



## Structural and functional analysis of the signaling lymphocytic activation molecule family 7 (*SLAMF7*) gene in response to infection with coagulase-negative and coagulase-positive staphylococci

A. Korwin-Kossakowska,<sup>1\*</sup> K. Ropka-Molik,<sup>2</sup> T. Ząbek,<sup>2</sup> T. Szmatoła,<sup>2,3</sup> P. Brzozowska,<sup>4</sup> B. Gralak,<sup>1</sup> E. Kawecka-Grochocka,<sup>5</sup> and E. Bagnicka<sup>4</sup>

<sup>1</sup>Department of Genomics and Biodiversity, Institute of Genetics and Animal Breeding, Postępu 36A str., 05-552 Jastrzębiec, Poland

<sup>2</sup>Department of Animal Molecular Biology, National Research Institute of Animal Production, Balice, Krakowska 1 str., 32-083 Balice, Poland

<sup>3</sup>University Centre of Veterinary Medicine, University of Agriculture in Krakow, Mickiewicza 24/28, 30-059 Krakow, Poland

<sup>4</sup>Department of Animal Improvement and Nutrigenomics, Institute of Genetics and Animal Breeding, Postępu 36A str., 05-552 Jastrzębiec, Poland

<sup>5</sup>Department of Preclinical Sciences, Faculty of Veterinary Medicine, Warsaw University of Life Sciences (SGGW), Nowoursynowska St. 166, Warsaw, Poland

### ABSTRACT

Splice variants of the signaling lymphocytic activation molecule family 7 (*SLAMF7*) gene have been identified, and differences in the expression of this gene have been demonstrated at the mRNA level in the mammary glands of healthy and mastitis-infected dairy cows. At the same time, significant associations have been found between a deletion in the *SLAMF7* gene exon, the occurrence of different splice variants, and the occurrence of mastitis in one group of dairy cows. An expression study was conducted on 40 Polish Holstein-Friesian dairy cows of the Black and White variety (group I). Milk samples were taken for microbiological analysis 2 d before slaughter and examined for the presence of bacteria. Immediately after slaughter, mammary tissue samples were taken and divided into 3 groups according to the health status of the mammary gland: healthy (without pathogenic bacteria in milk), coagulase-negative staphylococci (CNS), and coagulase-positive staphylococci (CPS). Based on different *SLAMF7* gene DNA fragments, 2 alternative variants of this gene (V1 and V2) and complete gene expression were identified. Separate analyses performed for each isoform showed that the health status of the cow was strongly associated with the expression level of individual variants. The highest expression was detected for the *SLAMF7* complete amplicon in healthy cows, and in the CNS and CPS cows the expression of this variant was also higher than V1 and V2. Sanger sequencing was applied to detect the polymorphism/indel variant in the second

exon of the *SLAMF7* gene probably having the greatest effect on the protein structure and function of *SLAMF7*. Two genotypes were detected: AA (wild-type) and AB (insertion A). In healthy cows, the frequency of homozygotes AA was higher than the heterozygotes, whereas in the infected animals, the genotypic distribution was the opposite. An association analysis between the identified polymorphism and production traits—including somatic cell count, as well as lactose, protein, and casein content and yield as indicators of subclinical mastitis occurrence—was performed on the group II cows (166 Polish Holstein-Friesian dairy cows). Unfortunately, due to the low number of AB animals, no relationship was demonstrated between genotype in the second exon and the health status of cows. Additionally, the difference in the percentage of *SLAMF7*-targeted DNA methylation between the groups of animals was not significant, with an average of ~66 to 68%.

**Key words:** dairy cow mastitis, *SLAMF7*, gene expression, alternative splicing, methylation

### INTRODUCTION

Mastitis is an infectious disease of the mammary gland caused by the invasion and colonization of pathogens and is considered one of the most economically important diseases in the dairy industry. In addition to sanitary care, one promising approach to reducing the problems caused by infectious and contagious diseases is the selection of animals that are resistant to disease and the incorporation of this trait into herds. Within this context, studies aimed at better understanding the biological processes involved in determining resistance to disease are fundamental to overcoming these problems and to developing technological solutions. In the past few decades, research has focused on determin-

Received August 4, 2019.

Accepted April 28, 2020.

\*Corresponding author: [a.kossakowska@ighz.pl](mailto:a.kossakowska@ighz.pl)

ing the genetic basis of mastitis/mammary infection (Kościuczuk et al., 2014). Resistance to mastitis is a complex trait, and the genes involved in the immune response have been considered strong candidate markers for this trait. A vast number of association studies have been conducted in the past decade, revealing some of the SNP that serve as markers of resistance to mastitis (Pawlik et al., 2015a,b).

Of particular interest is the study of gene expression. Regulating gene expression is essential for linking genotypes with phenotypes, which ultimately drive biological processes. The use of RNA sequencing as a screening method has provided valuable insights into various aspects of transcriptomics, which are essential for interpreting the functional elements of the genome. Medrano et al. (2014), using the RNA sequencing technique, compared milk transcriptomes and found a total of 3,566 genes that were differentially expressed in the somatic cells of healthy and mastitic milk. Moreover, Kościuczuk et al. (2017), in a microarray analysis, found between 418 and 2,200 differentially expressed genes in mastitic mammary gland parenchyma samples, depending on the type of bacteria and parity. One of the main regulatory processes during gene expression is alternative splicing (**AS**), in which different mature transcripts are produced from the same primary RNA sequence. An increasing amount of research is showing that AS plays a crucial role in the development and variability of species and individuals. Genome-wide analyses of this phenomenon indicate that most human genes have alternative splice variants, suggesting that AS is one of the most significant components of the functional complexity of the human genome (Modrek and Lee, 2002); AS is also a feature of numerous diseases (Tazi et al., 2009). The study conducted by Pan et al. (2008) was one of the first to use this technique on a genome-wide scale; the authors detected AS events in about 20% of multiexon genes. López-Bigas et al. (2005) showed that even for a hereditary disease caused by a single mutation, in most cases, the mutation affected splicing of the gene rather than directly affecting the translated gene sequence. All of these findings indicate that the ability to detect AS is vital to further enhancing our knowledge about biological processes.

Of the 21,755 bovine genes, 4,567 (21%) are known to be alternatively spliced (Yang et al., 2012). One of the genes that undergoes AS is the signaling lymphocyte activation molecule family 7 (*SLAMF7*) gene. The *SLAMF* is in the family group of immune cell-specific receptors that can regulate the function of several immune cell types. The *SLAMF7* is a member of the *SLAM* family, which contains an immunoglobulin-like domain and is expressed in many tissues. The bovine *SLAMF7* gene comprises 7 exons and encodes a protein

with 336 AA. Ju et al. (2012) characterized 3 novel splice variants of bovine *SLAMF7*, which are expressed in the mammary tissues of cattle.

The central epigenetic modification in most eukaryote genomes is DNA methylation. This process plays an important role in pre-transcriptional regulation. The epigenetic mark links the interactions among genetic, epigenetic, and environmental factors, and it is associated with many biological processes. The DNA methylation map is also important to understand the changes in DNA methylation during disease progression (Song et al., 2016).

Wang et al. (2016) hypothesized that the differences induced by the AS of genes can play a role in the cow's immune system regulation: processing complex information to initiate a host response to invading pathogens. Therefore, it is important to distinguish the transcriptomic characteristics and differential patterns of AS in the bovine mammary glands of healthy cows versus mastitic cows naturally infected with *Staphylococcus aureus*. Based on the results of Ju et al. (2012), the hypotheses for the present work were that different splice variants of the *SLAMF7* gene exist and that they are expressed differentially in the mammary gland tissues of healthy and mastitis-infected dairy cows. Also, we hypothesized that *SLAMF7* is a novel DNA methylation target gene that is strongly correlated with susceptibility to mastitis in Holstein-Friesian cows. The aims of the present study were to (1) describe different splice variants and demonstrate their relationship to mastitis infection in dairy cows, (2) analyze the association between a single nucleotide deletion in an exon of the *SLAMF7* gene and occurrence of mastitis in dairy cows, and (3) assess the relevance of methylation in CG content for *SLAMF7* differential expression observed in healthy samples and samples infected with coagulase-positive staphylococci (**CPS**) and **CNS**.

## MATERIALS AND METHODS

### Animals Used for *SLAMF7* Expression Analysis

The analysis of *SLAMF7* expression levels was conducted in 40 Polish Holstein-Friesian dairy cows of the Black and White variety. The animals were born in and maintained at the Experimental Farm of the Institute of Genetics and Animal Breeding in Jastrzębiec, Poland. The animals were kept under identical conditions in a loose barn with free access to water. They were fed the same TMR diet ad libitum, consisting of corn silage (75%), concentrates (20%), and hay (5%), supplemented with a mineral and vitamin mixture, according to the system developed by the Institut National de la Recherche Agronomique (INRA) of France and

adopted by the National Research Institute of Animal Production (IZ PIB), Poland (Strzetelski and Śliwiński, 2009). The selected cows were between their first and fourth lactations. The animals were slaughtered at the end of lactation (286 d, SD = 27) due to reproduction problems or chronic, usually asymptomatic, recurrent, and incurable mammary gland inflammation. They were slaughtered at least 1 mo after the last antibiotic administration at the registered slaughterhouse under constant monitoring by authorized authorities. All procedures involving animals were performed in accordance with the Guiding Principles for the Care and Use of Research Animals and were approved by the III Local Ethics Commission (Warsaw University of Life Sciences; Permission No. 15/2010).

### Sample Collection

Milk samples for microbiological analysis were taken from each quarter of the udder 2 d before slaughter and were examined for the presence of bacteria. Samples were collected aseptically according to the methods recommended by the International Dairy Federation. Ten microliters of the “first milk” was streaked on agar with 5% sheep blood (Columbia, bioMérieux, Craponne, France; Quinn et al., 2002) and Chapman-Mannitol Salt Agar (MSA, bioMérieux) and incubated at 37°C for 18 to 24 h. The isolates were phenotypically evaluated for colony morphology, cell morphology, and biochemical properties. The production of coagulase by staphylococci was detected using a tube test with rabbit plasma. Additionally, *S. aureus* strains were identified using a Slidex Staph-Kit (bioMérieux). Immediately after slaughter, mammary tissue samples were taken from deep secretory parts of the gland (the parenchyma) of each quarter. The samples were rinsed in PBS to remove, as accurately as possible, the milk and blood.

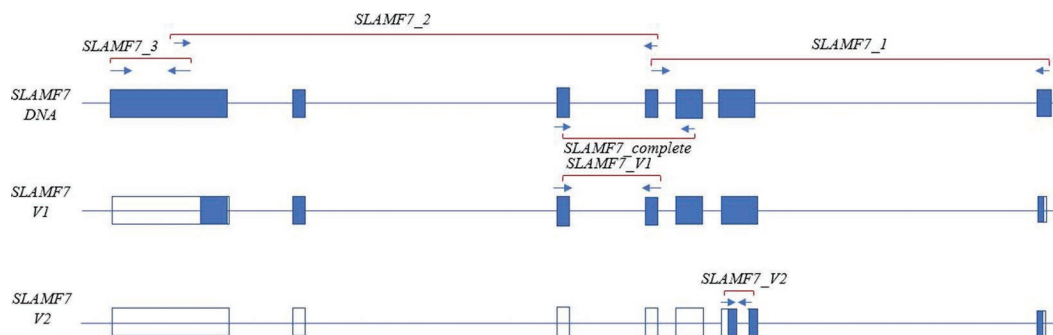
Then the samples were frozen in liquid nitrogen and stored at −80°C until further analysis.

### Gene Expression Measurements

The RNA from all samples was isolated using a Direct-zol RNA Mini Prep (Zymo Research, Irvine, CA). The quality and quantity of the obtained RNA were evaluated using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA), as well as a TapeStation 2200 system with RNAScreen Tape (Agilent Technologies, Santa Clara, CA). The 400 ng of total RNA was reverse transcribed to cDNA using a Maxima First Strand cDNA Synthesis Kit for RT-qPCR (Thermo Fisher Scientific) according to the protocol.

The primers for the PCR products were designed for the complete transcript (ENSBTAT00000001585.6; ENSBTAG00000001197), and for the V1 (ENSBTAT00000001585.6) and V2 (ENSBTAT000000079658.1) transcripts of the *SLAMF7* gene and according to the reference sequences for individual genes using Primer 3 (v. 0.4.0; <http://primer3.ut.ee/>) freeware (Figure 1). It was not possible to design specific primers for V3 transcript (ENSBTAT000000084603.1). Previous reports indicate a very low expression level of the V3 variant (Ju et al., 2012) in relation to whole *SLAMF7* gene; thus, its expression was evaluated in complete gene expression linking all 3 variants (*SLAMF7*\_complete gene). Detailed information of the methodology for all analyzed transcripts is shown in Table 1.

The specificity of the obtained PCR products was checked using PCR with AmpliTaq Gold 360 Master Mix (Applied Biosystems, Thermo Fisher Scientific) according to the manufacturer's protocol and on a 3% agarose gel electrophoresis.



**Figure 1.** The exact localization of the analyzed amplicons of signaling lymphocyte activation molecule family 7 gene (*SLAMF7*). The amplicons *SLAMF7*\_1, *SLAMF7*\_2, and *SLAMF7*\_3 represent fragments of the *SLAMF7* gene used for the sequence analysis. *SLAMF7* V1 and *SLAMF7* V2 represent amplicons used for gene expression measurements. *SLAMF7*\_complete represents amplicons used to evaluate whole gene expression (for both transcripts together). The blue boxes represent coding sequence in each transcript, and the white boxes represent the exons belonging to untranslated regions of transcripts (5' and 3' untranslated regions).

The expression levels of the analyzed transcripts were evaluated using Sensitive RT HS-PCR Mix EvaGreen (A&A Biotechnology, Gdynia, Poland) on a QuantStudio 7 Flex instrument (Applied Biosystems, Thermo Fisher Scientific), according to the following temperature profile: initial denaturation 95°C for 5 min and 45 cycles of 95°C for 15 s, 57°C for 30 s, and 72°C for 30 s. The reaction was carried out in 3 replications for each sample and for each amplicon. Two genes were used as endogenous controls: *RPS15* and  $\beta$ -actin (*ACTB*; Bionaz and Loo, 2007). Expression was calculated using the delta-delta Ct method according to Pfaffl (2001). The efficiency of the real-time PCR reactions for each amplicon was defined by using the standard curve method. The normalization factor was calculated based on the geometric mean of the normalized quantity of the endogenous genes used.

### Methylation Analysis

Converted DNA samples from tissue sections representing all investigated groups were subjected to *SLAMF7* targeted methylation analysis. The DNA was prepared using a commercial kit (A&A Biotechnology, Gdynia, Poland) and bisulfite converted using an EpiTect Bisulfite Conversion kit (Qiagen, Hilden, Germany). During this process, all unmethylated cytosines are converted to uracine, which is amplified as thymine during PCR. Thereafter, the presence of cytosines or thymines, or both, next to guanines after cycle sequenc-

ing of the PCR amplicons determines the presence of complete, partial, or lack of methylation at inspected CpG sites. The bisulfite primers were designed in a region located upstream of the *SLAMF7* gene using Methyl Primer Express software (Life Technologies, Warsaw, Poland; Table 1). The investigated region spanned 6 CpG sites residing –132 to –736 bp from the transcriptional start site of *SLAMF7* (coordinates in Table 1). Converted DNA was amplified with the use of Hot Start Taq DNA polymerase (Qiagen, Hilden, Germany), applying a 2-step PCR protocol of 5 cycles of 62°C and 35 cycles of 60°C of annealing temperature. The obtained amplicons were purified and subjected to Sanger sequencing using bisulfite primers (Table 1). Bisulfite sequencing reads were then used to calculate the relative methylation percent based on the method proposed by Leakey et al. (2008). To perform this calculation, 4 dye-trace value outputs generated by BioEdit software (Hall, 1999) were used to read and calculate the primary and secondary peaks, representing cytosines or thymines from the electrophoregrams.

### Analysis of the Coding Sequence Contents and Untranslated Regions

All coding sequences and 5' and 3' untranslated regions (**UTR**) of the bovine *SLAMF7* gene were screened to find the most important polymorphism(s) that may affect expression levels or protein function. Based on an mRNA reference (ENSBTAT00000001585.6), the cod-

**Table 1.** Detailed information about the reference sequences used and the method designed

Gene/isoform	Primers (F = forward; R = reverse)	PCR (bp)	E qPCR <sup>1</sup>	Accession number/coordinates
Gene expression analysis				
<i>SLAMF7_V1</i> <sup>2</sup>	F: GCTCCCTACGTGGTCTCTCTA R: ATGAGTGTCCAGTTCCTTCT	117	1.97	ENSBTAT00000001585.6
<i>SLAMF7_V2</i> <sup>2</sup>	F: AGCTGGTTGGTGCCATTGCCAA R: CCACACGGTAGATACCTGAGTCG	154	2	ENSBTAT00000079658.1
<i>SLAMF7_Complete gene</i> <sup>2</sup>	F: TGGAGGTGGGAGAAAAGTGA R: CTTCTCTTCAATGATCTCTTCAC	135	1.86	ENSBTAT00000001585.6; ENSBTAG00000001197;
<i>RPS15</i>	F: ACAACGGCAAGACCTTCAAC R: AGGTTACTTGAGGGGGATG	201	2	ENSBTAG00000019718
<i>ACTB</i>	F: GCACAATGAAGATCAAGATCATC R: CTAACAGTCCGCCTAGAAGCA	173	1.98	ENSBTAG00000026199
<i>SLAMF7_1</i>	F1: CGGTGGTTAACACCCAGAAA R1: GGAACAGAGCAGGAAGAACG	745		ENSBTAT00000001585.6 ENSBTAG00000001197 <sup>3</sup>
<i>SLAMF7_2</i>	F2: GCTCCCTACGTGGTCTCTCTA R2: GCTTCTAAGCCCCCTTGCTTT	689		
<i>SLAMF7_3</i>	F3: TCAAAGCAAGGGGCTTAGAA R3: TACACTTCCCCAATGCACAA	581		
BS-PCR/Sanger cycle sequencing <sup>4</sup>	F: GGATTTTTTTAGAGATTTAGAAGAGAGA R: TCCAATACTTTAACCACCTCATAT	604		8,985,883–8,986,487 bp of chromosome 3 of cow UMD3.1 genomic sequence

<sup>1</sup>E qPCR = efficiency of quantitative PCR reaction.

<sup>2</sup>Primers designed by Ju et al. (2012).

<sup>3</sup>These accession numbers apply to *SLAMF7\_1*, *SLAMF7\_2*, and *SLAMF7\_3*.

<sup>4</sup>BS-PCR = bisulfite PCR.



ing sequence and UTR were divided into 3 amplicons spanning 1,949 nucleotides from all 1,982 bp, including *SLAMF7* mRNA (Figure 1, Table 1). The whole coding sequence was sequenced for all animals used in the *SLAMF7* gene expression analysis ( $n = 40$ ). The PCR products were amplified with AmpliTaq Gold 360 Master Mix (Applied Biosystems, Thermo Fisher Scientific) according to the manufacturer's instructions and with the use of cDNA as a matrix. Next, the obtained products were purified using EPPiC–Enzymatic Post-PCR Immediate Cleanup (A&A Biotechnology) according to the manufacturer's instructions. Sanger sequencing was performed on a 3500xL Genetic Analyzer (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA) with the use of a BigDye Terminator v3.1 Cycle Sequencing Kit and a BigDye XTerminator Purification Kit (Applied Biosystems, Thermo Fisher Scientific). The data were analyzed using 3500/3500xL Data Collection Software, v3.1 (Applied Biosystems, Thermo Fisher Scientific), and FinchTV v1.4.0 software (Geospiza Inc., Seattle, WA).

### Genotyping of the Most Interesting Polymorphism

Based on the strong effect of one identified polymorphism on the protein sequence and, as a result, on protein function, this polymorphism was chosen and a larger group of animals was genotyped. The animals were divided into 2 groups: group I consisted of all 40 cows used in the gene expression experiment, and group II consisted of 167 Polish Holstein-Friesian dairy cows of the Black and White variety from 3 herds in central Poland ( $n = 59, 59$ , and  $49$  per herd). All herds were evaluated for milk yield (AT4 method). Extraction of DNA was performed: in group I, the biological material for DNA extraction was whole blood, whereas in the group II, the material was hair follicles. Isolation of DNA was performed using a suitable DNA isolation kit (Macherey Nagel, Düren, Germany) depending on the material. Information on the health status of the mammary tissue of the animals in groups I and II was collected. The detected indel polymorphism was genotyped in the animals of these 2 groups using the Sanger sequencing method, according to the procedure described previously and using primers dedicated for transcript *SLAMF7\_V2* (Table 1). For the selected polymorphism, an association analysis with mastitis occurrence was performed for group II.

### Statistical Analysis

The normality of the distribution of expression data was estimated using the Shapiro-Wilk test, and differences between the groups were analyzed using

the Kruskal-Wallis test. The relative quantities of the transcript forms were compared among the biological groups: healthy cows (without pathogenic bacteria in milk) and both mastitic cow groups: CNS and CPS. The second analysis considered the differences between the expression levels in the biological groups for each *SLAMF7* splice variant.

Production traits were obtained from phenotypic data from the Polish Federation of Cattle Breeders and Dairy Farmers, including SCC, and lactose, protein, and casein content and yield as indicators of subclinical mastitis in group II. The association between the identified polymorphism and 30 production traits (Supplemental Table S1; <https://doi.org/10.3168/jds.2019-17398>) was calculated by a MIXED procedure package for a one-trait repeatability model (SAS/STAT9.4, v.13.1, 2002–2012, SAS Institute Inc., Cary, NC) with Bonferroni correction for multiple testing, as a preliminary analysis based on 167 cows. In the ANOVA, the association between the studied indel polymorphism and 3 traits—daily milk yield, lactose yield, and casein content—turned out to be significant; therefore, these were included in the final step of the analysis. The one-trait repeatability animal model was applied for estimating the genetic parameters of these traits, along with the best linear unbiased estimators for fixed effects using the average information REML method with the DMU4 package (Denmark) of Madsen and Jensen (2000). There were 12 yr of calving and 4 seasons (winter: December, January, and February; spring: March, April, and May; summer: June, July, and August; and autumn: September, October, and November). Finally, 44 classes of the herd-year-season of calving interaction, and 70 classes of the herd-year-month of milking interactions were created. Four classes of parity were established (with each parity as a separate class and with the 4th covering lactations later than the 4th). The effect of the lactation stage was modeled using Legendre polynomials nested in lactation (Brotherstone et al., 2000) to the 5th power. The pedigree file consisted of information on 2,057 sires and 4,634 dams. The average inbred and co-ancestry coefficients were 12% and 4%, respectively. A multiple comparison Duncan's test with Bonferroni adjustment was performed to establish the differences between the gene variants. The genetic relationship of the animals was not considered.

The following model was used in the statistical analyses:

$$y_{ijklmn} = \alpha + a_i + p_i + hys_j + HYM_k + P_l + IND_m + lpDIM_{ijklmn} + lpDIM^2_{ijklmn} + lpDIM^3_{ijklmn} + lpDIM^4_{ijklmn} + lpDIM^5_{ijklmn} + e_{ijklmn},$$

where  $y_{ijklmn}$  = studied traits,  $\alpha$  = intercept for trait,  $a_i$  = random additive direct genetic effect of  $i$ th animal,  $p_i$  = random effect of the animal's permanent environment effect,  $hys_j$  = the random effect of the herd-year-season of calving,  $HYM_k$  = the fixed effect of the herd-year-month of milking,  $P_l$  = the fixed effect of the parity,  $IND_m$  = the fixed effect of  $m$ th gene variant,  $lpDIM^5_{ijklmn}$  = fixed regression effect on the Legendre polynomials up to the 5th degree of standardized days; and  $e_{ijklmn}$  = random residual.

The obtained values of the percentage of methylation (PM) were the basis of statistical pairwise comparisons among the 3 groups (2 types of bacteria and healthy tissue). A Wilcoxon rank-sum test was applied to assess the level of significance of the pairwise PM differences.

## RESULTS

### Performance of Cows Used in the Study

The dairy cows used in the first part of the study (i.e., the *SLAMF7* expression analysis) were in their last stage of lactation, which is characterized by low milk yield (below 10 kg of milk per day). The milk infected with CPS contained approximately  $5 \times 10^6$  somatic cells (SD =  $3 \times 10^6$ ), 3.70% fat (SD = 1.49), 3.64% total protein (SD = 0.51), 2.24% casein (SD = 0.66), 3.85% lactose (SD = 0.87), and 0.10% citric acid (SD = 0.05), whereas the milk infected with CNS contained  $2.5 \times 10^6$  (SD =  $3.5 \times 10^6$ ), 3.98% (SD = 1.44), 3.70% (SD = 0.49), 2.25% (SD = 0.51), 4.27% (SD = 0.90), and 0.11% (SD = 0.04) of somatic cells, fat, protein, casein, lactose, and citric acid, respectively. Milk derived from healthy quarters contained approximately  $5 \times 10^4$  somatic cells (SD =  $2 \times 10^4$ ), 3.53% fat (SD = 0.92), 3.22% protein (SD = 0.31), 2.34% casein (SD = 0.31), 4.54% lactose (SD = 0.25), and 0.14% citric acid (SD = 0.03). The higher content of milk somatic cells and total protein and the lower content of casein, lactose, and citric acid in the milk of the CPS and CNS groups versus the healthy (H) group clearly indicated the pathological changes in the udders of the cows belonging to the first 2 groups.

The second part of the study (i.e., the association analysis) was conducted based on phenotypic information for 30 traits. The average values of the traits and their standard deviations are shown in Supplemental Table S1 (<https://doi.org/10.3168/jds.2019-17398>).

### Results of Milk Microbiological Analysis and Sample Numbers

The tissue samples were divided into 3 groups according to the health of the mammary gland, which

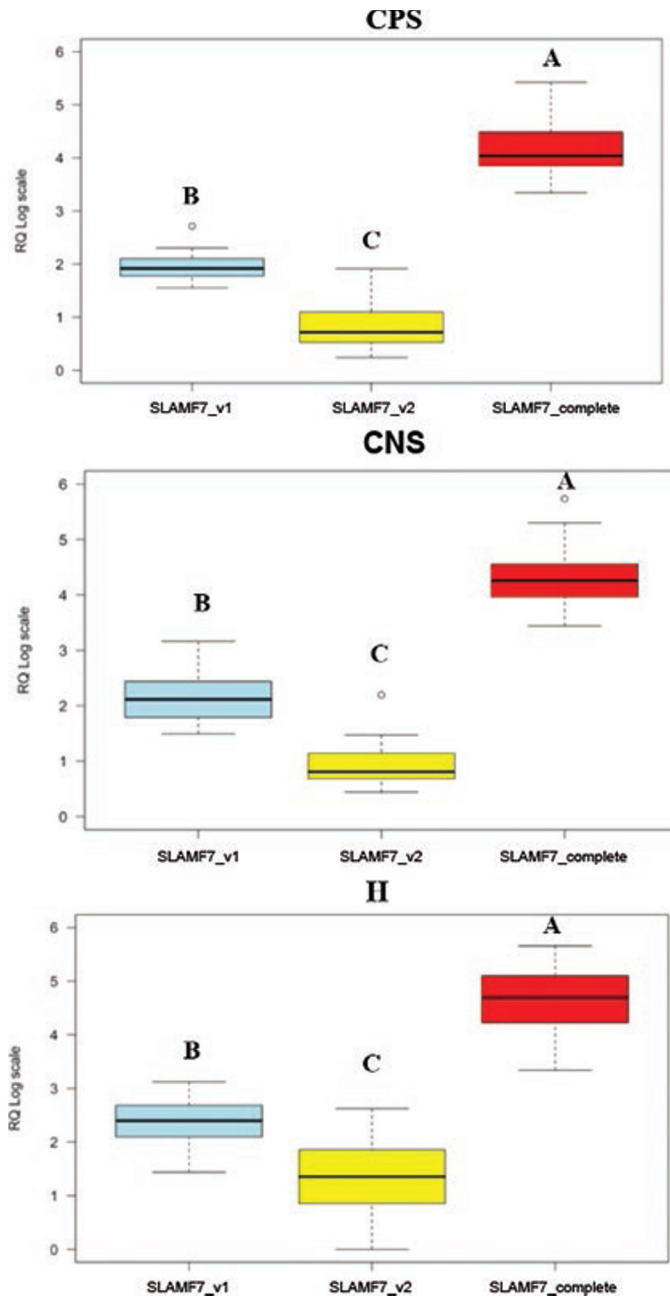
was based on an analysis of the microbiological status of the quarter milk. A prior statistical analysis (variance analysis with GLM procedure in SAS/STAT software) had revealed no differences in gene expression level among the lactations; therefore, the number of lactations was excluded from further analysis. Thus, 3 groups of samples were used (1) the healthy (H or control) group—cows with no pathogenic bacteria in their milk, (2) cows with subclinical mastitis caused by CPS (*S. aureus* mainly), and (3) cows with subclinical mastitis caused by CNS. Of a total of 160 samples (40 slaughtered cows  $\times$  4 quarters of the mammary gland), only part of the samples was selected on the basis of the milk microbiological examination. No more than 2 quarter samples were taken from one cow for the analysis. Samples of the mammary gland infected with more than one type of bacteria were excluded from the analysis. There were 6 samples in the H group, 22 samples in the CPS group, and 10 samples in the CNS group.

### SLAMF7 Expression Analysis

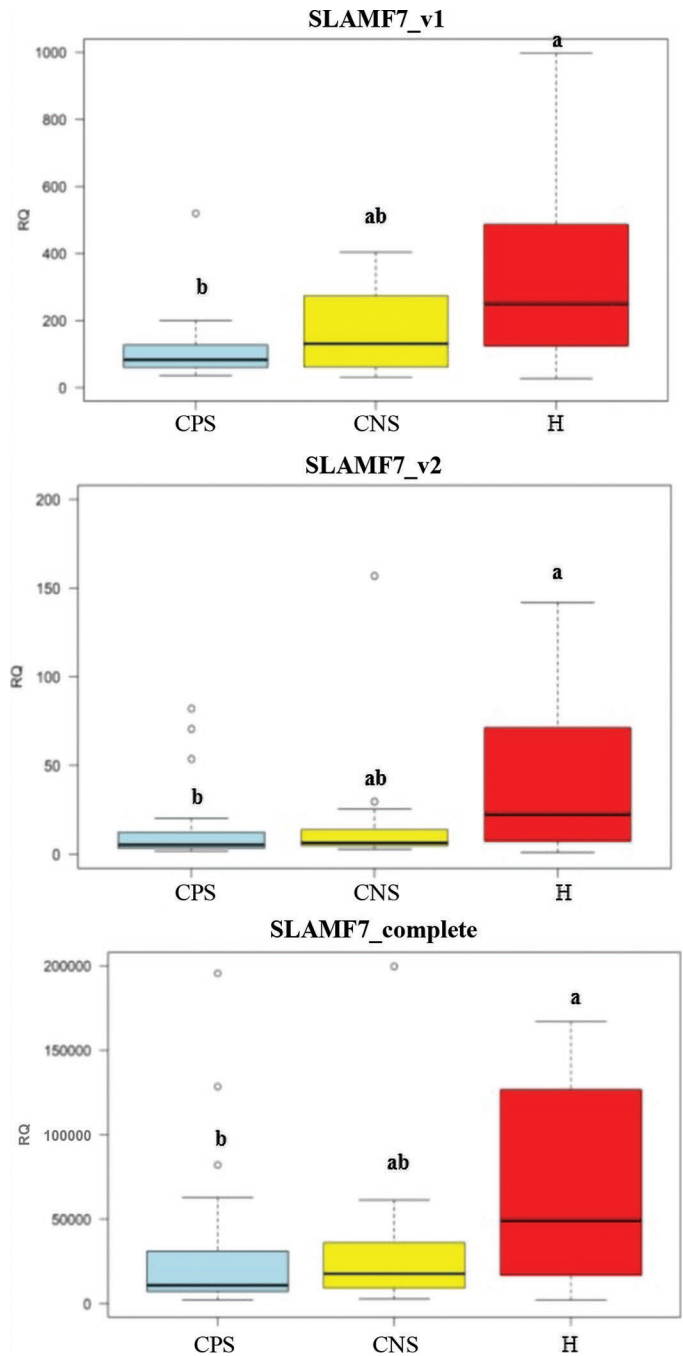
To check the specificity of the amplified *SLAMF7* transcript variants, Sanger sequencing was performed. On this basis, the presence of 2 alternative variants were confirmed: *SLAMF7\_V1*, *SLAMF7\_V2*, and sequence included whole *SLAMF7* gene expression (all transcripts together V1, V2, and not amplified separately V3—ENSBTAT00000084603.1), *SLAMF7\_complete* was identified (Supplemental Figure S1; <https://doi.org/10.3168/jds.2019-17398>). Selective amplification of these alternative variants was achieved by reverse transcription-quantitative PCR using primers specific to each transcript (Table 1) and described previously by Ju et al. (2012).

Relative quantification of the analyzed variants in mammary gland tissues was performed by reverse transcription-quantitative PCR amplification, with the conditions optimized for each of the *SLAMF7* gene isoforms (Table 1). The analysis of each *SLAMF7* isoform in each biological group confirmed the fact that despite mastitis occurrence, the significantly ( $P < 0.01$ ) highest expression level was detected for the *SLAMF7\_complete* amplicon. The lowest expression level was observed for the *SLAMF7\_V2* isoform, whereas an intermediate, yet significantly different, expression level value was obtained for the *SLAMF7\_V1* isoform (Figure 2).

The separate analyses performed for each isoform showed that health status was strongly associated with the expression levels of individual variants. Differences were observed between the H group and the mastitic mammary gland groups, both CPS and CNS ( $P < 0.05$ ; Figure 3). The expression levels of each isoform showed intermediate values for the tissue affected by CNS. The



**Figure 2.** Differences between the expression levels [shown as relative quantity values (RQ)] of the individual splice variants V1, V2, and complete within the CPS, CNS, and H groups. Different letters (A–C) indicate the differences at  $P < 0.01$ . CPS = coagulase-positive staphylococci group; H = healthy group (i.e., no pathogenic bacteria present in milk). Box plots show data distribution within each group, midline indicating median; whiskers represent 25% of data variability; dots represent outliers.



**Figure 3.** Differences in the expression levels [shown as relative quantity values (RQ)] of the individual splice variants in all 3 groups of animals: CPS, CNS, and H. Different letters (a,b) indicate the differences at  $P < 0.05$ . CPS = coagulase-positive staphylococci group; H = healthy group (i.e., no pathogenic bacteria present in milk). Box plots show data distribution within each group, midline indicating median; whiskers represent 25% of data variability; dots represent outliers.

**Table 2.** Identified polymorphisms with signaling lymphocyte activation molecule family 7 (*SLAMF7*) gene locus in the analyzed Holstein-Friesian cows

Gene region	SNP	Nomenclature	Variant	Amplicon
Exon 2	A>C	ENSBTAP00000001585.6:p.LyS64Thr	Missense variant	SLAMF7_1
Exon 2	Ins A	ENSBTAP00000001585.6:c.233insA	Frame shift variant	SLAMF7_1
Exon 2	T>C	ENSBTAP00000001585.6:p.Leu80Met rs110114287	Missense variant	SLAMF7_1
Exon 5	G>A	ENSBTAT00000001585.6:c.878G>A	Missense variant; splice region variant	SLAMF7_2
3' UTR <sup>1</sup>	Ins A	ENSBTAP00000001585.6:p.Gly293Asp rs110442050	3 prime UTR variant	SLAMF7_2
3' UTR	C>G	ENSBTAT00000001585.6:c.*439C>G rs110239497	3 prime UTR variant	SLAMF7_3

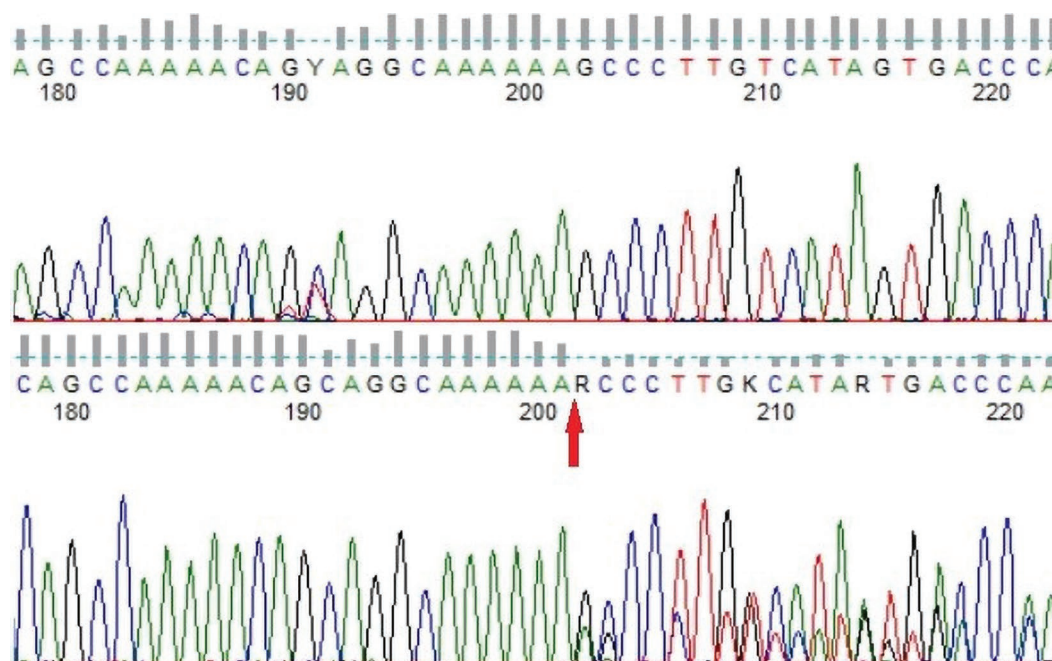
<sup>1</sup>UTR = untranslated region.

H group exhibited the greatest variation in expression levels of all tested isoforms.

### Identification of Polymorphisms

An analysis of the whole coding sequence and UTR identified 6 polymorphisms in the investigated Holstein-Friesian cow population (Table 2). Three polymorphisms were localized in the second exon, 1 in the fifth exon, and 2 in the 3' UTR. The second exon, together with third, encodes the Ig-like domain in 2 *SLAMF7* splice variants (ENSBTAT000000084603.1, the third *SLAMF7* splice variant annotated in Ensembl database, and ENSBTAT00000001585.6, which corresponds to our variant *SLAMF7\_V1*). Thus, for the next

genotyping analysis, we selected the frame shift variant localized within the second exon, which probably has the greatest effect on the protein structure and its function ENSBTAP00000001585.6:c.233insA (Figure 4, Table 2). The occurrence of the identified polymorphism caused a shift of reading frame and production of a protein with a different structure (Figure 5). For both analyzed groups of cows (group I and group II), 2 genotypes were detected: homozygotes AA (wild-type) without insertion, and heterozygotes AB with insertion of nucleotide A (Figure 6). In group II, the frequencies were AA: 97.7% and AB: 2.7% (Table 3). According to the Hardy-Weinberg principle, the population was imbalanced ( $P = 0.003$ ); thus, the null hypothesis that the population is in Hardy-Weinberg equilibrium was

**Figure 4.** A chromatogram depicting insertion of A nucleotides (ENSBTAP00000001585.6:c.233insA), resulting in a frame shift (the red arrow shows the location of the insertion).



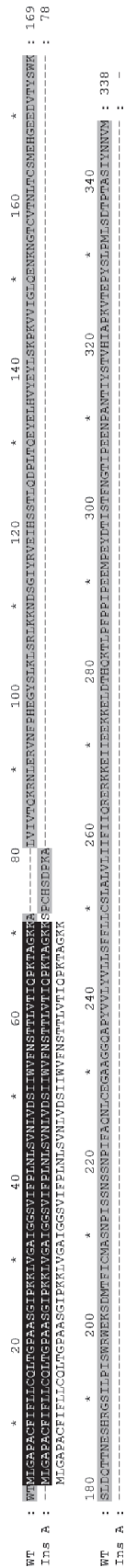


Figure 5. The polymorphism causes a shift of the reading frame and production of a protein with a different structure. WT = wild type; Ins A = insertion A.

Table 3. Frequencies of the signaling lymphocyte activation molecule family 7 (*SLAMF7*) genotype in the cows of group II<sup>1</sup>

Frequency (%)	
Genotype AB (n = 6)	Genotype AA (n = 174)
2.3	97.7

<sup>1</sup>n = number of observed cows.

rejected. Hardy-Weinberg disequilibrium means that the genotype frequencies in a population will not remain constant from generation to generation.

A different situation regarding genotype distribution was observed in group I. Significant differences in genotype distribution were observed between the cows with mastitis and healthy cows. In the groups of cows in which bacteria were present in milk, regardless of the type of pathogen, a low homozygote AA number was observed, 6.25% (in the CPS group) and 12.50% (in the CNS group), whereas the number of homozygotes in healthy cows (the H group) was ~66.60% (Table 4). It should be noted that the populations were adequately numerous (i.e., the groups were large enough) for the gene expression studies, whereas for the genotyping studies, they were too small. In group II, the distribution of genotypes was different. Within the group of 167 genotyped animals, only 6 were heterozygotes AB, and the rest were AA; however, according to the Hardy-Weinberg principle, the population was in balance ( $P > 0.05$ ), and thus, the null hypothesis that the population is in Hardy-Weinberg equilibrium was not rejected. This finding explains why, from generation to generation, it is mainly the cows with the AA genotype that are selected for breeding. Differences ( $P < 0.01$ ) were observed between the AA and AB genotypes in milk and lactose yield and in casein content, with higher yield traits and lower casein content associated with the AA genotype (Table 5). Despite the low number of genotyped animals, the estimated genetic parameters did not differ from those obtained for the worldwide population of dairy cattle (Table 6; Costa et al., 2019).

Table 4. Frequencies of the signaling lymphocyte activation molecule family 7 (*SLAMF7*) genotype in healthy and affected cows (group I)<sup>1</sup>

Group	N	Frequency (%)	
		Genotype AB	Genotype AA
1: CPS	22	93.75	6.25
2: CNS	10	87.50	12.50
3: Healthy	8	33.30	66.60

<sup>1</sup>N = number of observed cows; CPS = coagulase-positive staphylococci.

**Table 5.** Effect of the indel in the signaling lymphocyte activation molecule family 7 (*SLAMF7*) gene on milk yield, lactose yield, and casein content, with SE and differences between the estimates assessed in group II

Trait	N <sup>1</sup>	Genotype variant	Estimates	SE
Daily milk yield	3,283	AA	30.23 <sup>A</sup>	1.59
	144	AB	27.28 <sup>B</sup>	2.37
Lactose yield	3,283	AA	1.38 <sup>A</sup>	0.08
	144	AB	1.24 <sup>B</sup>	0.11
Casein content	2,065	AA	2.79 <sup>B</sup>	0.08
	83	AB	2.88 <sup>A</sup>	0.11

<sup>A,B</sup>Different superscripts within a column and within a trait indicate the differences at  $P < 0.01$ .

<sup>1</sup>N = number of observations.

### Methylation Analysis

The PM differences between the groups of inspected samples were not significant (Table 7). The averaged PM values for all of the investigated groups pointed to hypermethylation of an upstream element of *SLAMF7* (CPS: 68.62%, CNS: 68.37%, and H: 66.40% of methylation). Results of the PM calculations are presented in Figures 7, 8, and 9.

## DISCUSSION

Alternative splicing of eukaryotic pre-mRNA is a key mechanism for potentially producing many transcript isoforms from a single gene. In many cases, AS can affect the coding sequence, leading to the production of diverse proteins (Jacob and Smith, 2017). Alternative splice variants of bovine *SLAMF7* have been reported by Ju et al. (2012), which are consistent with the reports concerning humans and mice. In their study, Ju et al. (2012) characterized 3 novel splice variants of bovine *SLAMF7* that are expressed in the mammary tissues of cattle: complete, AS1, and AS2. The patterns of these variants are exon skipping and alternative 5' splice sites. The predicted protein encoded by 2 bovine *SLAMF7* fragments—complete and AS1—also showed that they share the same extracellular and transmembrane domains, but differ in their cytoplasmic regions, which is consistent with the human isoforms. Ju et al. (2012) showed that the *SLAMF7* gene was differentially expressed in healthy and mastitis-affected mammary

tissue and was investigated to understand its role in regulating mammary gland defense in dairy cows. The expression level of the variant AS1 was higher in *S. aureus*-induced mastitis mammary tissue than in healthy mammary tissue and was comparable to that of the complete variant and the AS2 variant. In this current study, we characterized 2 splicing fragments of the *SLAMF7* gene: V1, V2, and complete gene expression. Our results contrasted with those of Ju et al. (2012): the *SLAMF7* mRNA expression levels of the V1, V2, and complete variants were significantly higher in healthy mammary tissues than in infected mammary tissues, and the highest expression was detected in the *SLAMF7* complete amplicon (total gene expression V1–V3 transcripts).

The results relating to the indel type polymorphism in exon 2 of the *SLAMF7* gene showed 2 different genotypes observed in the study populations: AA and AB. Unfortunately, we did not observe any BB individuals. The distribution of the genotypes did not allow us to establish an unambiguous association between the polymorphism and the presence of mastitis. When this insertion occurs in the sequence of the *SLAMF7* gene, after 9 new nucleotides (Figure 5), a stop codon follows. In this case, a fragmentarily different protein will be formed, which consists of 78 AA from the 336 presented in native protein (1/3 of the version without the insertion). Such a huge modification of protein structure can result in differences in *SLAMF7* function and cell immunity. In connection with such a large effect of this change, it seems that homozygous BB individuals may not be present in the population due to the total lack of functional protein, which may be lethal.

The association analysis was implemented using data from group II. Despite the small number of animals, especially those carrying the AB genotype, the standard error was small. Moreover, as mentioned above, the heritability coefficients of all of these traits were comparable to those obtained in the whole Polish dairy cattle population. The very low frequency of the AB genotype in the studied herds, as well as the Hardy-Weinberg principle disequilibrium in the whole population, can be connected with the choice of mate and the assortative mating of breeding animals. Cows carrying the AB genotype probably had a lower lactation milk yield and therefore were probably excluded from breed-

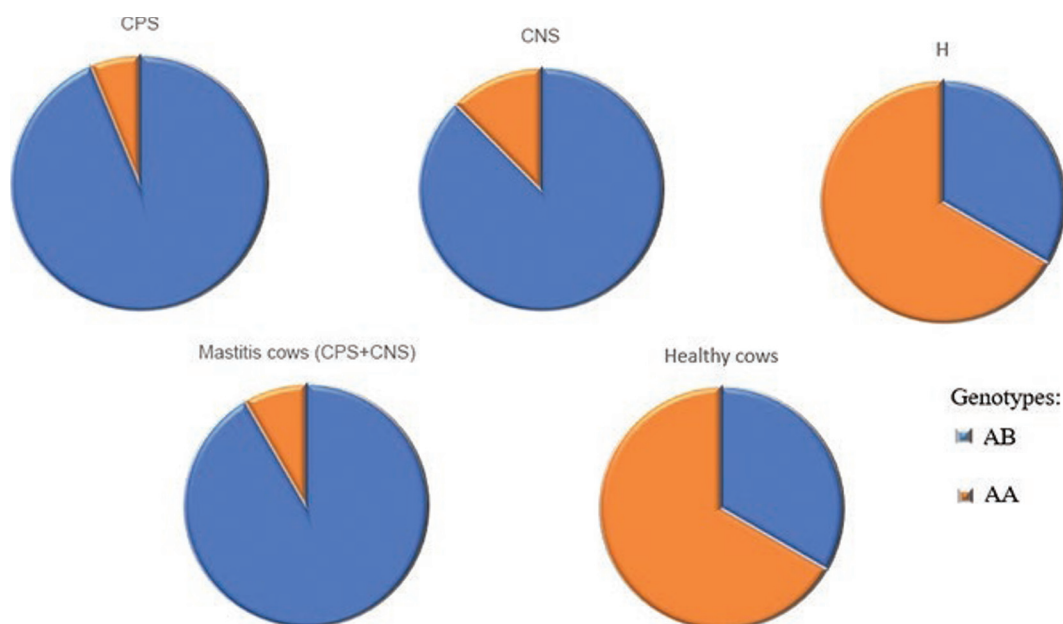
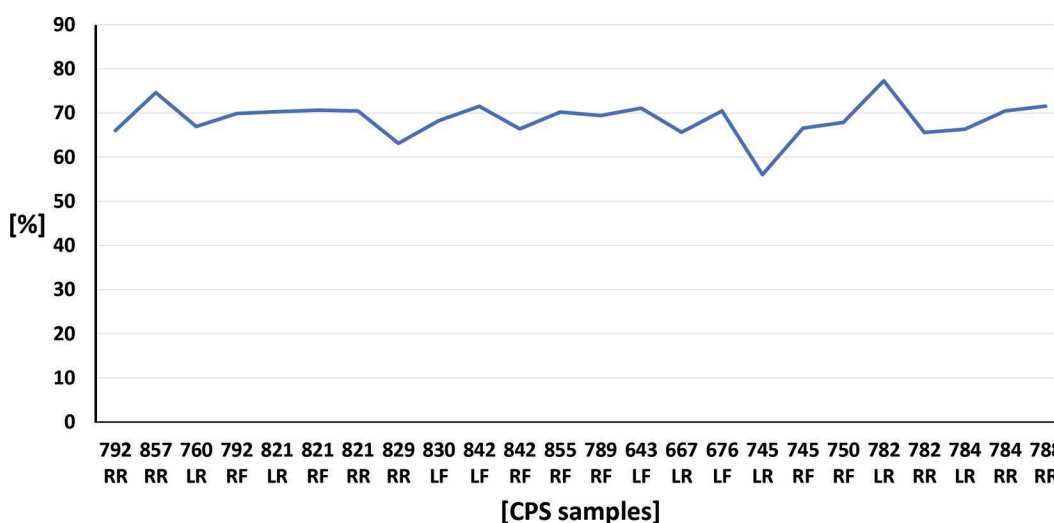
**Table 6.** Genetic parameters for milk yield, lactose yield, and casein content estimated in group II<sup>1</sup>

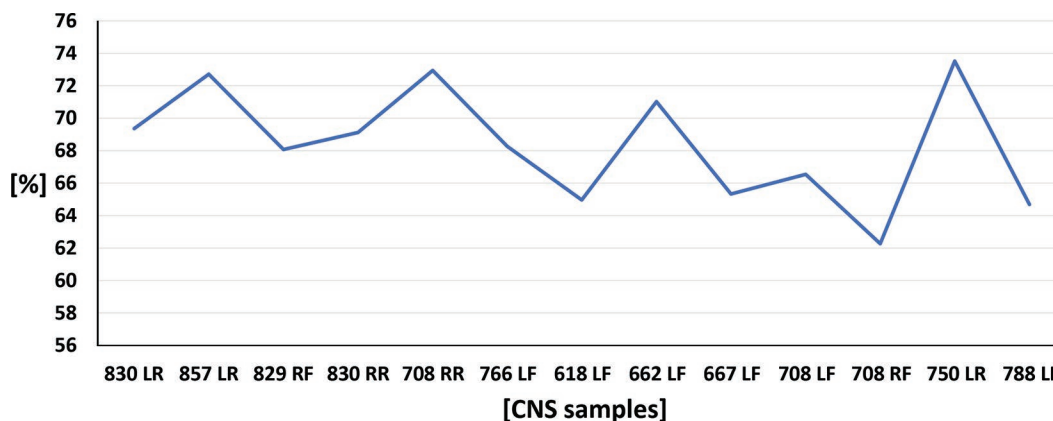
Trait	N	$\sigma^2$	Mean	SD	$h^2$	$r^2$
Daily milk yield	3,427	12.80	32.67	8.87	0.31	0.41
Lactose yield	3,427	0.03	1.57	0.44	0.34	0.41
Casein content	2,148	0.01	2.72	0.33	0.22	0.60

<sup>1</sup>N = number of observations;  $\sigma^2$  = variance;  $h^2$  = coefficient of heritability;  $r^2$  = coefficient of repeatability.

**Table 7.** Significance level of percentage of methylation (PM) differences with regard to the health status of the inspected tissue sections

Item	CpG2	CpG3	CpG4	CpG5	CpG6	Averaged PM across all CpG sites
Group 1 vs. 2	0.369	0.766	0.561	0.519	0.387	0.649
Group 1 vs. 3	0.152	0.107	0.123	0.612	0.563	0.056
Group 2 vs. 3	0.91	0.277	0.134	0.91	0.82	0.167

**Figure 6.** The signaling lymphocyte activation molecule family 7 gene (*SLAMF7*) genotype frequency in mastitis-infected and healthy cows (group I). CPS = coagulase-positive staphylococci group; H = healthy group (i.e., no pathogenic bacteria present in milk).**Figure 7.** Average methylation percentage across 6 CpG sites of an upstream element of signaling lymphocyte activation molecule family 7 gene (*SLAMF7*), in mammary glands infected with coagulase-positive staphylococci (CPS). RR, LR, RF, and LF = different sections of mammary gland.



**Figure 8.** Average methylation percentage across 6 CpG sites of an upstream element of signaling lymphocyte activation molecule family 7 gene (*SLAMF7*), in mammary glands infected with CNS. RR, LR, RF, and LF = different sections of mammary gland.

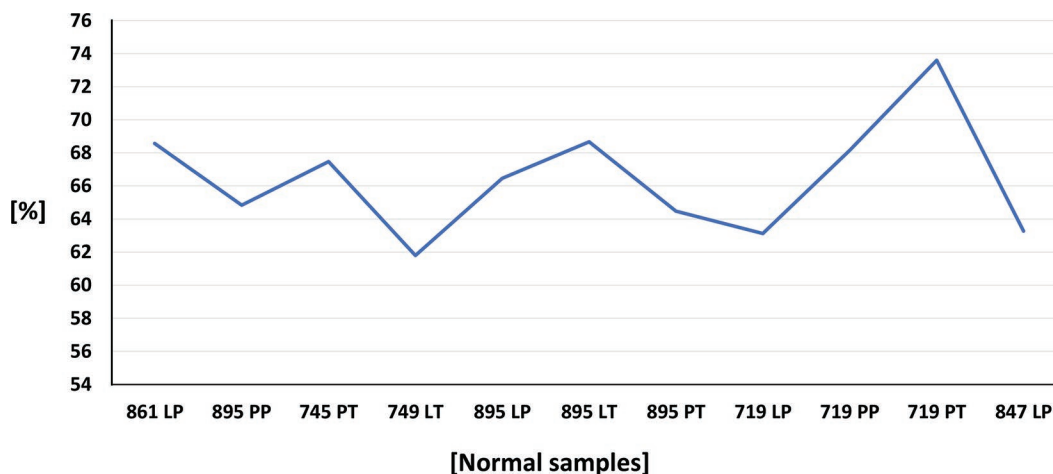
ing. Thus, in herds of an intensive production system, animals carrying the AA genotype are mainly used.

Demethylation of regulatory DNA sequences in a permissive environment are often associated with gene up-regulation in particular biological systems (Deaton and Bird, 2011). The *SLAMF7* plays an important role in the differentiation of immune cells, and demethylation of its promoter, associated with altered *SLAMF7* expression, has been demonstrated during hematopoiesis (Bocker et al., 2011). In this report, we did not detect any significant methylation alterations associated with *SLAMF7* upregulation relevant to a possible immune response to a pathogenic agent (i.e., CPS and CNS) in the mammary glands of milk cows. Classical regulatory elements of genes undergoing differential methylation associated with expression changes are often enriched with CpG dinucleotides (Deaton and Bird, 2011). In

contrast to this idea, the inspected upstream DNA fragment indicated low CG content. Moreover, inspection of the 26 kbp region (BTA3:8964738–8990751), including the *SLAMF7* gene plus upstream and downstream sequences located nearby, did not reveal the presence of any CpG-enriched sequences being potential targets for differential methylation. It seems that the analyzed DNA fragment instead represents the numerous DNA sequences undergoing hypermethylation across the genome, which is important for genomic stability in normal cells (Deaton and Bird, 2011).

## CONCLUSIONS

The specificity of the *SLAMF7* gene is very interesting, and knowledge about this gene is valuable in many different aspects. The analysis performed separately for



**Figure 9.** Average methylation percentage across 6 CpG sites of an upstream element of signaling lymphocyte activation molecule family 7 gene (*SLAMF7*), in sections of uninfected mammary gland (the healthy group). Different sections of mammary gland: PT = right rear quarter, LT = left rear quarter, PP = right front quarter, LP = left front quarter.



each isoform showed that health status was strongly associated with the expression levels of individual variants. Associations between the polymorphism in exon 2 of this gene and milk yield, lactose yield, and casein content were found. Nevertheless, our study did not reveal unambiguous associations between the polymorphism of this gene and the occurrence of mastitis, nor did we find differences in DNA methylation patterns among the groups of animals. However, finding new splicing forms, linking them with the mRNA expression levels and the polymorphisms of this gene, and evaluating protein expression could help answer the questions raised in the study.

## ACKNOWLEDGMENTS

The research was funded by the Polish National Science Center (Krakow; grant no. 2015/17/B/NZ9/01561). The authors have not stated any conflicts of interest.

## REFERENCES

- Bionaz, M., and J. J. Loo. 2007. Identification of reference genes for quantitative real-time PCR in the bovine mammary gland during the lactation cycle. *Physiol. Genomics* 29:312–319. <https://doi.org/10.1152/physiolgenomics.00223.2006>.
- Bocker, M. T., I. Hellwig, A. Breiling, V. Eckstein, A. D. Ho, and F. Lyko. 2011. Genome-wide promoter DNA methylation dynamics of human hematopoietic progenitor cells during differentiation and aging. *Blood* 117:e182–e189. <https://doi.org/10.1182/blood-2011-01-331926>.
- Brotherstone, S. I. M. S., I. M. S. White, and K. Meyer. 2000. Genetic modeling of dairy milk yield using orthogonal polynomials and parametric curves. *Anim. Sci.* 70:407–415. <https://doi.org/10.1017/S1357729800051754>.
- Costa, A., N. Lopez-Villalobos, G. Visentin, M. De Marchi, M. Cas-sandro, and M. Penasa. 2019. Heritability and repeatability of milk lactose and its relationships with traditional milk traits, somatic cell score and freezing point in Holstein cows. *Animal* 13:909–916. <https://doi.org/10.1017/S1751731118002094>.
- Deaton, A. M., and A. Bird. 2011. CpG islands and the regulation of transcription. *Genes Dev.* 25:1010–1022. <https://doi.org/10.1101/gad.203751>.
- Hall, T. A. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41:95–98.
- Jacob, A. G., and C. W. J. Smith. 2017. Intron retention as a component of regulated gene expression programs. *Hum. Genet.* 136:1043–1057. <https://doi.org/10.1007/s00439-017-1791-x>.
- Ju, Z., C. H. Wang, Q. Li, M. Hou, S. Gao, Q. Hou, J. Li, J. Huang, and J. Zhong. 2012. Alternative splicing and mRNA expression analysis of bovine *SLAMF7* gene in healthy and mastitis mammary tissues. *Mol. Biol. Rep.* 39:4155–4161. <https://doi.org/10.1007/s11033-011-1198-z>.
- Kościczuk, E. M., P. Lisowski, J. Jarczak, J. Krzyżewski, L. Zwierzchowski, and E. Bagnicka. 2014. Expression patterns of  $\beta$ -defensin and cathelicidin genes in parenchyma of bovine mammary gland infected with coagulase-positive or coagulase-negative staphylococci. *BMC Vet. Res.* 10:246. <https://doi.org/10.1186/s12917-014-0246-z>.
- Kościczuk, E. M., P. Lisowski, J. Jarczak, A. Majewska, M. Rzewuska, L. Zwierzchowski, and E. Bagnicka. 2017. Transcriptome profiling of staphylococci-infected cow mammary gland parenchyma. *BMC Vet. Res.* 13:161. <https://doi.org/10.1186/s12917-017-1088-2>.
- Leakey, T. I., J. Zielinski, R. N. Siegfried, E. R. Siegel, C. Y. Fan, and C. A. Cooney. 2008. A simple algorithm for quantifying DNA methylation levels on multiple independent CpG sites in bisulfite genomic sequencing electropherograms. *Nucleic Acids Res.* 36:e64. <https://doi.org/10.1093/nar/gkn210>.
- López-Bigas, N., B. Audit, C. Ouzounis, G. Parra, and R. Guigó. 2005. Are splicing mutations the most frequent cause of hereditary disease? *FEBS Lett.* 579:1900–1903. <https://doi.org/10.1016/j.febslet.2005.02.047>.
- Madsen, P. O., and J. Jensen. 2000. A user's guide to DMU. A package for analyzing multivariate mixed models. Version 6, Release 4. Accessed Sep. 2012. [http://dmu.agrsci.dk/dmuv6\\_guide-R4-6-7pdf](http://dmu.agrsci.dk/dmuv6_guide-R4-6-7pdf).
- Medrano, J. F., A. Canovas, and A. Islas-Trejo. 2014. RNA sequencing for the analysis of complex traits in milk. Page 17 in *Proceedings, 10th World Congress of Genetics Applied to Livestock Production*. Vancouver, Canada. <https://asas.confex.com/asas/WCGALP14/webprogram/Paper10461.html>.
- Modrek, B., and C. Lee. 2002. A genomic view of alternative splicing. *Nat. Genet.* 30:13–19. <https://doi.org/10.1038/ng0102-13>.
- Pan, Q., O. Shai, L. J. Lee, B. J. Frey, and B. J. Blencowe. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* 40:1413–1415. <https://doi.org/10.1038/ng.259>.
- Pawlik, A., G. Sender, M. Kapera, and A. Korwin-Kossakowska. 2015a. Association between interleukin 8 receptor alpha gene (*CXCR1*) and mastitis in dairy cattle. *Cent. Eur. J. Immunol.* 40:153–158. <https://doi.org/10.5114/ceji.2015.52828>.
- Pawlik, A., G. Sender, M. Sobczyńska, A. Korwin-Kossakowska, H. Lassa, and J. Oprządek. 2015b. Lactoferrin gene variants, their expression in the udder and mastitis susceptibility in dairy cattle. *Anim. Prod. Sci.* 55:999–1004. <https://doi.org/10.1071/AN13389>.
- Pfaffl, M. W. 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29:45e. <https://doi.org/10.1093/nar/29.9.e45>.
- Quinn, P. J., B. K. Markey, F. C. Leonard, P. Hartigan, S. Fanning, and E. S. Fitzpatrick. 2002. *Veterinary Microbiology and Microbial Disease*. Blackwell Science, Oxford, UK.
- Song, M., Y. He, H. Zhou, Y. Zhang, X. Li, and Y. Yu. 2016. Combined analysis of DNA methylome and transcriptome reveal novel candidate genes with susceptibility to bovine *Staphylococcus aureus* subclinical mastitis. *Sci. Rep.* 6:29390. <https://doi.org/10.1038/srep29390>.
- Strzetelski, J., and B. Śliwiński. 2009. *Normy żywienia bydła*. W: IZ PIB-INRA. *Normy żywienia przeżuwaczy: wartość pokarmowa francuskich i krajowych pasz dla przeżuwaczy* [In Polish], (Standard of cattle feeding. Pages 21–49 in NRIAP-INRA, Standard of ruminants' feeding: Nutrient value of French and domestic fodders for ruminants). J. Strzelecki, ed. Research Institute of Animal Production, Cracow, Poland.
- Tazi, J., N. Bakkour, and S. Stamm. 2009. Alternative splicing and disease. *Biochim. Biophys. Acta* 1792:14–26. <https://doi.org/10.1016/j.bbdis.2008.09.017>.
- Wang, X. G., Z. H. Ju, M. H. Hou, Q. Jiang, C. H. Yang, Y. Zhang, Y. Sun, R. L. Li, C. F. Wang, J. F. Zhong, and J. M. Huang. 2016. Deciphering transcriptome and complex alternative splicing transcripts in mammary gland tissues from cows naturally infected with *Staphylococcus aureus* mastitis. *PLoS One* 11:e0159719. <https://doi.org/10.1371/journal.pone.0159719>.
- Yang, Y., J. M. Huang, Z. H. Ju, Q. L. Li, L. Zhou, R. L. Li, J. B. Li, F. X. Shi, J. F. Zhong, and C. F. Wang. 2012. Increased expression of a novel splice variant of the complement component 4 (*C4A*) gene in mastitis-infected dairy cattle. *Genet. Mol. Res.* 11:2909–2916. <https://doi.org/10.4238/2012.May.18.12>.