Association of Toll-Like Receptor 4 Polymorphisms with Somatic Cell Score and Lactation Persistency in Holstein Bulls

B. S. Sharma, I. Leyva, F. Schenkel, and N. A. Karrow¹

Centre for Genetic Improvement of Livestock, Department of Animal and Poultry Science, University of Guelph, Guelph, N1G 2W1, Canada

ABSTRACT

Mastitis, an inflammatory disease of the mammary gland generally caused by intramammary infections, is the most frequently occurring disease in the North American dairy industry. Reduced milk yield, milk quality, and lactation persistency as well as early culling contribute to the economic losses associated with this disease. During intramammary infections, cells of the innate immune system become activated through pattern recognition receptors that recognize conserved molecular signatures associated with the invading pathogen. The quality, timing, and intensity of the host inflammatory and subsequent immune response determine the fate of this disease. Toll-like receptor 4 (TLR4) is an important pattern recognition receptor that recognizes endotoxins associated with gram-negative bacterial infections. Its role in pathogen recognition and subsequent initiation of the inflammatory and immune response makes it a suitable candidate gene for enhancing disease resistance in Canadian Holsteins. In this study, polymorphisms in the *TLR4* gene were identified in the Canadian Holstein bull population. Genotypes and haplotypes were constructed, and their associations with somatic cell score and lactation persistency were determined. Sequencing of selective DNA pools was used to reveal polymorphisms in *TLR4*. Two DNA pools were constituted based on high and low estimated breeding values for somatic cell scores. A total of 3 single nucleotide polymorphisms (SNP), including 1 SNP in a putative promoter region (P-226) and 2 SNP in exon3 (E3+1656 and E3+2021) of *TLR4* were detected. A total of 388 bulls were genotyped for the SNP, haplotypes were reconstructed, and their frequencies were obtained. Polymorphisms in these regions were found to be associated with estimated breeding values for lactation persistency, and somatic cell scores in the Canadian Holstein bull population. The unfavorable alleles at P-226 and E3+1656 were found at a frequency of 40 and 37%, respectively; hence, selection against

these alleles is promising in Canadian Holsteins. Selection against the unfavorable allele, T at E3+2021, is limited because of its low frequency (7%). Two frequently occurring haplotypes (GCC and CTC) occurred in 86% of the Canadian Holstein bull population chosen for genotyping. The most frequent haplotype (GCC; 54%) was found to be associated with higher lactation persistency and lower somatic cell scores. The transversion SNP in the putative promoter region (P-226) was in a potential DNA binding site.

Key words: Toll-like receptor 4, health trait, somatic cell score, Canadian Holstein

INTRODUCTION

Mastitis is an inflammatory disease of the mammary gland that is generally caused by IMI. It is the most frequently occurring disease in North American dairy cows, and economic losses associated with this disease are attributed to reduced milk yield and quality, reduced lactation persistency, and early culling (Fetrow, 2000; Capuco et al., 2003). During IMI, mammary epithelial cells and tissue macrophages become activated via their pattern recognition receptors, which recognize molecular signatures associated with the invading pathogen, called pathogen-associated molecular patterns (PAMP). Pattern recognition receptor–PAMP ligation induces the transcription and secretion of cytokines and chemokines that elicit first the recruitment of blood neutrophils and then monocytes to the site of infection. These cells are the effector cells of the innate immune system and are involved in the inflammatory response and in regulating the subsequent immune response.

The Toll-like receptors (TLR) are a multigene family of pattern recognition receptors that are members of the TLR–interleukin 1 superfamily. These receptors recognize a great variety of PAMP, and therefore play a central role in the initiation of inflammatory response and subsequent adaptive immune response to microbial pathogens (Sabroe et al., 2003; Takeda et al., 2003). Currently, at least 13 members of the TLR family have been identified in mammals; genes encoding 10 of these receptors have recently been mapped to the bovine genome (McGuire et al., 2006).

Received December 14, 2005.

Accepted May 5, 2006.

¹Corresponding author: nkarow@uoguelph.ca

Each member of the TLR family is capable of recognizing a unique set of PAMP. Lipopolysaccharide endotoxin from *Escherichia coli* and other gram-negative bacteria, for example, are recognized by TLR4. Toll-like receptor 4 also recognizes PAMP associated with *Mycobacterium tuberculosis*, *Aspergillus fumigatus*, *Cryptococcus neoformans*, and *Candida albicans* (Poltorak et al., 1998; Lien and Ingalls, 2002; Netea et al., 2002), and host inflammagens such as heat-shock protein (Hsp60), fibrinogen, fibronectin, and oligosaccharides of hyaluronan (Lien and Ingalls, 2002). Peptidoglycan and lipoteichoic acid from *Staphylococcus aureus* and other gram-positive bacteria are uniquely recognized by TLR2 (Bannerman et al., 2004). Recent evidence showing that concentrations of important facilitators of TLR2 and TLR4 signaling, specifically soluble CD14 and LPS-binding protein, are increased in milk during *E. coli* and *Staph. aureus* IMI (Bannerman et al., 2004). That the *TLR2* and *TLR4* genes (but not the *TLR9* gene) are strongly expressed during mastitis caused by *Staph. aureus* (Goldammer et al., 2004) suggests that TLR2 and TLR4 may play a role in the host response to IMI.

Polymorphisms in genes encoding receptors associated with the innate immune system are likely to contribute to the overall variation in the resistance or susceptibility to mastitis in dairy cattle. Polymorphisms in the *CXCR2* chemokine receptor gene, for example, have recently been associated with mastitis susceptibility and neutrophil function (Youngerman et al., 2004; Rambeaud and Pighetti, 2005). Additionally, polymorphisms in *BoLA-DR3* have been associated with clinical mastitis and may play critical role in altered antigen-binding affinity to the major histocompatibility complex class II molecule and recognition of the peptide by T cells (Sharif et al., 2000).

Several lines of evidence suggest that lack of, or polymorphism in, *TLR4* can compromise immune responses to certain pathogens. When *TLR4*^{-/-} knockout mice were challenged with LPS, for example, they were unable to mount a cellular response, and no detectable change in tumor necrosis factor- α was observed (Hoshino et al., 1999). Polymorphisms in the coding and promoter regions of *TLR4* can also determine different host resistance or susceptibility patterns to various infectious diseases. A recently described association between the Asp299Gly polymorphism in human *TLR4* and gram-negative-induced septic shock suggests that a functional defect in *TLR4* leads to an increased susceptibility to gram-negative bacteremia (Lorenz et al., 2002). The same mutation is also associated with hyporesponsiveness to LPS in humans (Arbour et al., 2000). Last, analysis of the truncated and mutated promoter sequence of *TLR4* in mice has revealed several positive

and negative regulatory elements considerably affecting promoter activity and *TLR4* mRNA levels (Roger et al., 2005).

Given that *TLR4* is involved in PAMP recognition, mutations in *TLR4* can compromise the host immune response to certain pathogens, that because *TLR4* is highly polymorphic in the bovine species (White et al., 2003), and *TLR4* expression is associated with IMI, this gene may be a potential candidate for use in marker-assisted selection to enhance mastitis resistance in dairy cattle. Therefore, the intent of this study was to reveal the 5'-upstream promoter sequence to detect single nucleotide polymorphisms (SNP) in the putative promoter region of *TLR4*, and to evaluate the possible association of SNP in the coding and promoter regions with health-related traits such as SCS and persistence of lactation in Canadian Holsteins.

MATERIALS AND METHODS

Resource Population

A total of 388 bull semen samples were selected from a resource population of 2,166 Holstein bulls. The EBV of these bulls were obtained from a national genetic evaluation database generated in October 2003 by the Canadian Dairy Network. Among the 388 bulls, a total of 48 bulls were selected that were among the most frequently used sires in an Ontario sentinel herd study (Kelton et al., 1999). The remaining 340 bulls were selected within families on the basis of extreme EBV for either protein yield or SCS. These bulls were the progeny of 30 sires. The majority of the selected population was composed of half-sib families of 20 sires, the size of which ranged from 2 to 30 bulls. Semen samples were kindly provided by Semex Alliance (Guelph, Ontario, Canada).

In the present study, the health-related traits considered in the analyses were SCS and LP because QTL for SCS exist on *Bos taurus* autosome 8 (**BTAS**). Somatic cell score represents a log score of the milk SCC and has been genetically correlated with clinical mastitis (Shook and Schutz, 1993). Macrophages are the dominant somatic cell population found in uninfected milk and represent a primary line of defense against IMI (Lee et al., 1980; Sordillo and Nickerson, 1988). Lactation persistency is defined as the expected milk yield on d 280, expressed as a percentage of the average actual base population yield on d 60 (Jamrozik et al., 1998). Animals with a greater LP are generally less susceptible to metabolic disorders, health problems, and fertility problems (Dekkers et al., 1998), and clinical mastitis can increase mammary epithelial cell apoptosis, thereby reducing LP (Capuco et al., 2003).

Figure 1. Diagram showing 3 single nucleotide polymorphisms (SNP) on sequencing electropherograms of DNA pools in the *TLR4* gene.

The EBV for SCS and LP were available for 387 and 359 bulls, respectively.

DNA Isolation

The DNA was extracted from semen samples following the standard phenol-chloroform method with minor modifications (Winfrey et al., 1997). The DNA concentration and quality were assessed based on the absorbance of UV light at 260 (A_{260}) and 280 nm (A_{280}) with an Eppendorf BioPhotometer (Berlin, Germany).

Determination and Analysis of the 5'-Upstream Sequence to Exon 1 of *TLR4*

The 5'-flanking sequence upstream of exon1 of *TLR4* was obtained using PCR and adaptor-ligated genomic *B. taurus* DNA restriction enzyme fragment libraries, generated using the Universal GenomeWalker Kit (BD Biosciences Clontech, Palo Alto, CA). Briefly, 4 separate walker libraries were constructed by ligating a specially designed adapter sequence to bovine genomic DNA (BD Biosciences Clontech), and each was digested by 4 different restriction enzymes (*Dra*I, *Eco*R V, *Pvu*II, and *Stu*I). The gene-specific nested primers were designed from the *TLR4* exon1 and available upstream sequence (GenBank accession no. AY297041). The major distinct PCR products were then sequenced by gene-specific primers (GSP). An initial ~800-bp fragment was amplified by 2 rounds of nested PCR using the library made with *Stu*I. The 2 GSP used were TLR4W0 (5'-AGC ATG ACC CTC TGT CTG TGC CGG CCA G-3') and TRL4W1 (5'-CCG GCC AGG GCA GCC TCC GAG GC-3'). The first round of PCR was 2 min at 94°C, followed by 30 s at 94°C and 3 min at 68°C for 35 cycles, and ending at 68°C for 7 min. The second round of PCR was 10 min at 94°C, followed by a touchdown profile. The initial 7 cycles were 30 s at 94°C, 30 s at 68°C, and 2 min at

72°C, followed by 30 cycles for 30 s at 94°C, 30 s at 61°C, and 2 min at 72°C for 30 cycles, and a final extension of 10 min at 72°C. Gene-specific nested primers were designed from the newly obtained sequence and were used for subsequent walking. Similarly, an additional 1.3-kb upstream sequence was amplified by 2 additional rounds of walking. Finally, using genomic bovine DNA as a template, a total of 2,112 bp of *TLR4* promoter region was amplified and sequenced.

Construction of DNA Pools for SNP Detection

Twenty bulls with high EBV for SCS and 20 bulls with low EBV for SCS were selected for creating DNA pools. DNA quantification of these samples was further reassessed using the PicoGreen dsDNA quantification procedure (Molecular Probes; Invitrogen, Carlsbad, CA) on a Victor 3 fluorescent plate reader (PerkinElmer, Wellesley, MA). Two DNA pools were constructed by aliquoting an equal amount of DNA from each selected bull in each group. These pools were used for the amplification of 3 exons of *TLR4* and the newly obtained 5'-upstream putative promoter region of *TLR4* to detect SNP. Three SNP were detected on the sequencing electropherograms (Figure 1), including one SNP at the 226-bp position in the putative promoter region (P-226) and 2 SNP in exon3 of *TLR4* at the 1,656-bp (E3+1656) and 2,021-bp (E3+2021) positions.

SNP Genotyping

The tetra-primer Amplification Refractory Mutation System-PCR procedure (Ye et al., 2001) was used to genotype the 3 SNP. This is a simple and economical method involving a single PCR reaction. The method uses 4 primers (Table 1) to amplify a larger control fragment from a DNA sequence containing a given SNP, producing amplicons representing each of the 2 allelic

Table 1. The tetra-primer Amplification Refractory Mutation System–PCR primers and conditions for genotyping *TLR4* single nucleotide polymorphisms

Position	Primer name	Sequence, 5' → 3'	T _m , °C	Amplicon size, bp
P-226	Forward inner primer (G allele)	TTTCTTCTTCTTCCTCTAACTTCCCCACG	68	G allele: 185
	Reverse inner primer (C allele)	TTTCTCTGATATTTATCTCCTCTGCCAGCG	68	C allele: 277
	Forward outer primer	CCAGGGTATTTTGTATGGCTGGAACAT	68	
	Reverse outer primer	CTCCGAGGCACTGTAGTGTCGTCTCT	68	Two outer primers: 403
E3+1656	Forward inner primer (T allele)	ACTCGCTCCGGATCCTAGACTGCCGT	73	T allele: 202
	Reverse inner primer (C allele)	GCTCCTTAGAGGCCATGATACGTTGCAG	73	C allele: 269
	Forward outer primer	TCTGCAAACCTTGAAAATGGCAGGCAAC	73	
	Reverse outer primer	CTGCACACATCATTTGCTCAGCTCCAC	73	Two outer primers: 416
E3+2021	Forward inner primer (T allele)	AAAAAGTATGGCAGGGGCGAGAGAAT	68	T allele: 272
	Reverse inner primer (C allele)	CGAGTAGATGACAAAGGCATCATCGG	68	C allele: 380
	Forward outer primer	TCATGGCCTCTAAGGAGCAAGAATACTACG	68	
	Reverse outer primer	GGGCAATCTCATACTCGAAGATACACCA	68	Two outer primers: 600

¹T_m = melting temperature.

forms (Figure 2). In this method, a mismatch is deliberately introduced at the 3' end of each of the 2 allele-specific primers to increase the specificity of the reaction. The primers can be designed to amplify fragments of differing sizes corresponding to each allele to enable easy resolution of the allele-specific bands on an agarose gel. For SNP genotyping, primers were designed using the primer design computer program made accessible by Ye et al. (2001) (http://cedar.genetics.soton.ac.uk/public_html/primer1.html). The method proved to be reliable, inexpensive, and easy to use. The PCR was performed in a total volume of 10 µL, containing 30 ng of template DNA, 10 pmol of each inner primer, 1 pmol of each outer primer, 200 µM deoxyribonucleotide triphosphates, 2.5 mM MgCl₂, 1× PCR buffer, and 1 U of AmpliTaq Gold DNA polymerase (Applied Biosystems, Foster City, CA). The touchdown profile was followed using 6°C higher than the annealing temperature in the first PCR cycle, subsequently decreasing the annealing temperature by 1°C per cycle in the following 6 cycles, and then keeping the annealing temperature at the final level thus obtained for the remaining 30 cycles. The PCR profile was as follows: 94°C for 10 min, 36 cycles of 30 s at 94°C, 30 s of annealing temperature (with touchdown profile), and 1 min of extension at

72°C, ending with 5 min at 72°C. The annealing temperatures were 62, 65, and 62°C for P-226, E3+1656, and E3+2021 SNP, respectively. The tetra-primer Amplification Refractory Mutation System–PCR products were observed on a 1.5% agarose gel stained with ethidium bromide.

Haplotype Construction

The haplotype probabilities were reconstructed using algorithm and HAPROB software (Boettcher et al., 2004). This software provides probabilities of haplotype combinations for members of half-sib families. The procedure involves 2 steps and requires genotypes only from siblings, not from parents. In the first step, haplotype probabilities of the sires were determined based on the genotypes of the offspring and their allelic frequencies in the general population. In the second step, the haplotype probabilities of the offspring were calculated based on the sire haplotype frequencies. These 2 steps were alternated iteratively until the estimated population frequencies converged to stable values. The final results were a set of estimated haplotype probabilities for each animal, expressed as the expected number of copies of each haplotype carried by each animal.

Genotype Analyses

Associations of 3 SNP of the *TLR4* gene with traits of interest were estimated individually, considering the genotype of SNP as a fixed effect. The following mathematical model was used and analyzed by SAS software (SAS Institute, 1999):

$$EBV_{ij} = \mu + G_i + e_{ij},$$

where EBV_{ij} is trait EBV of the *j*th animal with the *i*th genotype, μ is the overall mean, G_i is the fixed effect

Figure 2. An illustration of single nucleotide polymorphism (SNP) typing using the technique of tetra-primer Amplification Refractory Mutation System–PCR. Diagram shows genotyping of E3+1656 SNP. Four primers were used: The 2 outer primers amplify a control fragment (416 bp) of the gene that contains E3+1656 SNP. The inner primers were designed to amplify the 2 different allelic fragments of 269 and 202 bp representing the C and T alleles, respectively.

of the i th genotype of *TLR4* SNP, and e_{ij} are random residual effects.

Haplotype Analyses

Haplotype effects were estimated by regressing EBV on haplotype probabilities, which are expressed as the expected number of copies of each haplotype. Three haplotypes were rare (joint frequency = 0.02) and were therefore pooled together (i.e., HapP) into a single common effect. The following model was used for statistical analyses using SAS software (SAS Institute, 1999):

$$EBV_{ij} = \mu + \sum_{i=1}^h \beta_i \text{Hap}_{ij} + e_{ij},$$

where EBV_{ij} is trait EBV of the j th animal, μ is the overall mean, β_i is the linear regression coefficient for the i th haplotype, Hap_{ij} is the probability of the i th haplotype of the j th individual, h is the number of haplotypes evaluated, and e_{ij} are random residual effects.

RESULTS

TLR4 Promoter Sequence

The 5'-upstream promoter region of the *TLR4* gene was sequenced using the genome walker technique (Clontech). After a total of 3 walking steps, a 2.1-kbp sequence in the promoter region of *TLR4* was obtained upstream of the translation start site. This sequence was compared with recently available sequence information from GenBank and was found to be 100% identical to the *B. taurus* chromosome Un genomic contig sequence (GenBank accession no. NW_482804) obtained by whole genome shotgun sequencing and released in September 2004 by the Baylor College of Medicine Human Genome Sequencing Center.

Potential DNA binding sites within the 700-bp region of the *TLR4* promoter sequence were identified by web-based computer analysis (<http://www.motif.genome.ad.jp>). Several DNA binding sites were revealed by this analysis, including sites for c-Ets-1 (p54), AP-1, AP-4, MZF1, ADR1, GATA-1, GATA-2, GATA-3, Oct-1, and CRE-binding protein. A TATA box, CCAAT sequences, and GC-rich regions were lacking in this region (Figure 3).

SNP Detection

A total of 3 SNP were detected, including a transversion C $\leftrightarrow$ G SNP at position -226 bp in the putative promoter region (P-226) and 2 transition C $\leftrightarrow$ T SNP at position 1,656 bp in exon3 (E3+1656) and position

2,021 bp in exon3 (E3+2021; Figure 1). The novel P-226 SNP was submitted to the National Center for Biotechnology Information (accession no. ss49839679), which is released in the dbSNP Build 126. This SNP is in a binding site of c-Ets-1, MZF1, and ADR1 transcription factors (Figure 3).

Genotypic and Allelic Frequencies of *TLR4* SNP in Canadian Holstein Bulls

The genotypic frequencies of CC, CG, and GG at P-226 were observed as 58 (15%), 193 (50%), and 137 (35%), respectively. The frequency of allele C was 40%, compared with 60% of allele G. At position E3+1656, the frequency of the CC, CT, and TT genotypes were 144 (37%), 200 (52%), and 44 (11%), respectively. The C allele was predominant over the T allele (63 vs. 37%). The largest differences in genotype frequencies were shown by SNP E3+2021. For this SNP, genotype TT was found at a very low frequency (3, 1%), whereas the majority of the population was either homozygous for CC (338, 87%), or heterozygous (47, 12%). The T allele was rare compared with the C allele (7 vs. 93%).

The frequencies of genotypes were in agreement with the Hardy-Weinberg equilibrium (Falconer and Mackay, 1996) within all 3 SNP (the probabilities of the χ^2 tests for deviation from the equilibrium were $P = 0.75$, $P = 0.12$, and $P = 0.64$ for P-226, E3+1656, and E3+2021, respectively). When considering 2 SNP together, the equilibrium in genotypic frequencies was tested by a χ^2 test of expected and observed frequencies of gametic types (Falconer and Mackay, 1996). All pairwise combination tests showed a significant disequilibrium ($P < 0.01$).

Genotype Analyses

The focus of the genotype analyses was on finding associations of *TLR4* polymorphisms with health-related traits, in particular, SCS and LP. Lactation persistency was influenced by genotypes of P-226 ($P < 0.004$), E3+1656 ($P < 0.04$), and E3+2021 ($P < 0.002$). Somatic cell score was associated with the E3+2021 genotype ($P < 0.03$).

The least squares means for the P-226, E3+1656, and E3+2021 genotypes for LP and SCS are presented in Table 2. Genotypes GG and CG of P-226 were highly associated with increased LP. Similar findings were observed at E3+1656, where allele C was associated with higher LP and genotypes CC and CT had higher LP than the genotype TT. Only 3 animals had the genotype TT, for the E3+2021 SNP. Since these were excluded from the analyses, only solutions for the genotypes CC and CT were obtained. The C allele of E3+2021

_____Oct-1_____
 TGAGGTTATAGATTAGGGTGGCTCTA**ACCC**ATTATGACTTGTGTGCTTATAAGAA
 c-Ets-1_____
 _____NF-AT_____
 _____GATA-2_____
 _____GATA-2/3_____
 _____CREB_____
 AAGGAAATTAAGACCCAGCACATGCAGAGAGAAGATGACGTGAAGGCACAGGG
 _____AP-4_____
 _____STRE_____
 _____GATA_____
 AGAAGACAGCCATCTATAAGCCAAGGGGACAGCATGAAAAATTAAGCCTGCTGA
 _____CREB_____
 _____1_____
 TCCCTTGAGCACAGACGTCTAGCCTCCAGAACTTCGTGAGAAAATGAATGTTTAA
 _____c-Ets-1_____
 _____AP-1_____
 _____NF-AT_____
 GTCACCTAGTCCAGGGTATTTTGTTATGGCTGGAACATAGGAAATGAACATTTAA
 _____NF-AT_____
 CAGTAAGTTTCCACTACAAAACGAAAGCAGAAAGCCACAGTTCCTGGGTAGTCA
 _____GATA-1_____
 _____CREB_____
 CACAGGCTGTTTGACCAGTCTGATTGCTTTCCTAAATTCATAACCCTGAGGTCA
 _____c-Ets-1_____
 _____AP-1_____
 _____c-Ets-1_____
 _____GCR1_____
 _____GATA-X_____
 _____GCR1_____
 _____STRE_____
 GATGACTTCCTGGGATAAAAGCCAGCTTCCTCTTGTTGTTACTTCAGCCCCTGGT
 _____ADR1_____
 _____MZF1_____
 _____GATA_____
 _____c-Ets-1_____
 _____GATA-1_____
 CTGCAGACGTTTTCTTCTTCTTCTCCTCTAACTTCCCCTCG/CGGTGGCAGAGGAGAT
 _____-X_____
 _____STRE_____
 _____c-Ets-1_____
 AAATATCAGAGAAACGAAAGTCCAGAAAGCTAAGGCTGCCCTCTCACTTCCTC
 _____STRE_____
 TCCTCCCGTAACCCAGCACTGCTTTGAATAGGGGCAGTGCCCTGGGGCACCAAG
 _____MZF1_____
 CCGGGGAGAGACGACACTACAGTGCCTCGGAGGCTGCCCTGGCCGGCACAGACA
 _____CF1_____
 GAGGGTCATGCTTTCACAGAGCCACTTCTGGTTCACAGAAAATGCCAGG+1

Figure 3. Bovine *TLR4* 5'-flanking region sequence (–700 bp), and putative binding sites for bovine transcription factors (overlined). The start codon is localized at +1. Transversion single nucleotide polymorphism (SNP; G/C) P-226 was represented by bold letters. (Oct = octamer factor; NF-AT = nuclear factor of activated T cells; GATA = GATA-binding factor; CREB = cAMP-responsive element-binding protein; STRE = stress-response element; AP = activator protein; ADR1 = alcohol dehydrogenase gene regulator 1; GCR = glucocorticoid receptor; MZF1 = myeloid zinc finger gene 1; CF = chorion factor).

Table 2. Least squares means ($\pm$ SE) for the genotypes of single nucleotide polymorphisms (SNP) in the *TLR4* gene for lactation persistency (LP) and SCS

SNP	Genotype	Trait	
		LP (%)	SCS
P-226	CC	65.87 $\pm$ 0.38 ^b	—
	CG	67.20 $\pm$ 0.21 ^a	—
	GG	67.25 $\pm$ 0.24 ^a	—
E3+1656	CC	67.17 $\pm$ 0.24 ^a	—
	CT	67.14 $\pm$ 0.20 ^a	—
	TT	66.00 $\pm$ 0.44 ^b	—
E3+2021	CC	67.20 $\pm$ 0.16 ^a	2.98 $\pm$ 0.02 ^b
	CT	65.82 $\pm$ 0.42 ^b	3.08 $\pm$ 0.04 ^a

^{a,b}The least squares means within a column for each SNP that lack a common superscript letter differ from each other at the 5% level of significance by the Bonferroni test.

was associated with higher LP and lower SCS. This nonsynonymous SNP changes an AA from Thr (C allele) to Ile (T allele) in the predicted transmembrane–cytoplasmic domain (White et al., 2003). The estimated differences between homozygous CC and heterozygous genotypes were 1.4% and -0.1 points for LP and SCS, respectively, which corresponds to 0.52 and -0.40 SD units of the respective traits.

Haplotype Construction

All 3 SNP were used for haplotype reconstruction. The estimated frequencies of all the haplotypes are presented in Table 3. The most common haplotype among Holstein bulls was GCC (Hap3), with an estimated frequency of 0.54. More than 86% of the haplotypes present were either GCC (Hap3) or CTC (Hap4). The rest of the 6 haplotypes had a combined estimated frequency of about 14%. Two haplotypes, Hap5 and Hap6, both bearing a T allele at E3+2021, were rare ($<0.005\%$) because this allele is at a very low frequency (7%) in the population.

Haplotype Effects

Because haplotypes 5, 6, and 7 together represented only 2% of all haplotypes, they were pooled (HapP) as

Table 3. Probabilities of single nucleotide polymorphism (SNP) haplotypes (Hap) at the *TLR4* gene in Canadian Holstein bulls

Haplotype	SNP			Probability
	P-226	E3+1656	E3+2021	
Hap1	C	C	C	0.02759
Hap2	G	T	C	0.04766
Hap3	G	C	C	0.53983
Hap4	C	T	C	0.31791
Hap5	G	T	T	0.00168
Hap6	C	T	T	0.00468
Hap7	G	C	T	0.01392
Hap8	C	C	T	0.04673

Table 4. Contrast between haplotypes (Hap) and the most frequent haplotype (Hap3) for lactation persistency (LP) and SCS¹

Haplotype	Trait	
	LP (%)	SCS
Hap1	-1.98 ± 0.86 ($P = 0.02$)	0.19 ± 0.09 ($P = 0.04$)
Hap4	-0.51 ± 0.23 ($P = 0.02$)	— ($P = 0.82$)
HapP	-2.07 ± 0.78 ($P = 0.009$)	0.16 ± 0.09 ($P = 0.07$)
Hap2	-1.02 ± 0.55 ($P = 0.07$)	— ($P = 0.27$)
Hap8	-0.89 ± 0.47 ($P = 0.06$)	— ($P = 0.20$)

¹Significance is indicated by P -values in parenthesis.

a single effect for further analyses. Following this pooling, the linear effect of the 6 haplotypes was estimated. These haplotypes had effects on both LP and SCS that were also found to be associated with various SNP. The effect of the most common haplotype (Hap3) was used as a control, and rest of the haplotypes were contrasted against this haplotype.

Hap1, Hap4, and HapP were found to be significantly different from Hap3 for LP, whereas Hap1 was different for SCS also. However, Hap2 and Hap8 were not different from Hap3 for any of these traits, although differences were observed for LP at a lower level of significance (Table 4). Interestingly, Hap3 was superior in terms of both traits when compared with other haplotypes. The second most common haplotype, Hap4, which represents approximately 32% of the population, showed 0.51% less LP as compared with Hap3. Haplotype 4 was found to carry the C allele at P-226 and the T allele at E3+1656, whereas at both the positions, Hap3 showed the alternative alleles. In agreement with the genotype analyses, the G allele at P-226 was associated with increased LP as compared with the C allele. A similar finding was also obtained for the C allele at E3+1656 when compared with the T allele.

Haplotype 1 was significantly different from Hap3 for LP and SCS. For LP, to some extent the observation was in agreement with genotype analyses, since both of these haplotypes carried different alleles for P-226, which are associated with LP differences. In the genotype analyses, E3+2021 was found to be associated with SCS; however, both the haplotypes shared the same allele for this SNP. In the genotype analyses, the T allele for E3+2021 was found to be associated with higher SCS; however, Hap8 composed of the T allele at this location combined with the C allele for P-226 was not different from Hap3 in terms of SCS. The observations obtained from both of these comparisons (Hap1 vs. Hap3 and Hap8 vs. Hap3) suggest that SNP in the putative promoter region might have an effect on SCS. Haplotype P (Hap5, Hap6, and Hap7), which carries a T allele for E3+2021 with either allele for P-226, also showed an increased SCS at slightly lower significance

($P < 0.07$). The observed differences between genotype analyses and haplotype analyses for SCS cannot be compared as such because the contrast between haplotype effects estimates only the additive linear effect of alleles (Schenkel et al. 2005).

DISCUSSION

The bovine *TLR4* putative promoter region was sequenced for 2.1 kbp located upstream of the first exon of *TLR4* and screened for DNA polymorphisms. In addition, all 3 exons of bovine *TLR4* were screened for DNA polymorphisms. Three SNP were detected in these regions and were evaluated for their association with health-related traits by genotype and haplotype analyses.

Sequence analysis of the 5'-upstream *TLR4* promoter sequence revealed several potential DNA binding sites, including c-Ets-1 (p54), AP-1, AP-4, MZF1, ADR1, GATA-like binding factors, Oct-1, and CRE-binding protein. Conversely, the region lacked a TATA box, a CCAAT box sequence, or a GC-rich region. These results were compared with the promoter region of mice (GenBank accession no. AF177767) and humans (GenBank accession no. AF177765). Sequence analyses of the 5'-upstream 500 bp of mouse *TLR4* and 400 bp of human *TLR4* promoter regions also showed consensus-binding sites for the Ets family of transcription factors, octamer-binding factors, and the absence of a TATA box or a CCAAT box sequence (Rehli et al., 2000). Similar findings were also observed 600 bp 5'-upstream of the putative promoter region in porcine *TLR4* (Thomas et al., 2005).

The SNP detection was carried out by sequencing pooled DNA from bulls with extreme EBV for SCS. We successfully detected 3 SNP, 2 of which were in exon3 at locations 1,656 bp (E3+1656) and 2,021 bp (E3+2021). Both of these SNP were also reported in *Bos taurus taurus* (White et al., 2003). Exon1 and exon2 were found to be nonpolymorphic based on the sequencing of DNA pools constructed in this study. One SNP was discovered in the putative promoter region at position 266 bp 5'-upstream of the *TLR4* gene (P-226). After individual genotyping of each SNP, we observed that the frequencies of the C allele were 45, 57.5, and 90% in the high-SCS pool and 42.5, 55, and 87.5% in the low-SCS pool for P-226, E3+1656, and E3+2021, respectively. It is quite possible that other SNP remain undetected because identification was based on 2 DNA pools composed of 40 animals belonging to either high or low EBV for SCS.

In this study, P-226 SNP was found to be associated with LP, where the C allele was associated with a lower LP as compared with allele G. In fact, this SNP (G →

C) was in a potential DNA binding site either for c-Ets-1 (p54), or MZF1, or the ADR1 transcription factor (Figure 3). Haplotypes constituted by the C allele also showed higher SCS, which is an indication of higher susceptibility to clinical mastitis (Philipsson et al., 1995). Hence, we hypothesized that this SNP is functional and may have an impact on the transcriptional level of *TLR4* mRNA.

Genotype and haplotype analyses revealed that these SNP have effects on LP and SCS. Persistency of lactation generally refers to the rate of decline in daily milk yield after the peak yield. A cow with a flatter lactation curve is considered to be more persistent than a cow with the same total yield, but with a curve that decreases rapidly after the peak yield. Persistency has direct economic value because persistent cows are more efficient at utilizing roughage, suffer less metabolic stress because of high peak yield, and are thus more disease resistant (Dekkers et al., 1996; Solkner and Fuchs, 1987). It is noticeable that the most frequent haplotype (Hap3) composed of favorable alleles at all locations (GCC), was more persistent compared with the other haplotypes. The next most frequent haplotype, Hap4 (32%), was associated with low LP, which suggests that selection against this haplotype may be promising to improve this trait in Canadian Holsteins. In the studied population few of the rare haplotypes were found, possibly because these haplotypes may not have actually existed in the population, even though they were identified as plausible haplotypes for some individuals by the reconstruction program (Boettcher et al., 2004).

The association of the *TLR4* polymorphism with SCS may be attributed to the high genetic correlation between SCS and occurrence of clinical mastitis (Shook and Schutz, 1993.) Somatic cell counts are currently used as an indicator of mastitis and as a tool for selection for mastitis resistance (Koivula et al., 2005). It is also possible that the association between SNP in *TLR4* and SCS and LP may be due to linkage with other gene(s) that truly influence these traits, because linkage disequilibrium extends over large distances in the North American Holstein cattle population (Vallejo et al. 2003). *TLR4*, for example, has been mapped to the distal end of BTA8 (White et al., 2003) and QTL affecting SCS, and clinical mastitis has also been found on this chromosome. Klungland et al. (2001) found a single QTL affecting SCC localized at position 54 cM, and a QTL for clinical mastitis at position 46 cM in one Norwegian sire family. Although QTL for LP have not been identified, clinical mastitis can increase mammary epithelial cell apoptosis, thereby reducing LP (Capuco et al., 2003). Our observation that the unfavorable CT genotype of SNP E3+2021, which was associated with

high SCS and low LP, supports this hypothesis. An analysis of an independent cattle population may help to determine whether these associations are truly attributed to linkage disequilibrium (Vallejo et al., 2003).

CONCLUSIONS

In conclusion, a total of 2.1 kbp 5'-upstream putative regulatory sequence of *TLR4* was revealed. Three SNP, including one at the putative promoter region (P-226) and 2 in exon3 (E3+1656 and E3+2021) of the *TLR4* gene were detected through sequencing of extreme DNA pools established on EBV for SCS. These SNP were found to be associated with LP and SCS. After confirmation of these effects on molecular mechanisms of immune function and disease resistance underlying these traits, marker-assisted selection for P-226 and E3+1656 may be a promising strategy for improving these traits. Haplotype frequencies showed that 2 haplotypes contributed 86% in the population; however, both were significantly different for their effects on LP. These haplotypes should be further assessed for immune function and disease resistance prior to their practical application. The SNP at the 5'-upstream sequence (P-226) is in a potential DNA binding site; hence, we hypothesize that it may alter gene function. Studies in this area are warranted.

ACKNOWLEDGMENTS

The financial support of the Natural Sciences and Engineering Research Council, DairyGen, and Dairy Farmers of Ontario are gratefully acknowledged.

REFERENCES

- Arbour, N. C., E. Lorenz, B. C. Schutte, J. Zabner, J. N. Kline, M. Jones, K. Frees, J. L. Watt, and D. A. Schwartz. 2000. *TLR4* mutations are associated with endotoxin hypo-responsiveness in humans. *Nat. Genet.* 25:187–191.
- Bannerman, D. D., M. J. Paape, J. W. Lee, X. Zhao, J. C. Hope, and P. Rainard. 2004. *Escherichia coli* and *Staphylococcus aureus* elicit differential innate immune responses following intramammary infection. *Clin. Diagn. Lab. Immunol.* 11:463–472.
- Boettcher, P. J., G. Pagnacco, and A. A. Stella. 2004. Monte Carlo approach for estimation of haplotype probabilities in half-sib families. *J. Dairy Sci.* 87:4303–4310.
- Capuco, A. V., S. E. Ellis, S. A. Hale, E. Long, R. A. Erdman, X. Zhao, and M. J. Paape. 2003. Lactation persistency: Insights from mammary cell proliferation studies. *J. Anim. Sci.* 81(Suppl. 3):18–31.
- Dekkers, J. C. M., J. Jamrozik, J. H. ten Hag, L. R. Schaeffer, and A. Weersink. 1996. Genetic and economic evaluation of persistency in dairy cattle. *Interbull Bull.* 12:97–102.
- Dekkers, J. C. M., J. H. Ten Hag, and A. Weersink. 1998. Economic aspects of persistency of lactation in dairy cattle. *Livest. Prod. Sci.* 53:237–252.
- Falconer, D. S., and T. F. C. Mackay. 1996. *Introduction to Quantitative Genetics*. Longman Group, Harlow, UK.
- Fetrow, J. 2000. Mastitis: An economic consideration. Pages 3–47 in *Proc. Natl. Mastitis Counc. Mtg.*, Atlanta, GA. Natl. Mastitis Council, Verona, WI.
- Goldammer, T., H. Zerbe, A. Molenaar, H. J. Schuberth, R. M. Brunner, S. R. Kata, and H. M. Seyfert. 2004. Mastitis increases mammary mRNA abundance of β -defensin 5, Toll-like-receptor 2 (TLR2), and TLR4 but not TLR9 in cattle. *Clin. Diagn. Lab. Immunol.* 11:174–185.
- Hoshino, K., O. Takeuchi, T. Kawai, H. Sanjo, T. Ogawa, Y. Takeda, K. Takeda, and S. Akira. 1999. Cutting edge: Toll-like receptor 4 (TLR4)-deficient mice are hyporesponsive to lipopolysaccharide: Evidence for TLR4 as the *Lps* gene product. *J. Immunol.* 162:3749–3752.
- Jamrozik, G., J. Jansen, L. R. Schaeffer, and Z. Liu. 1998. Analysis of persistency of lactation calculated from a random regression test day model. *Interbull Bull.* 17:64–69. <http://www-interbull.slu.se/bulletins/bulletin17/Jamrozik.pdf> Accessed Aug. 15, 2005.
- Kelton, D., A. Godkin, D. Alves, K. Lissemore, K. Leslie, N. Smart, C. Church, and P. Meadows. 1999. Prevalence of major mastitis pathogens on Ontario Sentinel dairy farms. Pages 257–258 in *Proc. 32nd Annual Conf. Am. Assoc. Bovine Practitioners*, Nashville, TN.
- Klungland, H., A. Sabry, B. Heringstad, H. G. Olsen, L. Gomez-Raya, D. I. Vage, I. Olsaker, J. Odegard, G. Klemetsdal, N. Schulman, J. Vilkkii, J. Ruane, M. Aasland, K. Ronningen, and S. Lien. 2001. Quantitative trait loci affecting clinical mastitis and somatic cell count in dairy cattle. *Mamm. Genome* 12:837–842.
- Koivula, M., E. A. Mantysaari, E. Negussie, and T. Serenius. 2005. Genetic and phenotypic relationship among milk yield and somatic cell count before and after clinical mastitis. *J. Dairy Sci.* 88:827–833.
- Lee, C. S., F. B. P. Wooding, and P. Kemp. 1980. Identification, properties, and differential counts of cell populations using electron microscopy of dry cow secretions, colostrum, and milk from normal cows. *J. Dairy Res.* 47:39–50.
- Lien, E., and R. R. Ingalls. 2002. Toll-like receptors. *Crit. Care Med.* 30(Suppl.):S1–11.
- Lorenz, E., J. P. Mira, K. L. Frees, and D. A. Schwartz. 2002. Relevance of mutations in the TLR4 receptor in patients with gram-negative septic shock. *Arch. Intern. Med.* 162:1028–1032.
- McGuire, K., M. Jones, D. Werling, J. L. Williams, E. J. Glass, and O. Jann. 2006. Radiation hybrid mapping of all 10 characterized bovine Toll-like receptors. *Anim. Genet.* 37:47–50.
- Netea, M. G., C. Van der Graaf, A. Vonk, I. Verschuere, J. W. M. Van der Meer, and B. J. Kullberg. 2002. The role of Toll-like receptors in the defense against disseminated candidiasis. *J. Infect. Dis.* 185:1483–1489.
- Philipsson, J., G. Ral, and B. Berglund. 1995. Somatic cell count as a selection criterion for mastitis resistance in dairy cattle. *Livest. Prod. Sci.* 41:195–200.
- Poltorak, A., X. He, I. Smirnova, M. Y. Liu, C. Van Huffel, X. Du, D. Birdwell, E. Alejos, M. Silva, G. Galanos, M. Freudenberg, P. Ricciardi-Castagnoli, B. Layton, and B. Beutler. 1998. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: Mutations in *Tlr4* gene. *Science* 282:2085–2088.
- Rambeaud, M., and G. M. Pighetti. 2005. Impaired neutrophil migration associated with specific bovine CXCR2 genotypes. *Infect. Immun.* 73:4955–4959.
- Rehli, M., A. Poltorak, L. Schwarzfischer, S. W. Krause, R. Andreesen, and B. Beutler. 2000. PU.1 and interferon consensus sequence-binding protein regulate the myeloid expression of the human Toll-like receptor 4 gene. *J. Biol. Chem.* 275:9773–9781.
- Roger, T., I. Miconnet, A. L. Schiesser, H. Kai, K. Miyake, and T. Calandra. 2005. Critical role for Ets, AP-1 and GATA-like transcription factors in regulating mouse Toll-like receptor 4 (*Tlr4*) gene expression. *Biochem. J.* 387:355–365.
- Sabroe, I., R. C. Read, M. K. Whyte, D. H. Dockrell, S. N. Vogel, and S. K. Dower. 2003. Toll-like receptors in health and disease: Complex questions remain. *J. Immunol.* 171:1630–1635.
- SAS Institute. 1999. *SAS User's Guide: Statistics, Version 8 Developer's Edition*. SAS Inst., Inc., Cary, NC.

- Schenkel, F. S., S. P. Miller, X. Ye, S. S. Moore, J. D. Nkrumah, C. Li, J. Yu, I. B. Mandell, J. W. Wilton, and J. L. Williams. 2005. Association of single nucleotide polymorphisms in the leptin gene with carcass and meat quality traits of beef cattle. *J. Anim. Sci.* 83:2009–2020.
- Sharif, S., B. A. Mallard, and J. M. Sargeant. 2000. Presence of glutamine at position 74 of pocket 4 in the BoLA-DR antigen binding groove is associated with occurrence of clinical mastitis caused by *Staphylococcus* species. *Vet. Immunol. Immunopathol.* 76:231–238.
- Shook, G. E., and M. M. Schutz. 1993. Selection on somatic cell score to improve resistance to mastitis in the United States. *J. Dairy Sci.* 77:648–658.
- Solkner, J., and W. Fuchs. 1987. A comparison of different measures of persistency with special respect to variation of test-day milk yields. *Livest. Prod. Sci.* 16:305–319.
- Sordillo, L. M., and S. C. Nickerson. 1988. Morphologic changes in the bovine mammary gland during involution and lactogenesis. *Am. J. Vet. Res.* 49:1112–1120.
- Takeda, K., T. Kaisho, and S. Akira. 2003. Toll-like receptors. *Annu. Rev. Immunol.* 21:335–376.
- Thomas, A. V., A. D. Broers, H. F. Vandegaart, and D. J. Desmecht. 2005. Genomic structure, promoter analysis and expression of the porcine (*Sus scrofa*) *TLR4* gene. *Mol. Immunol.*
- Vallejo, R. L., Y. L. Li, G. W. Rogers, and M. S. Ashwell. 2003. Genetic diversity and background linkage disequilibrium in the North American Holstein cattle population. *J. Dairy Sci.* 86:4137–4147.
- White, S. N., K. H. Taylor, C. A. Abbey, C. A. Gill, and J. E. Womack. 2003. Haplotype variation in bovine Toll-like receptor 4 and computational prediction of a positively selected ligand-binding domain. *Proc. Natl. Acad. Sci. USA* 100:10364–10369.
- Winfrey, M. R., M. A. Rott, and A. T. Wortman. 1997. *Unraveling DNA: Molecular Biology for the Laboratory*. Prentice Hall, Upper Saddle River, NJ.
- Ye, S., S. Dhillon, X. Ke, A. R. Collins, and I. N. M. Day. 2001. An efficient procedure for genotyping single nucleotide polymorphisms. *Nucleic Acids Res.* 29:e88.
- Youngerman, S. M., A. M. Saxton, S. P. Oliver, and G. M. Pighetti. 2004. Association of CXCR2 polymorphisms with subclinical and clinical mastitis in dairy cattle. *J. Dairy Sci.* 87:2442–2448.