Associations between polymorphism within regulatory and coding fragments of bovine kappa-casein gene and milk performance traits*

S. Kamiński

Department of Animal Genetics, University of Warmia and Masuria in Olsztyn
Oczapowskiego 5, 10-718 Olsztyn, Poland

(Received 13 January 2000; accepted 15 July 2000)

ABSTRACT

In Black-and-White dairy cattle, the kappa-casein gene (CASK) appears in three of the most common genotypes, AA, AB, and BB. Apart from the coding sequence, several point mutations within the CASK gene promoter have been found. In this report, the hypothesis on the relationship between A/B polymorphism within exon IV and the promoter of CASK gene with milk performance traits is verified.

One hundred and twenty four Black-and-White cows held in one herd were genotyped for two polymorphisms called CASK (exon IV, A and B allele) and CASK-R (kappa-casein promoter, Dde I restriction site in position -385, P and M allele). Sixty-nine AA, 48 AB and 7 BB CASK cows, and 92 PP, 30 PM and 2 MM CASK-R cows were identified. Associations between single CASK genotypes, single CASK-R genotypes as well as intragenic haplotypes CASK/CASK-R and milk performance traits were investigated. Significant differences (P<0.01) were found only for milk protein percentage. Cows with the PM CASK-R genotype and AA/PM intragenic haplotype showed the highest protein percentage.

This is the first report showing the associations between mutations within the promoter of the CASK gene, intragenic CASK haplotypes, and milk protein percentage. Our results support the hypothesis suggested in the literature that mutations within regulatory sequences of the CASK gene may influence the expression of milk protein genes.

KEY WORDS: bovine, kappa-casein, promoter, Dde I polymorphism, protein percentage

* Supported by the Committee of Scientific Research, Grant No 5 P06D 04212
INTRODUCTION

The bovine CASK gene belongs to the cluster of four casein genes (alpha S1, alpha S2, beta and kappa) located within a 250 kb fragment on chromosome 6 (Grosclaude et al., 1972; Ferretti et al., 1990; Threadgill and Womack 1990; Lien et al., 1993; Rijnkels et al., 1997). The overall length of the CASK gene is close to 13 kb, but most of the sequences coding the mature kappa-casein protein are contained in the fourth exon (Alexander et al., 1988). So far, six variants of the CASK protein (A, B, C, E, G, F) have been described (Grosclaude et al., 1972; Erhardt, 1989, 1996; Chianese et al., 1991; Prinzenberg et al., 1996). Variant B of the CASK protein is commonly known to be favourable for cheese making (Jakob and Puhan, 1992; Walsh et al., 1998). Studies on the associations between CASK genotypes and milk protein content are not consistent across the literature. Most authors have shown positive effects of allele B on protein percentage and other milk performance traits (Ng-Kwai-Hang et al., 1990; Mao et al., 1992; Bovenhuis and Weller, 1994; Walawski et al., 1994), while others have found weak or even no associations (Ehramann et al., 1997), especially for CASK alleles as parts of the casein haplotype (Lien et al., 1995; Velmala et al., 1995). These confusing reports suggested searching for putative quantitative trait loci (QTL) for milk performance parameters not in the coding sequences of milk protein genes but rather within their 5' regulatory regions (Van Eenennaam and Medrano, 1991; Bovenhuis and Weller, 1994; Lien et al., 1995). In the meantime, over 15 point mutations within the promoter of the CASK gene were detected (Shild et al., 1994; Kamiński, 1996b). So far, no results on potential associations between regulatory sequences of CASK gene and milk performance traits have been published.

This report is an attempt to determine the importance of Dde I polymorphism in position -385 of the CASK promoter for milk performance traits in Black-and-White dairy cattle.

MATERIAL AND METHODS

One hundred and twenty four Polish Black-and-White cows originating from one herd were included in the analysis. Approximately 4 ml of blood were taken from each animal. Genomic DNA was isolated from leukocytes by the MasterPure DNA Purification Kit (Epicentre).

Two DNA fragments were amplified using PCR: a 214 bp fragment of the CASK promoter, called CASK-R, and a 273 bp fragment of CASK exon IV-intron IV, called CASK. PCR amplifications of 214 bp and 273 bp CASK fragments and the whole procedure of genotyping were previously described in detail (Kamiński, 1996b, 1998).
Milk performance traits collected from first complete 305-day lactations included: milk yield (kg), fat and protein content (%), fat and protein yield (kg). These data originated from breeding documentation and were provided by the owner of the animals.

The statistical analysis included: statistical characteristics of the experimental material (x – arithmetic mean, s – standard deviation), and inferences on the significance of differences between the genotypic or halotypic means for single trait were performed by using the F-test and Duncan test incorporated in the software package for Windows (STATSOFT, 1995). Single CASK genotype (AA, AB, BB), and single CASK-R genotype (PP, PM) as well as haplotype CASK/CASK-R (AA/PM, AA/PP, AB/PM, AB/PP, BB/PP) were included in the statistical model.

RESULTS

Each of 124 cows included in the analysis was genotyped for both loci, CASK and CASK-R. For CASK, 69 AA, 48 AB and 7 BB cows, and for CASK-R, 92 PP, 30 PM and 2 MM cows were identified. Complete results of statistical analysis are presented in Table 1.

<table>
<thead>
<tr>
<th>Genotypes/Haplotypes</th>
<th>Performance traits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>milk, kg</td>
</tr>
<tr>
<td><strong>CASK</strong></td>
<td></td>
</tr>
<tr>
<td>AA (n = 69)</td>
<td>6494 ± 1228</td>
</tr>
<tr>
<td>AB (n = 48)</td>
<td>6400 ± 899</td>
</tr>
<tr>
<td>BB (n = 7)</td>
<td>6603 ± 1297</td>
</tr>
<tr>
<td><strong>CASK-R</strong></td>
<td></td>
</tr>
<tr>
<td>PP (n = 92)</td>
<td>6473 ± 1097</td>
</tr>
<tr>
<td>PM (n = 30)</td>
<td>6494 ± 1120</td>
</tr>
<tr>
<td><strong>CASK/CASK-R</strong></td>
<td></td>
</tr>
<tr>
<td>AA/PM (n = 16)</td>
<td>6810 ± 1277</td>
</tr>
<tr>
<td>AA/PP (n = 51)</td>
<td>6430 ± 1192</td>
</tr>
<tr>
<td>AB/PM (n = 14)</td>
<td>6132 ± 809</td>
</tr>
<tr>
<td>AB/PP (n = 34)</td>
<td>6511 ± 922</td>
</tr>
<tr>
<td>BB/PP (n = 7)</td>
<td>6603 ± 1297</td>
</tr>
</tbody>
</table>

x – arithmetic mean, s – standard deviation of arithmetic mean, n – number of animals investigated, ^A^B P<0.01
Statistical analysis revealed no significant associations between single CASK, single CASK-R genotypes as well as intragenic haplotypes CASK/CASK-R for all of the milk performance traits except milk protein percentage. The cows with the CASK-R PM genotype produced milk with a significantly (P<0.01) higher protein percentage (x=3.45) than cows with the PP genotype (x=3.32).

The animals were also subdivided into 6 groups according to intragenic CASK/CASK-R haplotype: AA/PM, AA/PP, AB/PM, AB/PP and BB/PP. The results shown in Table 1 indicate that the effect of both polymorphisms was significant (P<0.01) only for milk protein percentage. The milk from cows with the AA/PM haplotype showed the highest protein percentage (x=3.50) and BB/PP the lowest (x=3.18). The AA/PM cows also produced more kilograms of protein (x=238.4) than the other haplotype groups, but the differences were not significant.

Three possible CASK/CASK-R combinations, AB/MM, BB/PM and BB/MM, were not identified within the investigated group. Two cows creating a group of CASK-R MM and AA/MM haplotype were excluded from the statistical analysis (too small number of animals).

DISCUSSION

The positive effect of the CASK B allele on cheese-making properties and cheese yield is commonly known among cattle geneticists and breeders. This phenomenon may be explained by the position of amino acid substitutions caused by mutations in exon IV of the CASK gene within the entire kappa-casein protein. The A and B variants differ in amino acid 136 and 148. At position 136 threonine (ACC) is changed into isoleucine (ATC) and at position 148, aspartic acid (GAT) is changed into alanine (GCT), for A and B, respectively (Alexander et al., 1988). These two point mutations are located relatively close to several glycosylation sites (amino acids at positions 131, 133, 135 or 141, 142) and probably affect the structure of the protein and glycosylation patterns (Fox, 1992; Creamer et al., 1998). It can therefore be thought that this region of the CASK protein is probably very important for the structure of the casein micelle, initiation of milk coagulation, formation of curd, and in effect, for cheese making. Indeed, in many studies it was found that the B allele is superior for coagulation time, curd firmness and cheese yield (Jakob and Puhan, 1992; Kamiński, 1996a; Walsh et al., 1998).

However, in current dairy breeding strategies, cheese-making properties are not included in the evaluation and pricing of milk. The most desirable traits now are milk protein percentage and protein yield. Therefore, searching for the association between CASK genotypes and milk protein output could have a significant practical value. Mao et al. (1992) in a relatively large population of over eleven thousand Holstein cows have showed that the B allele is recessive to the
A allele in respect to milk and protein yield, and dominant for protein percentage in milk.

Ng-Kwai-Hang et al. (1990) found for approximately 8000 cows that the replacement of the A by the B allele at the kappa-casein locus increased protein levels in milk by 0.08, 0.06 and 0.04%, for the first, second, and third lactation, respectively. Velmala et al. (1995), however, found no relationships between casein haplotypes (including CASK gene) and milk protein percentage and yield within Finnish Ayrshire cattle. Lien et al. (1995) found that not the B, but the A CASK allele, being a member of casein haplotype 5 (consisting of alpha S1 casein allele C, beta-casein allele A5, microsatellite in intron III of kappa-casein allele 14, and kappa-casein exon IV allele A) was associated with a higher protein yield in Norwegian cattle. These rather controversial results were verified by the precise measurement of kappa-casein protein in milk from cows of different CASK genotypes (Ehrmann et al., 1997). They found no correlation between CASK genotype AA, AB and BB and the synthesis of kappa-casein protein as well as total protein content in milk. These results are probably most reliable because they show the most direct relationships between the CASK gene and its product. They are also in accordance with the generally accepted opinion in molecular biology that the substitution of amino acids may change the chemical and physical properties of a protein, but not its quantity. The latest achievements of molecular biology clearly indicate that quantitative variance of gene expression in most eucaryotic genes is coded within the 5’ regulatory regions of the gene (Lewin, 1997; Papavassiliou, 1997).

In the bovine CASK gene, the 5’ flanking region has been sequenced up to -1073 bp (EMBL GeneBank acc. no M75887; Schild et al., 1994). The most important part of this region necessary for expression was limited to -552 (Adachi et al., 1996). The entire 5’ flanking region contains experimentally confirmed or computationally deduced common as well as specific DNA motifs engaged in regulation of CASK gene transcription (Groenen and van der Poel, 1994; Schild et al., 1994; Adachi et al., 1996; Malewski, 1998).

The choice of Dde I polymorphism in position -385 within the CASK promoter was made for two reasons: 1. this mutation is inside the regulatory element of a 132 bp fragment (-439/-308) which was defined by deletion mutation analysis in a mammary epithelial cell line as preferentially expressed during pregnancy (Adachi et al., 1996); 2. within the amplified 214 bp CASK promoter fragment, three point mutations were reported by Schild et al. (1994). Dde I polymorphism refers to point mutation (C/T) in position -385 which is located within the recognition site for the AP-2 transcription factor. Based on these results it can be hypothesized that the point mutation responsible for Dde I polymorphism may influence AP-2 binding ability to the CASK promoter. If C appears in position -385, the AP-2 recognition sequence gains higher homology to the AP-2 consensus sequence (CCCCAG-
GC) (Mitchell et al., 1987), then the Dde I site is disrupted (MM genotype). If T occurs in position -385, the AP-2 site has a lower homology to the AP-2 consensus sequence and the Dde I site is sustained (PP genotype) (Figure 1).

As shown in Table 1, cows with the PM genotype had a significantly higher protein percentage compared with PP cows. It can be concluded that the higher homology of AP-2 may result in a formation of a more stable and efficient transcriptional complex, increased transcription yield, CASK protein synthesis yield, and in effect, total protein in milk yield. Similar assumptions were proved experimentally for bovine beta-casein and beta-lactoglobulin. In vitro studies showed that mutations within the binding site of the mammary gland-specific nuclear factor (MGF) strongly affected transcriptional activity of the beta-casein gene (Schmitt-Ney et al., 1991). The mutation within binding site for AP-2 was shown to influence the binding ability of the AP-2 protein to its recognition sequence located in the promoter of the bovine beta-lactoglobulin gene (Lum et al., 1997).

The results presented in Table 1 also indicate that the protein percentage in milk may be associated with the CASK/CASK-R intragenic haplotype. The AA/PM haplotype is significantly superior, however, only to BB/PP.

Other milk performance traits (milk, kg; fat percent; fat and protein, kg) turned out not to be associated with either CASK-R or CASK genotypes or with any haplotype. These results should not be surprising because these traits are commonly considered polygenic.

The small number of animals does not allow calculation the genotype frequency and may serve only as preliminary observations. However, these results show the predominance of the P allele.

Haplotype is the particular combination of alleles in a defined region of some chromosome (Lewin, 1997). Intragenic haplotypes limit the combination of alleles to one gene. Wagner et al. (1994) have analyzed such haplotypes for bovine beta-lactoglobulin (LGB). These authors identified haplotypes created by the combination of the point mutation located in bovine beta-lactoglobulin exon IV (responsible for protein variants A and B) and different point mutations localized within the promoter of this gene. Among 60 analyzed cows, they found 80% animals with regular haplotypes (for example AA pairs to BB or AB to AB, for exon IV and promoter, respectively). The remaining 20% showed irregular haplotypes. The results presented in this paper confirmed the occurrence of this irregularity. For example, 2 of 7 CASK BB cows had the MM genotype, the others had different CASK genotypes. This phenomenon may be explained by a linkage disequilibrium phase, which could occur between mutations in the coding regions and in the regulatory sequences of the CASK gene. Linkage disequilibrium can result from reduced recombination in the region or from a founder effect, in which there has been insufficient time to reach equilibrium since one of the markers had been introduced into the population. Differences in the linkage disequilibrium phase...
Figure 1. The hypothetical explanation of the molecular importance of Dde I polymorphism in position -385 of CASK gene promoter.

The CASK-R PCR product is flanked by two primers: RKAP 7 and RKAP 8. The transcription factor AP-2 (activator protein -2) may bind to consensus sequence CCCCAGGC (Mitchell et al., 1987) with different homology caused by mutation Y in position -385 (Schild et al., 1994). This mutation may be recognised by restriction enzyme Dde I (Kamiński, 1996b). Binding sites for transcription factor AP-2 is underlined and bold. Other mutations within CASK promoter reported by Schild et al. (1994) are displayed as shaded letters.

Dde I endonuclease recognizes the sequence CTNAG which is changed by nucleotide substitution C/T (Y). P allele (plus) indicates that Dde I enzyme digests the PCR product, and M allele (minus) means that the Dde I enzyme does not digest. Genotype PM refers to heterozygote. Cows with PM genotype produce milk with statistically higher protein % and cows with PP genotype, lower. Cows with MM genotype were not included in the analysis, because of too small number of animals (only 2 cows).

The sequence between nucleotide -439 and -308 refers to regulatory element of CASK promoter shown to be preferentially expressed during pregnancy (Adachi et al., 1996). In the left side of the figure, the example of CASK-R genotyping by PCR-RFLP method is shown.
could vary across individuals, populations and breeds and might explain the confusing data on the importance of CASK protein variants for milk performance traits reported in the literature (Cowan et al., 1992).

The effect of Dde I polymorphism on milk protein content may be helpful in modifying the explanation of the differential expression of kappa-casein alleles on protein level (Van Eenennaam and Medrano, 1991) and mRNA level (Robitaille and Petitclerc, 1998).

The milk protein percentage is probably determined not only by milk protein genes, but also by other single loci; some of them have been mapped on chromosomes 1, 6, 9, 10 and 20 (Georges et al., 1995). Nevertheless, regulatory sequences of milk protein genes and other genes directly engaged in milk protein synthesis seem to be the most suitable gene candidates for milk protein output QTLs.

CONCLUSIONS

This is the first report showing the relationships between mutations within the promoter of the CASK gene and intragenic CASK haplotypes and milk protein percentage. Cows with the PM CASK-R genotype and AA/PM CASK/CASK-R haplotypes showed the highest protein percentage. In comparison with mutations within exon IV of the CASK gene, the Dde I polymorphisms in CASK-R may be more reliable as a DNA marker for protein percentage. These results support the hypothesis that mutations within regulatory sequences may influence the expression of milk protein genes. The hypothetical effect of Dde I polymorphism on different AP-2 binding affinity affected by a mutation in position -385 should be confirmed experimentally. Further studies are necessary to show the importance of Dde I polymorphism in mRNA coding CASK protein, kappa-casein protein level and milk protein content. Moreover, our observations should be confirmed on a larger population of cows representing all possible genotypes and haplotypes.

ACKNOWLEDGEMENTS

The author is thankful to Dr. Józef Jaźdzewski (OHZ Pawłowice) and Zbigniew Lach (OHZ Osięciny) for help in collecting the blood samples, Dr. Jolanta Rymkiewicz for help in statistical analysis as well as Anna Ruść and Elżbieta Wójcik for excellent laboratory assistance.
REFERENCES


STATSOFT INC., 1995. STATISTICA for WINDOWS (user’s manual)
STRESZCZENIE

Związek pomiędzy polimorfizmem w regulacyjnych i kodujących fragmentach genu bydlęcej kappa-kazeiny a cechami użytkowości mlecznej

U bydła czarno-białego gen kappa-kazeiny (CASK) występuje najczęściej w postaci trzech genotypów: AA, AB i BB warunkowanych przez mutacje w części kodującej genu CASK (exon IV). Poza sekwencją kodującą genu CASK znanych jest wiele mutacji punktowych w obrębie promotora genu CASK. Weryfikowano hipotezę o związkach pomiędzy polimorfizmem A/B w eksonie IV oraz w promotorze genu CASK a cechami użytkowości mlecznej.

U stu dwudziestu czterech krów rasy czarno-białej, utrzymywanych w jednym stadzie, określono dwa genotypy oznaczone jako CASK (ekson IV, allele A i B) i CASK-R (promotor kappa-kazeiny, polimorfizm dla enzymu restrykcyjnego Dde I w pozycji -385, allele P i M). Zidentyfikowano 69 AA, 48 AB i 7 BB krów dla locus CASK i 92 PP, 30 PM i 2 MM krów dla CASK-R. Badano związki pomiędzy pojedynczymi genotypami CASK, pojedynczymi genotypami CASK-R oraz intragenicznymi haplotypami CASK/CASK-R a cechami użytkowości mlecznej. Stwierdzono statystycznie istotne różnice (P<0,01) tylko dla procentowej zawartości białka w mleku. Jego zawartość była największa u krów z genotypem CASK-R PM i haplotypem AA/PM.

Jest to pierwsze doniesienie wykazujące związek pomiędzy mutacjami w promotorze genu CASK intragenicznymi haplotypami genu CASK a zawartością białka w mleku. Wyniki te potwierdzają hipotezę wysuwaną w literaturze, według której mutacje w obrębie sekwencji regulacyjnych genu CASK mogą wpływać na ekspresję tego genu.