 0.95</td>
</tr>
<tr>
<td></td>
<td>(T/-)</td>
<td>5</td>
<td>(T) - 0.05</td>
</tr>
<tr>
<td></td>
<td>total: 54</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The T allele appeared slightly more frequently in PR than in BW cattle (0.12 vs. 0.05). The high frequency of allele (–) was also reported previously by Schild and Geldermann [28], who found only two TT animals amongst 13 animals belonging to 7 cattle breeds.

The C→G transversion at position –109 in the 5'-flanking region of the β-casein gene can be recognised by the NlaIII restriction nuclease, but was not found in this study, either in PC or in BW cattle. All analysed animals were CC homozygotes. This mutation was previously reported [28], but was found in only one animal out of thirteen tested.

3.1. Analysis of milk proteins
Genetic variants and the content of the calcium-sensitive caseins αs1, αs2 and β in milk were analysed by SDS-PAGE and HPLC in cows carrying various genotypes within the αs1- and αs2-casein gene promoters. Milk samples were taken from 30 PR cows carrying either CT or TT genotypes in the αs2-casein gene (position –186) and from 22 BW cows with either CC or CT genotypes (position –1084). Milk from both BW and PR cows was included in these studies, since both breeds differ in the frequency and distribution of αs2-casein gene promoter genotypes, thus enabling the study of different combinations of gene variants between promoters and coding sequences. The results of representative milk protein separations carried out with SDS-PAGE and HPLC are shown in Figures 1–4.

Bovine αs1-casein mostly occurs as genetic variants A or B with variant A differing from B by a deletion of 13 amino
acids (a.a.) from a.a. position 14 to 26 [11]. In polyacrylamide gel electrophoresis, allele B of \(\alpha_s1\)-casein appeared as a slower migrating band in comparison with the fast migrating allele A (Figs. 3, 4). In HPLC, variant A of \(\alpha_s1\)-casein was eluted earlier from the column than variant B (Figs. 1, 2). The most common variants of bovine \(\beta\)-casein are A1 and A2, differing by His\(\rightarrow\)Pro substitution at a.a. position 67 [11]. The A1A1 genotype of \(\beta\)-casein appeared as a single slow migrating band, whereas A1A2 genotype variants migrated in two bands in SDS-PAGE (Figs. 3, 4), and showed two distinctly separated peaks in HPLC (Figs. 4, 5).

Numbers of genotypes and allele frequencies of \(\alpha_s1\)- and \(\beta\)-casein in BW and PR cows, as determined by the use of SDS-PAGE and HPLC, are shown in Table III. In the case of \(\alpha_s1\)-casein, allele B was mostly found with a frequency of 1.0 and 0.83 in BW and PR cattle, respectively. Known in the literature are five genetic variants of bovine \(\alpha_s1\)-casein – A, B, C, D and E. In European cattle breeds, variant B is present at a very high frequency (0.9), and is highest in the Ayrshire breed (1.0). Allele B was also found predominantly in PR cattle, while variants C and D were rare [8, 9, 25].

The A1 and A2 alleles of \(\beta\)-casein existed mostly as A1A1 or A1A2 genotypes in this study (Tab. III). The frequency of allele A1 was 0.77 and 0.79 in BW and PR cattle, respectively. Only one A2A2 BW individual was found among the 52 animals analysed. However, allele A1 was slightly more frequent when compared with earlier reports.
Gene polymorphism and casein content in milk

for Polish, German and Danish Red cattle (0.61, 0.62, and 0.71, respectively; Michalak [25], Feleczak [9], Litwinczuk [22], Erhardt [8]). In different populations of the Black-and-White cattle, variant A1 of β-casein was found at a frequency of 0.36–0.72 [11].

In this study, only variant A of the αs2-casein was found among PR and BW cattle. In SDS-PAGE, αs2-casein migrated as a thin band seen at the front of αs1-casein (Fig. 3). In HPLC diagrams, αs2-casein formed a small peak migrating between κ- and αs1-casein fractions (Figs. 1, 2). Interestingly, in several milk samples (derived from 4 BW and 4 PR cows) no detectable amounts of αs2-casein were found (Fig. 1B). However, this phenomenon was apparently not associated with considered mutations within the αs2-casein gene promoter (C/T at positions –1084 and –186), since milk from animals with various 5’-flanking region genotypes were represented. Four genetic variants of bovine αs2-casein – A, B, C and D – are known. Only variant A was found in most European cattle breeds; variant C was found in yaks and variant B in Bos indicus (Zebu) [7]. In BW cattle, the frequency of variant A of αs2-casein was 1.0 [11], while...
in Polish and German Red cattle variant A was predominant (0.98), and D was rare (0.02) [8].

Genotyping of caseins in individual animals showed significant associations between specific variants of $\alpha_{s2}$-casein gene promoters and the genotypes of $\alpha_{s1}$- and $\beta$-caseins; some combinations were much more frequent than others (Tab. IV). In PR cattle, significant association (Chi-square, $P \leq 0.001$) was found between variants in the 5' flanking region (position –186) of the $\alpha_{s2}$-casein gene and the $\alpha_{s1}$-casein genotype. Allele A of $\alpha_{s1}$-casein was associated with genotype CT in the $\alpha_{s2}$ promoter, whereas allele B was associated with genotype TT (Tab. IV; Fig. 3). In BW cattle, such association ($P \leq 0.02$) was found between allele $A_2$ of casein $\beta$ and genotype CT in the $\alpha_{s2}$-casein gene promoter (position –1084), whereas allele $A_1$ was associated with genotype CC (Tab. IV; Fig. 4).

Associations between polymorphism in the $\alpha_{s1}$-casein gene promoter and $\alpha_{s1}$-casein variants were not observed (Tab. V). The present results suggest the existence of intergenic haplotypes within the casein locus. Associations CT/A, TT/B and CT/A2, CC/A1 between genotypes of the $\alpha_{s2}$-casein gene promoter and casein $\alpha_{s1}$ and $\beta$, respectively, is obviously a result of strong linkage of these genes and their common localisation on cattle chromosome 6 [10, 31].

Associations were searched for between polymorphisms in promoter regions of the $\alpha_{s2}$- and $\alpha_{s1}$-casein genes and casein contents in milk. SDS-PAGE gels with stained milk proteins were scanned and densitometry was performed. Moreover, quantitative analyses were done by HPLC.

Differences in $\alpha_{s2}$-casein content were shown in milk derived from cows with different $\alpha_{s2}$-casein gene promoter variants at positions –1084 and –186 (Figs. 5A, 5C). As shown in Figure 5A, the –1084 CT genotype is associated with a higher amount of $\alpha_{s2}$-casein in milk, as compared with the CC genotype. The difference approached the significance of $P \leq 0.08$. Moreover, animals with the $\alpha_{s2}$-casein CT genotype (–1084)
Gene polymorphism and casein content in milk

had statistically more β-casein in their milk ($P \leq 0.01$; Fig. 5B). The second C/T transition in the αs2-casein gene promoter (position –186) was associated with αs2-casein content. Cows with the TT genotype had more αs2-casein in their milk than CT heterozygotes ($P \leq 0.07$; Fig. 5C).

Ehrmann et al. [6] detected differences in the expression of milk proteins in the mammary glands of cows with variants in the 5’-flanking regions of the β-lactoglobulin gene. Association has also been described between alleles in coding regions of the milk protein genes, and the levels of their expression in milk [32]. This is possibly caused by linkage between variants of coding and regulatory regions. Black and Bremel [2] suggested that the point mutation at position +15 in the α-lactalbumin gene (region coding for 5’ UTR) was a direct cause of the differences in milk production, and proposed this polymorphism as being a quantitative trait locus (QTL) on chromosome 5 of dairy cattle. Lum et al. [23] studied differential expression of two common allelic variants of β-lactoglobulin – A and B. The authors hypothesised that the G to C substitution at position –430 is a potential candidate for allele-specific regulation of β-lactoglobulin expression by interfering with binding of the AP-2 transcription factor. Kuss et al. [19] have reported the association of a single nucleotide polymorphism in the β-lactoglobulin gene promoter (G or C

**Table V.** Expected and observed frequencies of intragenic haplotypes within bovine casein locus.

<table>
<thead>
<tr>
<th>PR cattle; αs1-casein gene polymorphism (pos. –728)</th>
<th>Genotype of αs1-casein promoter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allele A</td>
<td>T/–</td>
</tr>
<tr>
<td>Observed</td>
<td>3</td>
</tr>
<tr>
<td>Expected</td>
<td>2.75</td>
</tr>
<tr>
<td>Allele B</td>
<td>–/–</td>
</tr>
<tr>
<td>Observed</td>
<td>11</td>
</tr>
<tr>
<td>Expected</td>
<td>12.25</td>
</tr>
</tbody>
</table>

**Figure 4.** Polyacrylamide gel electrophoresis (SDS-PAGE) of bovine caseins. Milk samples were obtained from Black-and-White cows with CT and CC genotypes of the αs2-casein gene promoter (position –1084). Associations between the 5’-flanking variants CT and CC of the αs2-casein gene and casein variants are shown. Homozygous αs1-casein genotype BB (lines 1–6), homozygous β-casein A1A1 (lines 2, 3, 5 and 6), and heterozygous A1A2 (lines 1 and 4). κ-, β- and α-cas are casein standards.
at position –435) with amounts of β-lactoglobulin in milk. They also demonstrated the association between single nucleotide polymorphism (SNP) in the β-lactoglobulin gene promoter, and amounts of α-lactalbumin and αs1-, β- and κ-caseins. The association between the SNP in the AP-2 binding site of the β-lactoglobulin gene and its gene product can be explained as a result of differences in protein binding capacity to DNA, and allele-specific differences in gene expression. Geldermann et al. [12] showed that, in a cell culture model, variant AA of the β-lactoglobulin gene promoter produced up to 3.5 times higher reporter gene expression than the BB genotype of β-lactoglobulin. Recently, Cardak et al. [5] have reported that genotypes of milk protein coding gene loci had a significant effect on contents, as well as yields, of corresponding milk proteins – αs1-, β- and κ-casein and β-lactoglobulin – in the milk of Holstein-Friesian and Simmental cows.

A number of groups [2, 28, 29] have hypothesised that inheritable variations in nucleotide sequences in gene regulatory elements might lead to differences in transcription rates, by decreasing or increasing the abundance of specific mRNAs. Thus, this would lead to differences in the amount of relevant proteins produced in cow’s milk.

As we showed previously, nucleotide substitutions in the 5′-flanking region of the bovine casein genes affect binding of transcription factors [30]. The C/T transition in the αs2-casein gene promoter appeared to influence the casein gene transcript levels. More αs2-casein gene transcripts were found in the RNA isolated from the mammary gland tissue genotyped as the αs2-casein CT genotype than the CC genotype. In this study, we showed an association of sequence nucleotide polymorphisms in the 5′-flanking region of the bovine αs2-casein gene with αs2- and β-casein contents in milk. In particular, cows carrying CT genotypes at the αs2-casein gene promoter contained more αs2-casein in milk. Thus, a mutation in the gene regulatory region might affect the levels of gene products in cow’s milk.

**Figure 5.** Analysis of αs2- and β-casein contents in cow’s milk with SDS-PAGE and HPLC. Quantitative differences in casein contents were expressed in arbitrary units. Six cows representing different casein promoter genotypes were used in every experiment; values represent the average (LSM ± SE). Milk was obtained from 3 Black-and-White cows with CC genotype and 3 with CT genotype in the αs2-casein gene promoter (position –1084) and the content of αs2-casein (A) or of β-casein (B) was measured (combined PAGE/HPLC data). (C) The αs2-casein content in milk from Polish Red cows with TT and CT genotypes of the αs2-casein gene promoter (position –186).
Gene polymorphism and casein content in milk. However, it must be stressed that our study only showed associations of certain SNPs with protein isoforms and content in milk, but provided no direct evidence for a causative relationship. It cannot be excluded that other mutations, not studied here, contributed to the differences seen in the levels of proteins in milk. Nevertheless, we hypothesise that these changes might be caused by alterations in affinity between transcription factors and promoters, these being the proteins principally involved in gene transcription regulation.

Acknowledgements: The authors would like to thank Prof. M. Żukowski of the Research Station, Popielno; and P. Kowal (VD) of KCHZ, Krakow, for organisation and kind assistance with collecting blood samples from Polish Red cattle. The authors also thank E. Siadkowska for excellent technical assistance. This study was partially funded by the State Committee for Scientific Research (KBN) grant No 6 P04B 019 17 and IGHZ project S.I.-2.1.

REFERENCES

[20] Leveziel H., Metenier L., Guerin G., Cullen P., Provot C., Bertaud M., Mercier J.C., Restriction fragment length polymorphism of ovine casein genes: close linkage between the $\alpha_s^1$, $\alpha_s^2$, $\beta$- and $\kappa$-casein loci, Anim. Genet. 22 (1991) 1–10.


To access this journal online: www.edpsciences.org