Cloning and SNP screening of the TLR4 gene and the association between its polymorphism and somatic cell score in dairy cattle

Xing-ping Wang[1], Shang-zhong Xu[2#], Xue Gao[2], Jun-ya Li[2], Hong-yan Ren[2] and Zhuo-ma Luoren[1]

[1] Department of Life Sciences & Technology, Xinxiang Medicine University, Xinxiang 453003, P.R.China*
[2] Laboratory of Molecular Biology and Bovine Breeding, Institute of Animal Science,
Chinese Academy of Agricultural Sciences, Beijing 100094, P.R. China

_____

Abstract

Toll-like receptor 4 (TLR4) is a cell-surface receptor that activates innate and adaptive immune responses. Recognizing a broad class of pathogen-associated molecular patterns (PAMPs), TLR4 is implicated as a receptor mediating cellular activation in response to bacterial lipopolysaccharide (LPS), a cell-wall component of gram-negative bacteria, which is the principal PAMP of TLR4. TLR4 is a candidate gene for resistance to a large number of diseases. Human models, in particular, suggest TLR4 to be a candidate gene associated with resistance to sepsis, immunodeficiencies, atherosclerosis and asthma. In this study the bovine TLR4 mRNA and genomic DNA were sequenced and their gene structures were determined through alignment of the genomic DNA sequence to the mRNA sequence. Its mRNA sequence consisted of a 2526-bp open reading frame (ORF) flanked by a 470-bp 5′-untranslated region (UTR) and a 743-bp 3′-UTR; and its genomic DNA is 11,013-bp in length and includes the three exons and two introns. Thirty-one single nucleotide polymorphisms (SNPs) were detected by sequencing polymerase chain reaction (PCR) products from 15 cattle, consisting of five randomly selected individuals from Chinese Simmental, Holstein and Sanhe cattle breeds each. Sixteen of these SNPs were in coding regions (cSNP). However, the 3′-untranslated region was comparatively conservative and no SNP was found in it. The SNP (g.8664C>T) found in the Toll IL-1 receptor (TIR) domain of the bovine TLR4 gene leads to amino acid substitution from Thr (C allele) to Ile (T allele). Three hundred and ninety seven cows from Chinese Simmental, Holstein and Sanhe cattle breeds were genotyped for SNP (g.8664C>T) using the method of created restriction site PCR together with restriction fragment length polymorphism (RFLP) by *Hinf* I. An association study between the SNP and somatic cell score (SCS) in the three breeds showed that the SCS of individuals with a CC genotype was significantly lower than that of the TT genotype.

_____

Introduction

Mastitis, an inflammatory disease of the mammary gland generally caused by a broad spectrum of bacterial and fungal pathogen intramammary infections, is the most frequent occurring and costly disease in the dairy industry (Sordillo & Streicher, 2002). Invading pathogens activate the immune defence system in the udder, and that is manifested as an increase in the number of somatic cells in milk. Therefore, somatic cells score (SCS) is commonly used to indicate mastitis and as a tool in the selection for mastitis resistance (Koivula *et al*., 2005; Heringstad *et al*., 2006).

The toll-like receptors (TLRs) are a family of cell-surface signalling molecules that play a fundamental role in the immune response to recognize pathogens that bind to specific pathogen-associated molecular patterns (PAMPs) (Takeda *et al*., 2003). In mammals there are at least 10 members of the TLR family that recognize specific components conserved among microorganisms (Rock *et al*., 1998; Chuang *et al*., 2000). In the case of TLR4 they have been implicated as receptors, mediating cellular activation in response to bacterial lipopolysaccharide (LPS), a cell-wall component of gram-negative bacteria, which is the principal PAMP of TLR4 (McGuire *et al*., 2005). LPS-induced apoptosis and nuclear factor-Kappa B (NF-kB) activation can occur independently, with the pathways diverging downstream signalling molecules, including the adaptor molecule MyD88, IL-1R-associated protein kinases (IRAKs) and the Tumor Necrosis Factor

Receptor Associated Factor 6 (TRAF6) (Bannerman *et al*., 2002; Akira, 2003), which can induce the over-expression of several pro-inflammatory cytokines and chemokines.

Gram-negative bacteria such as *Escherichia coli* are one of the predominant pathogens causing bovine mastitis. TLR4 is the only important pattern recognition receptor of the TLRs family that recognizes endotoxins associated with gram-negative bacterial infections (Takeda *et al*., 2003; Werling & Jungi, 2003; McGuire *et al*., 2005; Sharma *et al.*, 2006). 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 dairy cattle (Sharma *et al.*, 2006). In this study the objective was to isolate the TLR4 gene using the reverse transcription polymerase chain reaction (RT-PCR) and rapid amplification of cDNA ends (RACE), and to screen the single nucleotide polymorphisms (SNPs) by re-sequencing. In addition, the association between polymorphism of SNP (g.8664C>T) for the TLR4 gene and SCS was tested to establish if the gene could be used in molecular marker assisted selection in the selection for mastitis resistance in cows.

Materials and Methods

In this study, all cattle investigated for gene isolation, SNP screening and association analyses were subjected to the same management conditions in accordance with the guidelines established by the Canadian Council on Animal Care (CCAC) and according to national legislation.

Five fresh udder tissue samples of Chinese Simmental cows were collected for the isolation of bovine cDNA of the TLR4 gene. The blood samples for screening SNPs were obtained from 15 individuals, five each from randomly selected individuals of the Chinese Simmental, Chinese Holstein and Sanhe cattle breeds. Three hundred and ninety seven blood samples were obtained from 104 Chinese Holstein, 197 Sanhe and 96 Chinese Simmental cattle for determining the association between g.8664C>T polymorphism and SCS.

Fresh udder tissue samples were aseptically collected from slaughtered cows according to the following procedure: The washed udder skin was incised with a disinfected (70% alcohol) knife. A piece of tissue (2 x 2 x 2 cm) was removed from the deeper area of an udder quarter. A small section (0.5 x 0.5 x 0.5 cm) was preserved for total RNA extraction, using the TRIZOL reagent kit (GIBCO/BRL, Bethesda, MA, USA). The RNA was prepared for RT-PCR and Rapid amplification of cDNA ends (RACE) experiments.

To obtain the cDNA of the bovine TLR4 gene, the EST database of GenBank was searched for the TLR4 gene by BLAST (http://www.ncbi.nlm.nih.gov/blast/) and the bovine contig was assembled. Primers (Table 1) were designed with Primer Premier 5.0 (Premier Biosoft International, Palo Alto, CA, USA) based on the bovine contig, to obtain the entire open reading frame (ORF) and the untranslated region (UTR) by RT-PCR and RACE. Among these primers, the pairs of primers B, C, D for RT-PCR were designed from the bovine contig sequence of TLR4, while primers A and E for RACE were designed from the sequence obtained by RT-PCR.

The first-strand cDNA used for RT-PCR was synthesized using 1 µg of DNaseI-treated total RNA, 0.5 µg of oligo(dT)$_{18}$, 500 µmol/L of each dNTPs, 10 U of RNAsine (Promega, Madison, WI, USA), 1×murine leukaemia virus (MLV) RT buffer and 300 U of M-MLV reverse transcriptase (Promega) in a total volume of 50 µL at 37 °C for 1 h. The enzyme was then inactivated at 95 °C for 5 min.

The PCR reactions were carried out in a total volume of 20 µL which contained 1 µL of the first-strand product, 1×buffer (Tris-HCl 100 mmol/L; pH 8.3; KCl 500 mmol/L), 0.25 µmol/L of each primers, 2.0 mmol/L of MgCl$_2$, 0.25 mmol/L of dNTPs and 0.5 U of *Taq* DNA polymerase (TaKaRa, Dalian, China). The optimal condition of the PCR was 94 °C for 3 min, followed by 40 cycles of 94 °C for 30 s, annealing for 30 s (temperatures as shown in Table 1), 72 °C extension for 2 min and a final extension at 72 °C for 10 min.

Rapid amplification of cDNA ends (RACE) was performed with a BD SMART™ RACE cDNA Amplification Kit (BD Clontech, Palo Alto, CA, USA) following the manufacturer's instructions. A touch-down PCR programme was applied with the following temperature profile: 95 °C for 2 min, followed by 4 cycles of 94 °C for 30 s, 72 °C for 3 min, followed by 4 cycles of 94 °C for 30 s, annealing at 70 °C for 30 s, extension at 72 °C for 3 min; followed by 30 cycles of 94 °C for 30 s, annealing at 68 °C for 30 s, extension at 72 °C for 3 min.

All the above-mentioned PCR products were purified with a Wizard PCR Preps DNA Purification System (Promega) and cloned into the pGEM T-easy vector (Promega), and then sequenced using an ABI3730 automatic sequencer (Applied Biosystems, Forster City, CA, USA). All sequences were assembled with SeqMan™ II (Lasergene version 6; DNAStar, Inc., Madison, WI, USA).

Table 1 Primers used for reverse transcription polymerase chain reaction (RT-PCR) and rapid amplification of cDNA ends (RACE) of the cDNA sequence of the Toll-like receptor 4 (TLR4) gene

Symbol	Primer sequence		Region of PCR * (bp)	Length (bp)	Annealing temp (°C)
A	Reverse:	5'CAGGCAGGAGAGGATGGCCGTGG 3'	1-534	534	68~72
B	Forward	5' ATGCTTTCACAGAGCCACT 3'	431-1668	1238	56
	Reverse	5' GGTTGTCCCAAAATCAGTGT 3'			
C	Forward	5' AATGGATTGACTCTGCGAAG 3'	1156-2742	1587	56.5
	Reverse	5' GGTCTGGGCAATCTCATACT 3'			
D	Forward	5' CACTGTGCTCCTGGTGTCTGT 3'	2391-3114	724	55.5
	Reverse	5' GCGTACCACTGAATCACCA 3'			
E	Forward:	5' CAAGAAGCGACAACCTCCACCT 3'	2974-3740	767	68~72

*The mRNA sequence has been submitted to GenBank (accession no. DQ839566).

The amino acid sequence deduced by ORF of TLR4 cDNA was obtained by the Editseq programme (Lasergene version 6; DNA Star, Inc., Madison, WI, USA), and the TLR4 protein domain was predicted by a website tool (http://smart.embl-heidelberg.de/).

Bovine genomic DNA was extracted from five randomly collected blood samples per breed from Chinese Simmental, Chinese Holstein and Sanhe cattle breeds, using the method as described in Vidal *et al.* (2005). The DNA fragments of the TLR4 gene were amplified from bovine genomic DNA using primers derived from the sequence obtained by RT-PCR and RACE. Starting from the sequenced fragment, the primers were designed to amplify five overlapping genomic fragments for the single nucleotide polymorphism (SNP) search (Table 2). PCR reactions were carried out in a volume of 20 μL which contained 50 ng of bovine genomic DNA, 1×buffer (100 mmol/L of Tris-HCl, pH 8.3; 500 mmol/L of KCl), 0.8 μmol/L of each primers, 2.0 mmol/L of $MgCl_2$, 0.25 mmol/L of each dNTPs and 2 U *Taq* DNA polymerase (TaKaRa, Dalian, China). The PCR optimal condition was 94 °C for 5 min followed by 40 cycles of 60 s at 94 °C, 30 s annealing (temperatures for each of the primer sets as shown in Table 2), 10 min extension at 72 °C, and a final extension of 10 min at 72 °C.

A PCR reaction was performed twice for each individual. Each PCR product was used for sequencing in the forward and reverse directions in order to reduce the risk of reporting PCR artifacts as polymorphisms. The sequences were assembled per individual for DNA sequence with SeqMan™ II (Lasergene version 6; DNA Star, Inc., Madison, WI, USA), and were aligned for search of SNPs with Megalign (Lasergene version 6; DNA Star, Inc., Madison, WI, USA).

Jugular blood samples (10 mL per dairy cattle) were collected from 104 Chinese Holstein, 197 Sanhe cattle and 96 Chinese Simmental, using acid citrate dextrose as an anticoagulant. Within breeds these dairy cattle were half-sibs from one sire, and were chosen at random. Genomic DNA was extracted from blood samples as described in Vidal *et al.* (2005), and kept at –20 °C.

The fragment containing a mutation (g.9788C>T) of the TLR4 gene was amplified with genomic DNA using created restriction site PCR (combination of the 25th base mismatch of forward primer with SNP was used to introduce a restriction site after PCR, which formed restriction site for endonuclease *Hinf* I (G^ANTC) in order to identify the SNPs). The primers were as follows: Forward 5'-CAAAAAGTATGGCAGGGGCG AGAGAA-3', reverse 5'-ATGAAGTGCTGGGACACCACGACAAT-3'. The PCR product was identified on 1.5% agarose gel (Figure 1). PCR reactions were carried out in a volume of 20 μL which contained 50 ng bovine genomic DNA, 1×buffer (100 mmol/L of Tris-HCl, pH 8.3; 500 mmol/L of KCl), 0.3 μmol/L of each primers, 2.0 mmol/L of $MgCl_2$, 0.25 mmol/L of each dNTPs and 0.5 U of *Taq* DNA polymerase (TaKaRa, Dalian, China). The optimum condition of PCR was 94 °C for 5 min followed by 36 cycles of 30 s at 94 °C, 30 s annealing at 57 °C, 30 s extension at 72 °C, and a final extension of 10 min at 72 °C. The fragment was digested at a constant temperature (37 °C) for 5 h in a 10 μL reaction containing 1×buffer (50 mmol/L of Tris-HCl, pH 7.5; 10 mmol/L of $MgCl_2$; 1 mmol/L of Dithiothreitol; 100

mmol/L of NaCl), 4 U of *Hinf* I (TaKaRa, Dalian, China) and 300 ng of PCR product. Then the mixture after digestion was visualized by 12% polyacrylamide gel electrophoresis at the following conditions: 200 V, 20 mA, 2.5 h. The gels were stained by Silver Stain. The patterns of DNA bands were observed and photographed (Figure 2).

Table 2 Primers used for isolation of DNA sequence of the Toll-like receptor 4 (TLR4) gene*

Location	Primer sequence	Region of PCR *(bp)	Length (bp)	Annealing temp (°C)
5' Flanking region	Forward:5' GGGTATTTTGTTATGGCTGG 3' Reverse: 5' CCATCATCCTGGCATTTT 3'	1-477	477	54.5
Exon1, Intron 1	Forward: 5' ATGCTTTCACAGAGCCACT 3' Reverse: 5' CCTGAAGAAGGGAGATAGC 3'	430-5449	5020	55
Intron 2	Forward: 5' AACCAAGATGCTGGACCTGA 3' Reverse: 5' GTTTCCCGTCAGTATCAAGG 3'	5183-8103	2931	56
Exon3	Forward:5' CCTTGATACTGACGGGAAAC 3' Reverse: 5' GCGTACCACTGAATCACCA 3'	8084-10410	2327	57
3'Flanking region	Forward:5'TCATCAGTGTGTCGGTGGTCA 3' Reverse:5'GGGGAATCTGAATCAGGAGTC 3'	9668-10788	1121	60

*The DNA sequence has been submitted to GenBank (accession no. DQ839567).

Mastitis information, consisting of somatic cell counts (SCC) and the number of antibiotic treatments during a 305 d lactation from 2003 to 2007, was obtained for each cow from DHI SCC reports and individual cow records of Chinese Holstein, Sanhe cattle and Chinese Simmental maintained at the farm. For association analysis, SCC was converted to linear SCS (SCS = $\log_2$SCC/100000+3) (Brown *et al*., 1986), and the means and standard deviations of SCS were calculated using SAS software (SAS, 1999).

Genotypic and allelic frequencies, Hardy-Weinberg equilibrium and polymorphism information content (PIC) were calculated with software of POPGENE (Ver. 1.31). The significance of differences in allele and genotype frequencies was determined by χ^2 test. Analysis of associations between genotypes of TLR4 and SCS reflecting on mastitis traits (Deluyker *et al*., 1993; Heringstad *et al*., 2006) was carried out by a general linear model (GLM) procedure of SAS (1999) for the following model:

$$y_{ijkh}=\mu+ B_i + A_j +F_k+ G_h +(BAFG)_{ijkh} + e_{ijkh}$$

Where, y_{ijkh} is phenotypic value of SCS from SCC of the cattle, μ is population mean, B_i is fixed effect of the breed (i = 1, 2, 3; 1 = Chinese Holstein, 2 = Sahne, 3 = Chinese Simmental), A_j is effect of age (j = 3, 4, 5 in years), F_k = parities (k = 1, 2, 3), G_h is the fixed effect of the genotype of the TLR4 gene, $(BAFG)_{ijkh}$ is interaction between the main effects of the model and e_{ijkh} is random error effect of each observation.

Results

The cDNA fragments, which were amplified by RT-PCR and RACE with bovine udder RNA as template, were sequenced. All sequences were assembled using SeqMan™ II, to obtain the bovine cDNA sequence (GenBank accession no. DQ839566). The bovine TLR4 mRNA contained a 5′-untranslated region (1-470 bp), a long open reading frame (471-2996 bp) and a 3′-untranslated region (2997-3739 bp).

The amino acid sequence inferred by ORF of TLR4 cDNA was analyzed, and the character and protein domain were predicted by bio-informatics. The results showed that the bovine TLR4 protein consisted of 841 amino acids, of which 1 - 23 amino acids were signal peptide, 24 - 634 amino acids were outside the membrane, 635 - 657 amino acids were transmembrane helices structure and 658 - 841 amino acids were inside the membrane. The TLR4 protein contained LRR (leucine-rich repeat), LRR-TYP (leucine-rich repeats, typical subfamily), LRRCT (leucine-rich repeat C-terminal domain) and TIR (Toll-interleukin1-resistance domain) domain. These were similar to the result of another study (White *et al.*, 2003).

The DNA fragments of bovine TLR4 were amplified and assembled in a DNA sequence (GenBank accession no. DQ839567), which has 11,013-bp long including 3 exons and 2 introns. The exon/intron

boundaries were determined through alignment of bovine TLR4 genomic DNA sequence to its cDNA sequence according to the GT-AG rule.

Sequence alignment of the PCR products from five randomly selected individuals of three different breeds revealed 31 SNPs (Table 3), 12 of which were novel SNPs, but no SNP was found in the 3′-untranslated region.

Table 3 List of single nucleotide polymorphisms (SNPs) observed in the bovine Toll-like receptor 4 (TLR4) gene

Location	DNA sequence variants*	Genbank database SNP ID	Amino acids substitution
5' Flanking region	g.245C>G	rs29017188	-
5' Flanking region	g.429T>C	unknown	-
Exon 1	g.480C>T	rs8193041	p.Cys4Arg
Exon 1	g.507C>T	unknown	p.Pro13Ser
Exon 1	g.534A>C	rs8193042	-
Exon 1	g.545T>C	rs8193044	-
Intron 1	g.610C>T	rs43578094	-
Intron 1	g.1196A>G	rs43578096	-
Intron 1	g.2256C>T	unknown	-
Intron 1	g.2320C>T	unknown	-
Intron 1	g.2462C>G	unknown	-
Intron 1	g.2618C>T	unknown	-
Intron 1	g.2625C>T	unknown	-
Intron 1	g.4597 del A	unknown	-
Intron 1	g.4669C>T	unknown	-
Intron 1	g.4986A>T	unknown	-
Intron 1	g.5088A>G	rs8193046	-
Exon 2	g.5166G>A	rs8193048	p.Asp50Asn
Intron 2	g.7201C>T	unknown	-
Intron 2	g.8000G>A	rs43578100	-
Exon 3	g. 8046C>T	unknown	-
Exon 3	g.8481G>C	rs8193050	p.Lys238Asn
Exon 3	g.8664C>T	rs8193052	-
Exon 3	g.8886A>G	rs8193054	-
Exon 3	g.8934T>G	rs8193057	-
Exon 3	g.9288A>G	rs8193059	-
Exon 3	g.9423T>C	rs8193060	-
Exon 3	g.9594T>C	unknown	-
Exon 3	g.9642C>T	rs8193064	-
Exon 3	g.9715G>A	rs8193065	-
Exon 3	g.9788C>T	rs8193069	p.Thr674Ile

*The DNA sequence has been submitted to GenBank (accession No. DQ839567), the symbol (-) means there was no amino acids substitution induced by this SNP.

Of 31 SNPs, g.9788C>T was included in the Toll IL-1 receptor (TIR) domain of TLR4, and it substitutes Thr for Ile, which may be very important for the function of the TLR4 gene. Therefore, this SNP was used for an association analysis with SCS. Created restriction site PCR was used to genotype all 104 Chinese Holstein, 197 Sanhe cattle and 96 Chinese Simmental for g.9788C>T. The 234 bp fragment of PCR was obtained (Figure 1), in which the mismatched A and T allele formed the restriction site (G^ANTC) which *Hinf* I can cut into two fragments (24 and 210 bp), but the mismatched A and C allele formed GANCC which *Hinf* I cannot cut (Figure 2). The allelic distribution clearly revealed a much higher frequency of the C
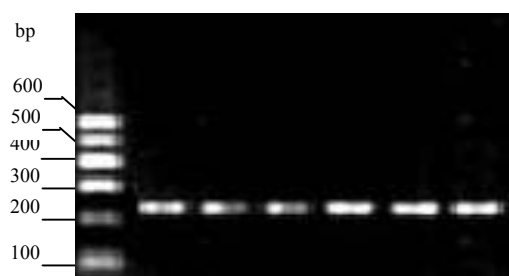
Figure 1 Image of PCR product including the SNP (g.9788C>T) for TLR4 gene about 234 bp detected on 1.5% agarose gel.
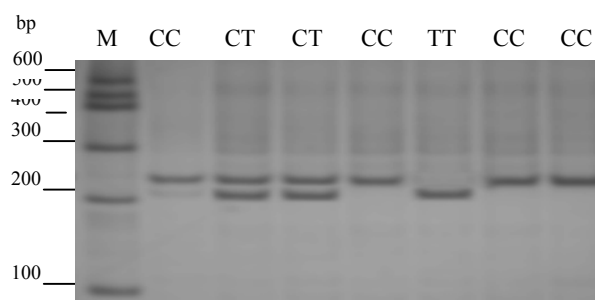

Figure 2 Image of the SNP (g.9788C>T) genotyping using created restriction site PCR together with RFLP. After the PCR product was digested by *Hinf* I and visualized on 12% polyacrylamide gel, the two different allelic fragments of 234 and 210 bp representing the C and T alleles, respectively.

Table 4 Genotypic frequencies, allelic frequencies and Hardy-Weinberg equilibrium test of single nucleotide polymorphisms (SNP) (g.9788C>T) in a population of three breeds

Breeds	Number	Genotypic frequency			Allelic frequency		χ^2
		CC	CT	TT	C	T	
Chinese Holstein	104	0.6538 (68)	0.3173 (33)	0.0288 (3)	0.8125	0.1875	0.4263
Sanhe cattle	197	0.5888 (116)	0.3959 (78)	0.0152 (3)	0.7868	0.2132	7.0888 [*]
Chinese Simmental	96	0.7500 (72)	0.2292 (22)	0.0280 (2)	0.8646	0.1354	0.0495

* means where χ^2 values differ at P <0.05.

allele in all three cattle breeds (Table 4). The frequencies of genotypes were in agreement with the Hardy-Weinberg equilibrium for g.9788C>T in Chinese Holstein and Chinese Simmental. Statistical results of the χ^2 test indicated that the polymorphism in Sanhe cattle was not in Hardy-Weinberg equilibrium (P <0.05).

The means and standard deviations for the phenotypes (SCS) investigated in the three populations were obtained using the method mentioned above, and the results are shown in Table 5, This indicates that the SCS of Chinese Holstein were the highest and that of Chinese Simmental the lowest of these three breeds. The effect of polymorphism g.9788C>T of the TLR4 gene on SCS was also analyzed using the GLM procedure of SAS (1999) described in the above section of association analysis. The results are shown in Table 6, which indicates that the animals in all three breeds with a CC genotype have lower SCS than animals of genotype TT (P <0.05).

Table 5 Means and their standard deviations (s.d.) for somatic cell score (SCS) investigated in different populations

Breeds	Number of datasets	Mean	s.d.
Chinese Holstein	2496	5.37	0.39
Sanhe cattle	4728	3.42	0.42
Chinese Simmental	2304	3.27	0.40

Table 6 Effect of different genotypes on somatic cell score (SCS)

Genotypes	Chinese Holstein	Sanhe cattle	Chinese Simmental	Total population
CC	4.08 [a] ± 0.19	3.28 [a] ± 0.23	3.17 [a] ± 0.40	3.45 [a] ± 0.18
CT	5.41 [ab] ± 0.52	4.38 [ab] ± 0.36	3.82 [ab] ± 0.31	4.68 [ab] ± 0.27
TT	6.52 [b] ± 0.67	5.50 [b] ± 0.29	5.01 [b] ± 0.37	5.74 [b] ± 0.99

Note: within columns means with different superscripts indicate significant differences at $P < 0.05$.

Discussion

Human TLR4 plays an important role in recognizing the conserved patterns unique to microbial surfaces (PAMPs), and LPS stimulates the expression of the bovine tracheal antibacterial peptide-encoding gene via the activation of NF-kB (Diamond *et al*., 2000; Ghosh & Karin, 2002), and induces inflammation that contributes to an immune response. It was proven that the TLR4 gene affects several diseases, including sepsis, immunodeficiencies, atherosclerosis and asthma (Donald *et al*., 2004).

The bovine TLR4 gene has been mapped to the distal end of BTA8 (White *et al*., 2003) and quantitative trait loci (QTL) affecting SCS and clinical mastitis have 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. There are, however, few references reporting gene sequence of the bovine TLR4 gene and characteristic of its coding protein.

The EST database was used to assemble the cDNA contig and designed the primers for amplification of the fragments using RT-PCR and RACE, and obtained the entire cDNA of the bovine TLR4 gene. The structure of the bovine TLR4 gene is very complex. It contains 5′-UTR, a complete coding sequence and 3′-UTR. The TLR4 gene encodes 841 amino acids. The protein domains include (5′ to 3′): a signal sequence (residue 1-23), putative coreceptor-binding region 1 (24-273), a putative ligand-binding region (274-368), a putative coreceptor-binding region 2 (369-632), a transmembrane region (633-653), a proximal cytoplasmic region (654-672), a Toll IL-1 receptor (TIR) domain (673-819) and a distal cytoplasmic region (820-841) (White *et al*., 2003). Among these regions, both putative coreceptor-binding regions are very important regions where the pathogen is recognized. The TIR domain of TLR4 interacts first with the TIR of MyD88 and then triggers the downstream signalling responses, then activates the immune system (Werling & Jungi, 2003; White *et al*., 2003). Alignments of the bovine TLR4 coding sequence (CDS) and human TLR4 CDS reveal a high similarity of 85%, and their corresponding proteins are 76% similar. The results of cDNA sequence analysis and protein domain of TLR4 inferred that bovine TLR4 might induce immune response for disease resistance of pathogen infection by the signal transduction pathway.

As in human and mice, TLR4 is a highly polymorphic gene (Smirnova *et al*., 2000; 2001). More than 40 SNPs of the bovine TLR4 gene have been found until now. This gives an average of 1 SNP per 90 bp for bovine TLR4. In our sample of 15 individuals from three breeds, 31 SNPs were observed, 12 of which are novel SNPs and five which resulted from the substitution of amino acids (Table 3). The SNP (g.9788C>T) in the Toll IL-1 receptor (TIR) domain of the TLR4 gene resulted in an amino acid change from Thr to Ile, and has a high genetic variation.

The literature on SNPs in the TLR4 gene and their association with mastitis related traits (SCS) is limited. The only reference is that three SNPs, named P-226 (the same as rs29017188 in Table 3), E3+1656, and E3+2021, were detected by sequencing pooled DNA from Canadian Holstein bulls with extreme estimated breeding values (EBV) for SCS (Sharma *et al*., 2006). These results showed that an unfavourable CT genotype of the last SNP is associated with high SCS. In the present study the distribution of the g.9788C>T polymorphism was not in Hardy-Weinberg equilibrium in Sanhe cattle, whereas it was in the Chinese Holstein and Chinese Simmental populations. This might reveal that artificial selection had a profound effect on the shaping of genetic diversities and their genetic differentiation in domestic animals. This is especially the situation in Sanhe cattle, because the Sanhe is a local breed that has been developed recently compared to the other two breeds. Furthermore, the Sanhe has been subjected to a high level of artificial selection.

The statistical results of association between polymorphism and SCS showed that in all three breeds the SCS of individuals with a CC genotype were significantly lower than those with the TT genotype (P <0.05). The C allele may be the favourable allele for mastitis resistance. A mutation of g.9788C>T changed amino acid Thr to Ile in the TIR domain of TLR4, which changes the protein conformation of TLR4. We inferred that the mutation of g.9788C>T may weaken the efficiency of downstream signal transduction when pathogen infected the body, so that SCS in cattle increases. Therefore, SNP of g.9788C>T may be used in future to construct a bovine haplotype map applied in the new strategy of QTL mapping and it may be regarded as a molecular marker, and guide breeders to select dairy cattle for resistance against mastitis at an early age.

Acknowledgments

References

Akira, S., 2003. Toll-like Receptor Signaling. J. Biol. Chem. 278, 38105-38108.

Bannerman, D.D., Tupper, J.C., Erwert, R.D., Winn, R.K. & Harlan, J.M., 2002. Divergence of bacterial lipopolysaccharide pro-apoptotic signaling downstream of IRAK-1. J. Biol. Chem. 277, 8048-8053.

Brown, C.A., Rischette, S.J. & Schultz, L.H., 1986. Relationship of milking rate to somatic cell count. J. Dairy Sci. 69, 850-854.

Chuang, T.H. & Ulevitch, R.J., 2000. Cloning and characterization of a sub-family of human Toll-like receptors: hTLR7, hTLR8 and hTLR9. Eur. Cytokine. Netw. 11, 372-378.

Deluyker, H.A., Gay, J.M. & Weaver, L.D., 1993. Interrelationships of somatic cell count, mastitis, and milk yield in a low somatic cell count herd. J. Dairy Sci. 76, 3445-3452.

Diamond, G., Kaiser, V., Rhodes, J., Russell, J.P. & Bevins, C. L., 2000. Transcriptional regulation of beta-defencing gene expression in tracheal epithelial cells. Infect. Immun. 68, 113-119.

Donald, N.C., David, S.P. & David, A.S., 2004. Toll-like receptors in the pathogenesis of human disease. Nat. Immunol. 10, 975-979.

Ghosh, S. & Karin, M., 2002. Missing pieces in the NF-kB puzzle. Cell. 109 (Suppl.), 81-96.

Heringstad, B., Gianola, D., Chang, Y.M., Odegard, J. & Klemetsdal, G., 2006. Genetic associations between clinical mastitis and somatic cell score in early first-lactation cows. J. Dairy. Sci. 89, 2236-2244.

Klungland, H., Sabry, A., Heringstad, B., Olsen, H.G., Gomez-Raya, L., Vage, D.I., Olsaker, I., Odegard, J., Klemetsdal, G., Schulman, N., Vilkki, J., Ruane, J., Aasland, M., Ronningen, K. & Lien, S., 2001. Quantitative trait loci affecting clinical mastitis and somatic cell count in dairy cattle. Mamm. Genome 12, 837-842.

Koivula, M., Mantysaari, E.A., Negussie, E. & Serenius, T., 2005. Genetic and phenotypic relationship among milk yield and somatic cell count before and after clinical mastitis. J. Dairy. Sci. 88, 827-833.

McGuire, K., Jones, M., Werling D., Williams, J.L., Glass, E.J. & Jann, O., 2005. Radiation hybrid mapping of all 10 characterized bovine Toll-like receptors. Anim. Genet. 37, 47-50.

Rock, F.L., Hardiman,G., Timans, J.C., Kastelein, R.A. & Bazan, J.F., 1998. A family of human receptors structurally related to Drosophila Toll-like receptors. Proc. Natl. Acad. Sci. 95, 588-593.

SAS, 1999. Statistical Analysis System user's guide, Ver. 8.2. SAS Institute Inc., Cary, N.C., USA.

Sharma, B.S., Leyva, I., Schenkel, F. & Karrow, N.A., 2006. Association of toll-like receptor 4 polymorphisms with somatic cell score and lactation persistency in Holstein bulls. J. Dairy. Sci. 89, 3626-3635.

Smirnova, I., Hamblin, M., McBride, C., Beutler, B. & Dirienzo, A., 2001. Excess of rare amino acid polymorphisms in the Toll-like receptor 4 in humans. Genetics 158, 1657-1664.

Smirnova, I., Poltorak, A., Chan, E.K., McBride, C. & Beutler, B., 2000. Phylogenetic variation and polymorphism at the toll-like receptor 4 locus (TLR4). Genome Biol. 1, 1-10.

Sordillo, L.M. & Streicher, K.L., 2002. Mammary gland immunity and mastitis susceptibility. J. Mamm. Gland Biol. Neoplasia. 55, 135-146.

Takeda, K., Kaisho, T. & Akira, S., 2003. Toll-like receptors. Ann. Rev. Immunol. 21, 335-376.

Vidal, O., Noguera, J.L., Amills, M., Varona, L., Gil, M., Jimenez, N., Davalos, G., Folch, J.M. & Sanchez, A., 2005. Identification of carcass and meat quality quantitative trait loci in a Landrace pig population selected for growth and leanness. J. Anim. Sci. 83, 293-300.

Werling, D. & Jungi, T.W., 2003. Toll-like receptors linking innate and adaptive immune response. Vet. Immunol. Immunopathol. 91, 1-12.

White, S.N., Taylor, K.H. Abbey, C.A., Gill, C.A. & Womack, J.E., 2003. Haplotype variation in bovine Toll-like receptor 4 and computational prediction of a positively selected ligand-binding domain. Proc. Natl. Acad. Sci. 100, 10364-10369.