The Effects of *MBL1* Gene Polymorphism on Subclinical Mastitis in Holstein Cows

Esma Gamze AKSEL 1,a Aytaç AKÇAY 2,b (*) Korhan ARSLAN 1,c Mahmodul Hasan SOHEL 1,d Güven GÜNGÖR 3,e Bilal AKYÜZ 3,f

1 Erciyes University, Faculty of Veterinary Medicine, Department of Genetics, TR-38039 Kayseri - TURKEY
2 Ankara University, Faculty of Veterinary Medicine, Department of Biostatistics, TR-06110 Ankara - TURKEY
3 Erciyes University, Faculty of Veterinary Medicine, Department of Biometrics, TR-38039 Kayseri - TURKEY

ORCIDs: a 0000-0002-0040-8933; b 0000-0001-6263-5181; c 0000-0002-2440-884X; d 0000-0003-2224-085X; e 0000-0003-3695-9443; f 0000-0001-7548-9830

Abstract

This study aims to investigate the effects of three single nucleotide (SNPs) polymorphisms of the *MBL1* gene on subclinical mastitis for Holstein cows. For this study, a total of 151 Holstein cows were selected in their third lactation. The subclinical mastitis conditions were determined using the California Mastitis Test (CMT). Genotyping was carried out with the PCR-RFLP method. The results revealed that the wild-type allele frequencies were 0.72 (G allele), 0.40 (G allele), and 0.37 (T allele) for 1252 G>A, 2534 G>A, and 2569 T>C SNPs, respectively. Among the investigated SNPs, only the 1252 G>A SNP was not found in Hardy-Weinberg equilibrium. The effects of SNPs on subclinical mastitis were modeled using multiple logistic regression analysis. The established model can identify cows with subclinical mastitis with a separation efficiency of 62.3%. It was thought that 1252 G>A and 2534 G>A SNPs may affect subclinical mastitis rates.

Keywords: Cow, *MBL1*, Multiple logistic regression analysis, SNP, Subclinical mastitis

INTRODUCTION

Infectious diseases lead to significant economic losses in both the dairy and meat industry because of the cost of treatment, loss of productivity, and finally the death of infected animals [1]. Among the infectious diseases, mastitis leads to huge economic losses in the dairy cattle industry, however, the total cost of economic losses is tried to be determined using several simulation models [2]. Mastitis is an infection occurring in the udder lobes as a result of an infection by different types of bacteria such as *Escherichia coli*, *Klebsiella spp.*, *Staphylococcus aureus*, *Streptococcus spp.* [3,5], and some species of yeast, fungi, and mycobacteria [6]. Mastitis is usually classified as clinical, subclinical, and chronic mastitis according to the symptoms of infected animals [6]. Clinical mastitis (CM) was characterized by swelling, pain, and redness of the mammary, it can cause decreased yield of milk, and altered appearance of milk [4].

How to cite this article?


(*) Corresponding Author

Tel: +90 312 317 0315/4475 Cellular Phone: +90 505 502 5823 Fax: +90 312 316 4472
E-mail: aytacakcay@ankara.edu.tr (A. Akçay)

This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)
On the other hand, subclinical mastitis does not create any visible changes, neither in the udder physiology nor in the appearance of milk. Besides the reduction in milk production, subclinical mastitis could transform into clinical mastitis if it is diagnosed at a later stage. Moreover, this leads to the early culling of productive animals from the herd that ultimately resulted in serious economic losses for farms [9,8]. It has been reported that subclinical mastitis cases are more common than clinical mastitis in Turkey, and about 70-80% of the loss of milk production cases concerning mastitis are caused by sub-clinical mastitis [9]. The prevalence rate of subclinical mastitis per cow was reported to be 46.4% in Brazil [10], 43.1% in Rwanda in Africa [11], 42.5% in Iran [12], and 40% in Turkey [13].

Animal care, udder and milking hygiene, and herd management play a significant role in the prevention of mastitis [14]. Nevertheless, most of the studies have concentrated on establishing genetically resistant herds against diseases [14,15]. The MBL is a collagenous serum lectin that can bind to many pathogens including bacteria, fungi, viruses, and parasites. Moreover, it participates in the activation of the innate and adaptive immune systems of organisms [17,18]. By binding with the surface carbohydrates of pathogens, MBL activates the lectin pathway, which facilitates the recognition of immune cells and mediates their phagocytosis [19]. It was reported that a low level of serum MBL protein was significantly associated with susceptibility to bacterial and fungal infections [17]. Eleven collagenous lectin genes were identified in cattle, including genes encoding MBL1 (A-type lectin) and MBL2 (C-type lectin) in the MBL pathway [1]. It was reported that several SNPs in MBL1 and MBL2 genes were associated with mastitis resistance [1]. Furthermore, it is reported to be associated with a missense mutation (Val24Ile, rs109492835) in the exon 2 of the MBL1 gene in cattle as well as with another mutation detected in the promoter region and the somatic cell score (SCS) in milk [14,15,20]. In another study was undertaken to identify whether the SNP rs109231409 (g.855G>A) of the MBL1 gene was associated with mastitis tolerance or not which were analyzed with CMT and SCS [21]. Also on the MBL1 gene, 1252 G>A, 2534 G>A, and 2569 T>C SNPs were examined to be associated with mastitis [14,15,20,22]. However, the association of these SNPs in the MBL1 gene with subclinical mastitis has not been investigated yet.

In the present study, we have genotyped three SNPs including in intron 1 (1252 G>A) and exon 2 (2534 G>A and 2569 T>C) of the MBL1 gene. This study was aimed to identify the separation efficiency of genotypes in positive and negative cases of subclinical mastitis at the end of genotyping by using a statistical model of Multiple Logistic Regression.

**MATERIAL AND METHODS**

**Animals and Mastitis Test**

A total of 151 Holstein dairy cows in their third lactation were randomly selected for this study. Blood samples of the cows used in the study were collected in accordance with the permission of Erciyes University Animal Experiments Local Ethics Committee, dated 14.06.2017 and numbered 17/056. The presence or absence of subclinical mastitis was determined by the CMT in farm conditions. According to CMT test results, the animals were categorized as positive (1) or negative (0).

**Sample Collection and Processing**

Blood samples were collected from the vena jugularis of cows. DNA was isolated from collected blood samples using the standard phenol:chloroform:isoamyl alcohol method [20]. DNA purity and concentrations were measured using a spectrophotometer (BioSpech-nano, SWEDEN).

**PCR Primers and PCR Conditions**

The sequences of primers [20] are presented in Table 1. The final volume of PCR reaction was found to be 20 µL consisting of 3 µL DNA (50 ng/µL), 1X buffer (NH₄)₂SO₄ 500 mmol/L), 2.0 mmol/L MgCl₂, 0.25 mmol/l dNTP, and 0.5 U Taq DNA polymerase. The thermal cycles were as follows: 1 cycle at 95°C for 5 min after pre-denaturation, followed by 35 cycles at 95°C for 30 sec, at the annealing temperatures

**Table 1. Information on genotyped SNPs**

<table>
<thead>
<tr>
<th>SNP</th>
<th>Forward and Reverse Primers</th>
<th>AL</th>
<th>TA</th>
<th>Genotype/bp</th>
<th>RE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2569 T&gt;C</td>
<td>F: GTGGTGGCAAAATGTTGGTCAAAAC R: TGGTCCTCCCTCCCTCCCTT</td>
<td>255 bp</td>
<td>63.5°C</td>
<td>TT: 255 TC: 255, 178, 77 CC: 178, 77</td>
<td>Haelll</td>
</tr>
</tbody>
</table>

AL: Amplicon Length; bp: base pair; TA: temperature of anningel; RE: Restriction enzyme
for 30 sec (Table 1), at 72°C for 30 sec, and finally, 1 cycle at 72°C for 7 min.

Genotyping of DNA

The resulting PCR products were digested using the restriction endonuclease enzymes (Thermo Scientific, USA) following the manufacturer’s instructions. Briefly, enzymatic digestion was conducted in a 20 μL final volume containing 5 U of restriction enzyme (Table 1) and 10 μL of PCR products. The reaction mix was placed at various incubation temperatures as reported in the manufacturer’s protocols for 4 h and inactivated at 80°C for 20 min. The digested products were visualized at 90 V and 300 mA on a 2% agarose gel dyed ethidium bromide for 40 min using an ultra-violet (UV) transilluminator. Expected restricted band sizes for each SNP after the restriction enzyme digestion are presented in Table 1.

Statistical Analysis

In the study, genotypic examinations were performed for the three SNPs in the intron-1 (1252 G> A) and exon 2 regions (2534 G> A, 2569 T> C) of the MBL1 gene. The genotype, allele frequencies of the studied samples, and whether the population was in Hardy Weinberg equilibrium in terms of three SNPs were analyzed using the Chi-square goodness of fit test. Moreover, the effect of the presence of subclinical mastitis on their genotypes was identified using the multivariate logistic regression analysis. Univariate logistic regression analysis was conducted to examine the relationship between the candidate variables, which were not included in the logistic model, and mastitis.

In the multiple logistic regression model, the presence of mastitis was defined as the dependent variable (Y) while the genotypic frequencies of 1252 G>A, 2534 G>A, 2569 T>C SNPs were defined as independent variables. The backward elimination method was used in the logistic regression analysis. The statistical significance of the variables in the multivariate model was tested using the Wald test. The success of the model was determined using the deviation and Chi-square goodness of fit criteria. The coefficients of the logistics model were calculated using the maximum likelihood estimation method. The goodness of fit test of the final model was evaluated using the Hosmer-Lemeshow test value (Ĉg*). The classification accuracy of the model (the separation efficiency of the model to separate subclinical mastitis positive and negative cows using the predicted probabilities) was determined. The final model was evaluated by using odds ratios [OR, Exp(β)]. The statistical analyses were carried out using SPSS for Windows 14.01 (License No: 9869264) software package.

Results

The size of the PCR product of the 1252G> A SNP in the intron-1 region was found to be 226 bp. The results revealed that there were three genotypes (Table 1) for the 1252G>A SNP which were restricted with the AvaII enzyme (Fig. 1). The size of the PCR product of the 2534 G>A SNP in the exon-2 region was found to be 217 bp. Three genotypes (Table 1) were determined after the digestion of the PCR products using the MaeII restriction enzyme (Fig. 2). The size of the PCR product of the 2569 T>C SNP in the exon-2 region was found to be 255 bp. Three genotypes (Table 1) were observed after the digestion of the PCR products using the HaeIII restriction enzyme (Fig. 3).

The Chi-square test was conducted to investigate whether the genotypes of the 1252 G>A, 2534 G>A, 2569 T>C SNPs were in the Hardy-Weinberg (H-W) equilibrium. The results revealed that the analyzed samples of the 1252 G>A SNP were not in the H-W equilibrium (P<0.05), while the 2534 G>A and 2569 T>C SNPs were in the H-W equilibrium (P>0.05) (Table 2). The genotype frequency of the 1252 G>A SNP was observed to be 11.9%, 31.8%, 56.3% for the AA, GA, and GG genotypes, respectively. On the other hand, the allele frequency was 0.28 and 0.72 for the A and
The details of the genotype and allele frequencies for all SNPs are presented in Table 2.

Genotype distribution of the 1252 G>A, 2534 G>A, and 2569 T>C SNPs in mastitis positive and negative cows in the study is presented in Table 3.

The effect sizes of the genotypes of the 1252 G>A, 2534 G>A, and 2569 T>C SNPs, which were thought to affect subclinical mastitis, were determined using the logistic regression analysis. Univariate logistic regression analysis was conducted to examine the relationship between the candidate variables, which were not included in the logistic model, with mastitis. According to the results, the 1252 G>A and 2534 G>A SNPs were found to be statistically significantly associated with subclinical mastitis (P<0.25) (Table 4).

The backward elimination method was used in the multiple logistics model. Initially, all variables were taken into the model. Then those found to be statistically non-significant were deleted from the model at each step. The iteration was completed in 2 steps. The results of the backward elimination method are presented in Table 5.

At the end of the second step of the logistic regression
models, the level of correlations between the dependent variable and independent variables were evaluated using the Cox-Snell’s R² and Nagelkerke’s R² (Pseudo R²) statistics. According to the results of this evaluation, the correlations between the dependent and independent variables in the second step were found to range between 8% and 12%.

The Hosmer-Lemeshow test value (Ĉg*) is one of the goodness of fit test statistics calculating the success of the model to explain the dependent variable. The Hosmer-Lemeshow test (Ĉg*) values (χ² = 0.063, P>0.05) calculated for the second step showed that the independent variables (genotypes) in the logistic regression model sufficiently explained the dependent variable (mastitis ratio).

The variable coefficients were evaluated for each step using the backward elimination method, and the non-significant variables were removed from the model in the next step. That the constant term and the coefficients of the independent variables were different from zero in the model obtained in the second step was found to be statistically significant (P<0.05). According to the final model, the independent variables “1252 G>A, 2534 G>A” were found to affect the mastitis rate, and the obtained model could predict the probability of mastitis (P) by using the following equation:

\[ \ln \left( \frac{P}{1-P} \right) = 0.120 + 4.220GA + 6.447GG + 3.839GA + 1.866GG \]

The classification rate of the model, i.e. the separation efficiency of the model in separating mastitis positive and mastitis negative cows using the predicted probabilities, was found to be 62.3%.

**Discussion**

The mannose-binding lectin pathway is one of three pathways that stimulate the complement system in the immune system [24]. This pathway binds to the mannose sugar of the pathogen, allowing early complement system members to be drawn into the pathogenic focus. Thereby, it plays a role in the activation of the complement system.
and thus in the pathogen phagocytosis of the natural immune cells [28]. Furthermore, the development of the membrane attack complex, and thus the direct lysis of the pathogen, is achieved through this pathway [25]. Mastitis is a disease leading to significant economic losses in the dairy cattle industry throughout the world since it is the most common disease caused by bacteria [15].

Molecular characterization of a gene and alleles are a useful tool to evaluate the effect of that gene on a certain character [24]. The present study investigates the relationship between the three SNPs in the MBL1 gene and the subclinical mastitis at a molecular level. The frequency of the G allele in the 2534 G>A SNP was found to be 0.40. In a study on Holstein cattle in China, the G allele frequency in the g.2651G>A coded SNP in the MBL1 gene (same as 2534 G>A SNP in the present study) was found to be 0.40. In a study on Holstein cattle in China, the G allele frequency in the g.2651G>A coded SNP in the MBL1 gene (same as 2534 G>A SNP used in the present study) was reported to be 0.58 [15], while another study conducted in China reported that the frequency was 0.61 [14]. A recent study on the Zavot, Native Black, Eastern Anatolian Red, Turkish Grey Cattle, Southern Anatolian Red, Brown Swiss, and Simmental cattle breeds, it was revealed that Brown Swiss had the highest GG estimated genotype frequency (0.83) of the 1262G>A SNP, while the Eastern Anatolian Red had the lowest estimated genotype frequency (0.68); also, Southern Anatolian Red was found to have the highest GG estimated genotype frequency (0.74) of the 2534G>A SNP, while Simmental (0.23) had the lowest estimated genotype frequency [28]. A similar study investigates that the GG estimated genotype frequency of the 1252 G>A SNP was the highest (0.50) in the Sahne breed, while the lowest frequency was observed in the Holstein breed (0.33).

Also, the Holstein breed was found to have the highest GG estimated genotype frequency of the 2534 G>A SNP (0.44) while the Sahne breed was found to have the lowest frequency (0.29). On the other hand, the TT estimated genotype frequency of 2569 T>C SNP was found to be the highest (0.45) in the Holstein breed while the lowest frequency (0.37) was observed in the Sahne breed [17].

In the present study, the G allele frequencies of the 1252 G>A SNP and the 2534 G>A SNP were found to be 0.72 and 0.40, respectively. Also, the T allele frequency of the 2569 T>C SNP was found to be 0.37. The wild-type allele frequencies of the two SNPs (2534 G>A SNP and 2569 T>C SNP) other than 1252 G>A SNP reported by Yuan et al. [20] were observed to be higher than those found in the present study [17]. The differences in the results might be due to the difference between the sample sizes since a low number of animals were used in the present study. Notably, unlike the other two SNPs, the 1252 G>A SNP was not found to be in the Hardy-Weinberg equilibrium. Therefore, it was considered that the lack of equilibrium in the population of the present study caused this difference in the wild-type alleles frequency of the 1252 G>A SNP compared to the study conducted by Yuan et al. [20].

Yuan et al. [17] investigated the relationship between three SNPs in the MBL1 gene and SCS in Holstein breed cattle in China, and they reported that the individuals with AA genotype in terms of 2534 G>A (g.2651G>A) have a higher SCS than individuals with GG and GA genotypes. The results of the correlation analysis between the genotypes of the SNPs that were utilized in the present study and the somatic cell score [17] revealed that there was a correlation particularly between the 2534 G>A SNP and the SCS in the Sahne, Holstein, and Simmental breed cattle, moreover, individuals with GG genotype had a lower somatic cell score than the individuals with GA and AA genotypes.

In the present study, the genotype frequency of the 2534 G>A coded SNP in the subclinical mastitis was found to be 45.1%, 40.2%, and 14.6% for AA, GA, and GG genotypes, respectively. These results are in accordance with the results found by Yuan et al. [20] and Wang et al. [19], where the authors reported that the individuals with GG genotype of 2534 G>A SNP may have a lower mastitis rate [15,17]. Moreover, we have investigated the association of subclinical mastitis with the genotype frequency in the Holstein breed cattle using multiple logistic regression analysis. According to the final model, the 1252 G>A, and the 2534 G>A coded SNPs were found to affect the subclinical mastitis, and the

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Genotype</th>
<th>β</th>
<th>SE (β)</th>
<th>Wald</th>
<th>P</th>
<th>OR</th>
<th>OR's 95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Min</td>
</tr>
<tr>
<td>1252 G&gt;A</td>
<td>AA</td>
<td>1.440</td>
<td>0.461</td>
<td>5.068</td>
<td>0.024</td>
<td>4.220</td>
<td>1.205–14.781</td>
</tr>
<tr>
<td></td>
<td>GA</td>
<td>1.864</td>
<td>0.630</td>
<td>8.759</td>
<td>0.003</td>
<td>6.447</td>
<td>1.877–22.146</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>-2.119</td>
<td>0.722</td>
<td>8.619</td>
<td>0.003</td>
<td>0.208</td>
<td>0.076–4.931</td>
</tr>
<tr>
<td>2534 G&gt;A</td>
<td>AA</td>
<td>1.345</td>
<td>0.530</td>
<td>6.435</td>
<td>0.011</td>
<td>3.839</td>
<td>1.358–10.852</td>
</tr>
<tr>
<td></td>
<td>GA</td>
<td>0.624</td>
<td>0.496</td>
<td>1.584</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td>-2.119</td>
<td>0.722</td>
<td>8.619</td>
<td>0.003</td>
<td>0.208</td>
<td>0.076–4.931</td>
</tr>
</tbody>
</table>

β: Estimated slope coefficient; SE (β): Standard error of the estimated slope coefficient; Wald: Wald statistic, which tests whether the slope coefficients for the model are equal to zero; P: P value of Wald statistics; OR: Estimated odds ratio and 95% confidence interval values are given.
separation efficiency of the model in separating cows with subclinical mastitis was determined as 62.3%.

It was concluded that the 1252 G>A SNPs in the intron region and the 2534 G>A SNPs in the exon region of the MBL1 gene may have a functional role that affected subclinical mastitis. Based on the obtained results, it was interpreted that the probability of subclinical mastitis could be determined considering the allelic structure of the cows. But further studies are required with more sample and also subclinic mastitis values should were verified with SCS for CMT tests.

**FINANCIAL SUPPORT**

This research received no grant from any funding agency/sector.

**CONFLICT OF INTEREST**

The authors declare that they have no conflict of interest.

**REFERENCES**


