

Identifying Polymorphisms in Bovine Interleukin-6

In partial fulfillment of the
requirements for the
CASNR Honors Program

Richard Trent Blake
Major: Animal Science/Pre-Veterinary Medicine
Department of Animal Science

Dr. Gina Pighetti
Professor, Faculty Mentor

May 2009

Abstract

Mastitis continues to plague the dairy industry even though significant steps have been taken to reduce the occurrence of this disease. Artificial selection to choose dairy cows less susceptible to mastitis has been utilized, though heritability remains low. Marker genes allow expedition of the selection process by studying loci that can be selected for or against a particular disease. As a potential marker for mastitis, Interleukin-6 (IL-6) induces antibody production, manufacture of acute phase proteins in response to inflammation, regulation of leukocyte populations, and stimulation of T cells. Single nucleotide polymorphisms (SNPs) in this immune system gene may make bacterial colonization and susceptibility to mastitis more or less favorable. The objective of this study was to sequence the bovine IL-6 gene and identify SNPs if present. To cover the entire length of the gene, specific overlapping primers were generated using the whole bovine genome shotgun sequence with reference number NW_001494874.2 as a reference sequence. To expedite DNA sequencing and polymorphism detection, two pools of isolated DNA from ten cows each were prepared with unrelated pedigrees as possible to paternal and maternal grandsire generations. Chromatogram analysis revealed a SNP within intron four, 152 bp from the start of exon five with exon five being +472 bp from the transcription start site. Conforming to proper naming guidelines, the transition SNP was named c.+472-152C>T. Specific primers were then generated to genotype cows at the c.+472-152C>T transition SNP using Tetra amplification refractory mutation system (ARMS). Genotype analysis via 2% gel electrophoresis found that the CC genotype had a frequency of 35%, CT genotype 65%, and TT genotype not represented within the sample population. The C allele was dominant in the population with a frequency of 67.5% with the T allele experiencing a frequency of 32.5%. Genotype frequencies depicted by the sample population failed to conform to Mendelian genetics assuming if two heterozygotes were crossed or a heterozygote and homozygote crossing. With the C allele being the highest frequency, a SNP may have occurred recently in the sample population with the T allele being new or the C allele has become more favorable within the sample population. As many studies have associated intronic SNPs with disease, further investigation of IL-6 expression in CC, CT, and TT cow genotypes in a larger population may be valid to determine whether cows with a certain genotype are more susceptible to mastitis or other diseases and assist in determining the reason for genotype frequencies lacking to follow Mendelian genetics.

Introduction

Mastitis, an inflammation of the mammary gland, continues to cost the dairy industry with estimations stemming from 1.2 to \$1.7 billion a year for the dairy producer (Shim et al., 2004). Several steps have been taken to prevent mastitis or lessen its severity such as vaccination programs, pre and post-milking teat disinfection, improved housing facilities and milking equipment, and antibiotic therapy during the lactation and dry periods (Lievaart et al., 2007; LeBlanc et al., 2006). Though all these practices have improved animal health some cows still remain susceptible. The ability for microorganisms to evade the cow's natural defenses and immune response is primary in establishing infection within the mammary gland. If a gene involved in the immune response is altered, a more favorable environment for bacterial colonization can occur. The possibility of studying marker genes has made it possible to determine if correlation exists between altered immune system genes and mastitis (Zhao and Lacasse 2008).

Marker genes make it possible to isolate particular alleles on loci among organisms and attribute them to phenotype variation (Griffiths et al., 2008). Particularly changes in certain single base sequences in the genome called single nucleotide polymorphisms (SNPs) have been associated with subclinical and clinical mastitis (Youngerman et al., 2004). SNPs can alter the expression and function of the gene and gene products associated with the SNP. Genetic markers and SNPs are useful in the dairy industry as producers can select from genetic lines that exclude these markers for mastitis. To study the virulence of mastitis further, the immune response to the disease must be studied at the molecular level. Studying genes of the immune system is ideal since the ability for this system to respond to disease has an ultimate outcome on the course of infection (Fraile-Diez et al., 2003). Cytokines, which are crucial components of the immune system, serve as inflammation mediators and cellular activators during innate and specific immune responses. Interleukins (IL), a group of cytokines, are secreted by macrophages, dendritic cells, and lymphocytes in response to an infection and can have extensive effects (Solomon et al., 2006; Dienz and Rincon 2009). Particularly, IL-6, an interleukin with multiple effects, assists in inducing antibody production, acute phase proteins in response to inflammation, regulating leukocyte populations, and T cell stimulation (Dienz and Rincon 2009). If SNPs are present within the IL-6 gene, this could alter the expression and function of the gene. Marker assisted selection could then be utilized in dairy cattle populations to exclude these animals with SNPs from breeding populations. As a preliminary study, the bovine IL-6 gene will be sequenced and SNPs identified if present.

Materials and Methods

Cow selection

Twenty Holstein dairy cows were chosen from the University of Tennessee, East Tennessee Research and Education Center (ETREC; Knoxville, TN). To maximize genetic diversity, cows were chosen with unrelated pedigrees to the paternal and maternal grandsire generation. Thirty percent of selected cows shared a common paternal grandsire with 35% sharing a common maternal grandsire. Pedigrees were generated by entering the cow's sire and maternal name code into the USDA Animal Improvement Programs Laboratory database (<http://aipl.arsusda.gov/>).

DNA isolation

DNA was isolated from the same twenty Holstein cows' whole blood at the ETREC Dairy Farm using UltraClean Bloodspin Kit (Mo Bio Laboratories, Inc., Carlsbad, CA). A 5.2 mg/ml Proteinase K solution (lyophilized Proteinase K reconstituted with 1.2 ml of distilled water) was first prepared. Two hundred μ l of whole blood was added to 20 μ l of Proteinase K in a two ml tube. Two hundred μ l of solution B1 (Guanidine HCl/Tween solution) was then added to the two ml tube and bump vortexed for 15 seconds followed by incubation at 65°C for 10 minutes. Following incubation, 200 μ l of solution B2 (100% ethanol) was added and vortexed for 15 seconds. The mixture was then transferred to a spin filter and centrifuged for one minute at 13,000 x g. The spin filter was then transferred to a new two ml tube following centrifugation. Next, 500 μ l of solution B3 (Ethanol/Guanidine HCl) was added to the spin filter followed by a 30 second centrifugation step at 13,000 x g. The spin filter was removed, flow through discarded,

and spin filter replaced. Five hundred μl of solution B4 (Ethanol/ Tris/ NaCl) was added to the spin filter and centrifuged for 30 seconds at 13,000 x g. Following centrifugation, the spin filter was removed, flow through discarded, and spin filter replaced. Centrifugation was repeated for 30 seconds at 13,000 x g. Following centrifugation, the spin filter was removed and transferred to a new two ml tube without contacting the wash. To the new two ml tube, 200 μl of solution B5 (10 mM Tris-HCl) was added. The tube was then incubated at 65°C for five minutes followed by centrifugation for one minute at 13,000 x g to elute the DNA from the spin filter. Following centrifugation, the spin filter was removed and lid closed.

DNA quantification

Isolated DNA samples were quantified using a Thermo Scientific NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA). One μl samples were loaded separately for each respective cow's isolated DNA and assayed for ratios of nucleic acid to protein absorbance (A260:A280), nucleic acid to ethanol absorbance (A260:A230), and nucleic acid concentration, measured in ng/ μl . An acceptable nucleic acid to protein absorbance ratio ranged approximately from 1.7 to 2.0. An acceptable concentration of nucleic acid to be sent for DNA sequencing was greater than 10.0 ng/ μl . Isolated DNA samples were quantified prior to use in PCR and DNA sequencing.

Primer Design

Specific primers were designed to amplify the IL-6 gene using Primer3 (<http://frodo.wi.mit.edu/>; NW_001494874.2). The primers were overlapping to cover the length of the gene. To verify forward and reverse sequence primers were specific, they were compared to other bovine DNA sequences using the basic align search tool (BLAST) option from NCBI (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) by input of a DNA reference sequence. Specific forward and reverse primers were selected on the basis of the closest primer melting temperatures. Shown in Table 1 are sequencing primers for the IL-6 gene. One SNP was detected at amplicon E +519 and primers generated to genotype individual cow DNA at this locus (c.+472-152C>T) as shown in Table 2. Genotyping of the same 20 Holstein dairy cows were conducted using Tetra ARMS primers generated using a publicly available software program (Ye et al., 2001).

Table1. Primer sequences for amplicons of the IL-6 gene

Amplicon	Primer Sequences	Annealing Temperature(°C)	Amplicon bp Size
A	Forward: 5'-GCCCTCCAGGAACAGCTATG-3' Reverse: 5'-CTCCCAGTTGCCTCTCCTGT-3'	57.8 59.0	985
B	Forward: 5'-TAGGTTATGCCAGCCCAGAG-3' Reverse: 5'-AAGTTGGCCTCCTCATGAAA-3'	56.5 54.3	958
C	Forward: 5'-CGGTTCTGATATTCTTTCCC-3' Reverse: 5'-GATCTGGATCAGTGTCTGA-3'	50.5 51.3	1,001
D	Forward: 5'-CTGTCTTTTCTCTTAGGCCGA-3' Reverse: 5'-ATTAAGAGTCAGCTGCTCTG-3'	51.8 51.9	1,001
E	Forward: 5'-GGTGCAGACACCTTCACCTA-3' Reverse: 5'-TAAGTTGTGTGCCCAGTGGA-3'	56.8 56.5	904

Table2. Primers used in amplification and genotyping Holstein dairy cows at IL-6 c.+472-152C>T SNP via 2% gel electrophoresis

Tetra ARMS Primer Sequence	Annealing Temperature(°C)	Amplicon bp Size
Forward inner primer C allele: 5'-GGGCTCAGAGCAGAGACCTCCCACC-3'	67.8	225
Reverse inner primer T allele: 5'-GCCACTGGCCTTGACTCGCCCAGCTA-3'	68.2	255
Forward outer ARMS primer: 5'-AGGCCCCCGAAGAACCATTAAATGCCT-3'	65.3	428(from both outer primers)
Reverse outer ARMS primer: 5'-TCCAGCAGGTCAGTGTGTGGCTGGAG-3'	65.6	

PCR

Selection of pools and preparation of master mix

To expedite DNA sequencing and SNP detection, two pools of DNA with a final concentration of ten ng/μl from ten cows each were prepared and used as the DNA template in PCR. A standard master mix of PCR ingredients was then prepared. The master mix formulation was standard for one DNA sample template and adjusted according to the amount of DNA samples used. For one PCR reaction tube, 31 μl of RNase/DNase free water (Gibco Invitrogen, Carlsbad, CA), ten μl of 5X green *Go Taq* reaction buffer (pH 8.5, 1.5 μM), four μl of 25 μM MgCl₂, one μl of ten μM dNTPs (Promega Corp., Madison, WI), one μl of each respective amplicon's forward and reverse primer (Integrated DNA Technologies, Coralville, IA), and 0.25 μl of *Go Taq* DNA polymerase (Promega Corporation, Madison, WI) was prepared in a single 500 ml tube. Two μl of isolated pooled DNA template was aliquotted into a PCR reaction tube and the master mix transferred.

Loading of samples into thermocycler

The mixture of pooled DNA template and master mix was then loaded into a PCR thermocycler (Eppendorf Corp., Hamburg, Germany). Polymerase chain reaction consisted of an initial 95 °C melt for two minutes, followed by 40 cycles of 95 °C for 15 seconds, annealing temperatures of 57 °C (amplicon A), 54 °C (amplicons B, C, D), 56 °C (amplicon E) for 15 seconds, and 72 °C for 60 seconds followed by a final step of 72 °C for ten minutes.

Verification of primer specificity via gel electrophoresis

Primer specificity for amplified DNA from PCR was determined via gel electrophoresis. A 1.4% agarose gel was prepared by adding 0.75 g of agarose (Fisher Scientific, Hampton, MA) to 50 ml of Tris 1X buffer, pH 8.5. Agarose crystals were dissolved by heating and allowed to cool followed by addition of five µl of 0.1% ethidium bromide (Sigma Aldrich, St. Louis, MO). The solution was then decanted into a pre-formed gel cast with comb. After casting, the gel was submerged into Tris 1X buffer, pH 8.5, in a horizontal gel electrophoresis chamber (Fisher Scientific, Hampton, MA). Ten µl samples of amplified DNA from PCR were loaded into each well with six µl of BenchTop PCR molecular weight markers (Promega Corp., Madison, Wisconsin) loaded into lane one. Gel electrophoresis was conducted at 90 V using a power supply unit (Bio-Rad laboratories, Hercules, CA). Following gel electrophoresis, gels were placed in a Fluor Chem 5500 ultraviolet light cabinet (Alpha Innotech, San Leandro, CA) to determine product sizes and reaction specificity.

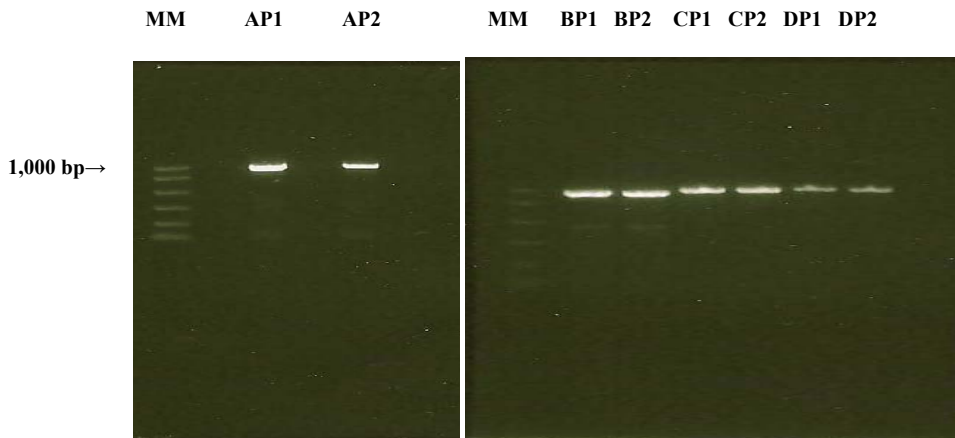


Figure 1. Verification of amplicon primer specificity via 1.4% gel electrophoresis. AP1 represents amplicon A pool one. AP2 represents amplicon A pool two. The same order can be inferred for amplicons B, C, and D. MM represents molecular weight markers. A product of approximately 1,000 base pairs represents a specific primer PCR product.

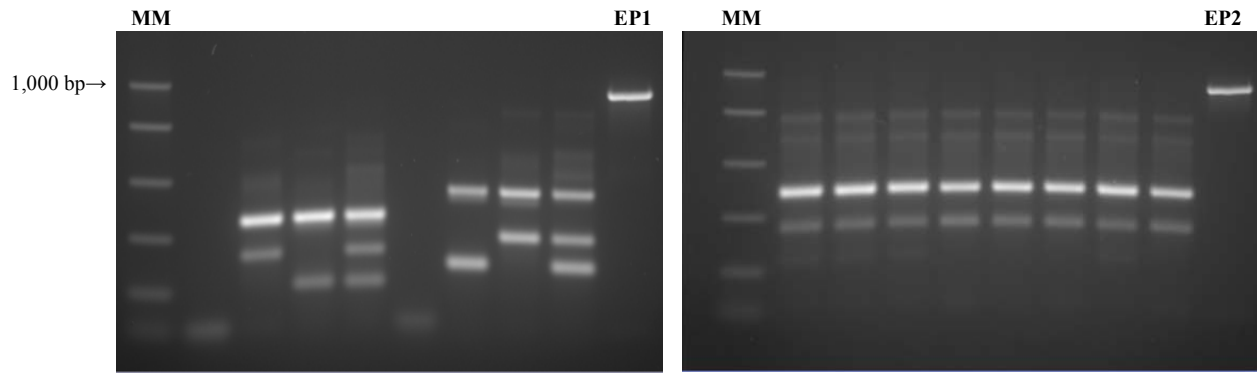


Figure 2. Verification of amplicon E primer specificity via 1.4% gel electrophoresis. EP1 represents amplicon E pool one. EP2 represents amplicon E pool two. A product of approximately 1,000 bp indicates a specific primer PCR product. Additional bands in the gel (lanes 2-9) represent DNA ran from another experiment.

DNA purification

Prior to DNA sequencing, primer specific DNA was purified and concentrated using DNA Clean & Concentrator-5 (Zymo Research Corp., Orange, CA). To a 1.5 ml microcentrifuge tube, 100 μ l of DNA Binding Buffer was added to the 50 μ l of DNA sample (2:1 ratio of DNA Binding Buffer to DNA sample) and vortexed briefly. This solution was then transferred to a spin filter placed in a microcentrifuge collection tube. The collection tube was centrifuged for 30 seconds at 10,000 rpm and flow-through removed. Two hundred μ l of DNA Wash Buffer was then added to the spin filter and centrifuged for 30 seconds at 10,000 rpm and step repeated. Following centrifugation, 10 μ l of DNase/RNase free water (Gibco Invitrogen, Carlsbad, CA) was added to spin filter. The spin filter was then transferred to a new 1.5 ml microcentrifuge tube and centrifuged for 30 seconds at 10,000 rpm. The eluted DNA was then assayed using the spectrophotometer prior to sequencing.

DNA sequencing and SNP detection

After DNA purification and concentration, sequences from amplicons A, B, C, D, and E of the IL-6 gene were taken to the University of Tennessee-Knoxville's Molecular Biology Resource Facility located at Walters Life Science building and sequenced with either ABI 3100 or 3730 genetic analyzers (Applied Biosystems, Foster City, CA). Sequences were received via email and entered into Vector NTI Advance 10 Contig Express (Invitrogen, Carlsbad, CA) software program and the chromatogram used to determine if SNPs were present in the IL-6 gene.

SNP genotyping

Individual cow DNA isolates were used as the DNA template instead of pools for PCR SNP genotyping. One PCR reaction tube contained 31 μ l of DNase/RNase free water, ten μ l of 5X green *Go Taq* reaction buffer, four μ l of 25 μ M MgCl₂, one μ l of dNTPs, two μ l of a mixed 2:1 volume ratio of inner primers to outer tetra ARMS primer (forward inner primer C allele, reverse inner primer T allele: forward outer ARMS primer, reverse outer ARMS primer), one μ l of isolated cow DNA template, and 0.25 μ l of *Go Taq* DNA Polymerase. Polymerase chain reaction of the SNP located on amplicon E +519 consisted of an initial 95 $^{\circ}$ C melt for two

minutes, followed by 40 cycles of 95 °C for 15 seconds, 68 °C for 20 seconds, and 72 °C for 40 seconds followed by a final step of 72 °C for ten minutes. Single nucleotide polymorphism genotyping was conducted via gel electrophoresis. A 2% agarose gel was prepared using the same reagents previously described. Ten µl PCR samples were loaded into each well and gel electrophoresis conducted at 70V.

Results

DNA sequence analysis

Genetic analysis of the sequenced IL-6 gene was conducted by comparing the IL-6 amplicons to the cow genome using BLAT from the UCSC browser (Kent 2002; <http://genome.ucsc.edu/>). A 4,344 bp IL-6 genomic region was revealed on bovine chromosome four using a whole shotgun sequence from bovine IL-6 (Genbank reference number NC_007302.3) as a query. Chromatogram analysis of amplicons A, B, C, and D revealed no SNPs in the pooled DNA sequences. A SNP was discovered within amplicon E with C>T transition at position +519 in the amplicon as shown in Figure 3. The position of this SNP relative to the transcription start site and exon/intron boundaries then was identified in order to better define the position of this SNP in the IL-6 gene and conform to standard naming guidelines (den Dunnen and Antonarakis 2000). Further analysis revealed this SNP, c.+472-152C>T, was located within intron four of the IL-6 genome as shown in Table 3. The SNP was located -152 bp from the start of exon five with exon five being +472 bp from the transcription start site.

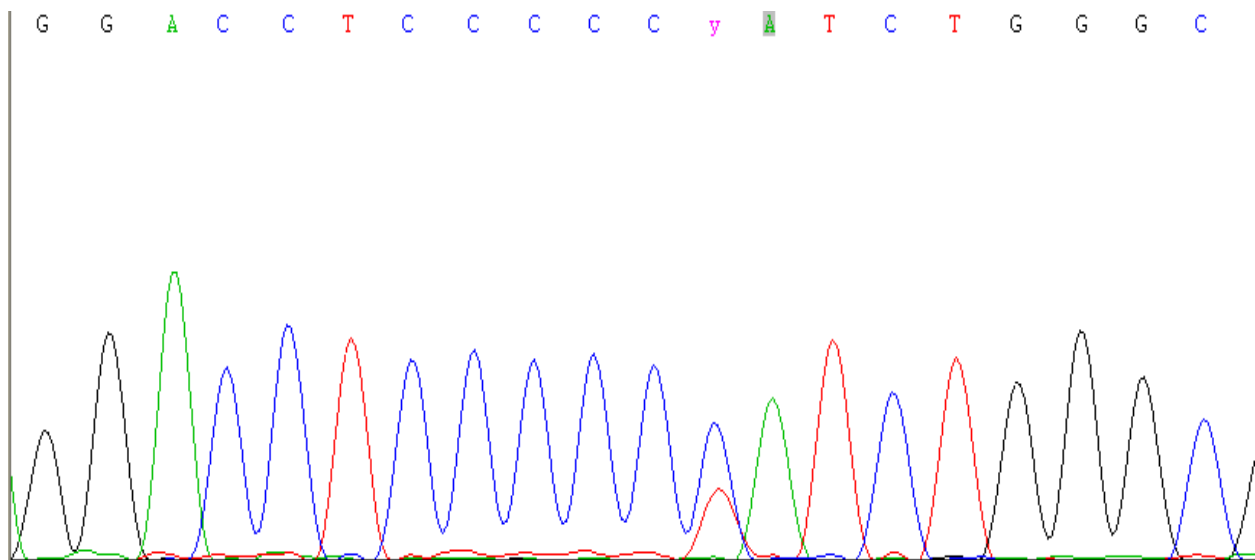


Figure 3. Chromatogram showing amplicon E SNP c. +472-152C>T

Table 3. Sequenced IL-6 genomic DNA with transcription start site highlighted blue, exon regions green, intron splice sites red and SNP c.+472-152C>T located at 3,606 bp of the genome red, and stop codon orange

GCCCTCCAGGAACAGCTATGAACTCCCGCTTCAAAAGTAAGTGAAGGAAATCCTTAGCCAGGAACGGCCTGGCAGCCGTGCGGC
 GAAGGAGGGGGTGTGTGTCGCTGTGAGGGCTGGCGGGCGGCCAGCAGCCAGAAGCACAGCTCCCCAGCGGCAGTCT
 GCTCACTGGCTGCTCCCTGTCTGTACAGGCGCCTTCACTCCATTGCTGCTCCCTGGGGCTGCTCTGGTATGACTTCT
 GCTTTCCTACCCTGGTCCCTGGGAGAAGATTTCAAAAAATGACCCACCCAGGCAGACTACTTCTGACCTCAGAGTAA
 CCGAAGCTCTCATTAAAGCGATGGTCGACAAAATCTCTGCAATGAGAAAGGAGGTGGGTAGGCTCGCTCCAGGGTATGAGCACC
 CATGTGGGCATCTATTTCCTTGTCTTGGATGTGAATGGGGCTTCTGCATTGGGAGGCTTTGTCTGGGTTGGAGGCAGTTCAGA
 CTTGAAGCCAAAGGGCAGACTATGCTCTGTAATTTCTCTCTGCTCTGGTCTCAGATGGTTCCTTTTCATCTGAAAAATACA
 GTGATGCTGGAACTCCCTGCCCCCTGGCCATTCTCAGTCCATTCTTGGCATGAGCTGTGGATCACACTGCCCTCAGGATTCTCT
 TTCCATAACCAAACTAAGGTCTGTCAATTTTTTGTGTGATTTGGACTTTAGGAGGATTAAGAACACCACCTGAAATGACTGT
 TCAGTCTCAGGACAGTACAGTACTTTTAGGTTATGCCAGCCAGAGGTTAGGGATTCCCCAAAGTATTCTGGCAGTAAAGAC
 GCATACCTGATAGTCCACCAAGTGGCAAAATCATCCATAGGCTACAGCAGCTCCAAGCTTTAAATGTAAGGTAGTTTGGCTCT
 GCTTTGGTAGAACTCCCTGCCCCCTGGCCATTCTCAGTCCATTCTTGGCATGAGCTGTAACTCTGATTCTGCTGCACTCAGGATTCT
 TGGGCAAAATTTTCAGCACCTCTGGGCTCCATCTGTAATAATCATGGAATTGGACTAGATGATCTGAAAGATCCTTCCAGCTGGGAC
 ATATAATTCTGGCTCAACTTATAGTCACTCAATATTATCAAAATTCCTATCATTATATGTTACTACTATGTCTGACACTTAT
 AAAATATTGTTTGTCTCACAATAAACTTAATGGGCAAGTTTTTCATTATCCCCAACTTGCAGATATAGGGCACTGAGACTGGG
 AATACATAACAACCTAAGTACGGGTGAGGCTAGAAGTGAACCATGTGCCATCTTCTGTAGAGCTACTGCTGGGTGCT
 AATGGACTGTGACTTCACTTTTCTTAGAGAATCTCTGGCCATGATAGATCATGCATCAGCCCTTAGTGGTGTGAGTTTTAGGG
 CATTATAAGATAACGGTCTGATATCTTCCCAGATATGTGAGAAGAATGATGAGTGTGAAAGCAGCAAGGAGACTGGCAG
 AAAATAAGCTGAATCTTCAAAAAATGGAGGAAAAGGACGGATGCTTCCAATCTGGGTCAATCAGGTACAAATGCATTACGTTCA
 CTTTTAGTACTCCCTAGTCAAAAAGTCTCCCTCTTGCATGAAGCATCTGTATATATAAAGACCAGGCAGGCAATGAAGAAGGGG
 ATTAATATAAAGAATACCATGTAGATTTTCATGAGGAGGCCAACTTAAATGATATTAAGGCAACTTATTTTTAAACAATCTCGTCA
 AATATAGCCCTCTGGGCTGGGATTTTTACCAGTGTCAAGTGTAGGAAACCCAGAGGTTAGCTAGTTCATCATGGGTAAGGTG
 GTCCAGGAACCTTTCTCTCTTGGCTGCCCTTAGCAGGATCTAGGCCTGCCCTCCCTGCCAGATCATTCTACACCCTTCTCCT
 TCAGACAGTGCAGTAACCTCCACTGACTTCTGCTTATTTCAAAAAGGAGGTTTACCCTGACATTGACAAGGGCAGGAGTGCAG
 CCAGCGGAAGAGCATTATGGGTTGTAACAGAGGATGGCCCTTGCATTAAGTATGATATTAAGGCAACTTATTTTTAAACAATCTCGTCA
 TGGTGCCTCTCGTTGTGCAGGATTTTTAAATGGGGCAGCTAATGCTGTCCAATGTCTCAAAACATGATGCTTAAGAAGTACTTGAA
 GTTCCTGGAGGGGAACAGGCAGAAAGTGTGAATTATAGCAATTTACATTTCTAAAATCAATTATGCCTGTCTGGCAATAACCAG
 TTTTCCACTGTCTTTCTCTTAGCGGATTTGCTGATCAGAACCTGCTGGTCTTCTGGAGTATCAGATATACCTGGACTACCTC
 CAGAAGGATGAGGGAAATCAGGAAAATGTCAGGATTTGAGGAAAAATCAGAACACTGATCCAGATCCCTGAAGCAAAAG
 GTGAGTGTCCCCTCGCTGACTCAAAGAAGGCGCCTAGACAACCTCGGCAGTGCAAAAGCCAATTTCAAAAAGAGATGGCTGTGTGTA
 AAGGAAAAGATCTGACAAATATTTCTCAAGTTAAAGCTTTTCTTTTCTGTCTTCTGCAAATGACATCAATAACTATATTTAAA
 ACTCAGTTAAGGGGAGAGTTTTAACATTAACCTTCAAAAAGTTTAAAGAAATAAGGGGAAAAAGTATCTGTAGAGGCAGAAAGGAA
 GCATATCCAACCTCAAGTAACAATCCATAATAACAAGAACAGAGAGCACAAAAACTATTTAAGATTGCCAAATGACATTTCTGT
 TATTATGGTAGTTTCAACTACTTTCCCATACTTTTTGCCAAGGAGTTAATGACAGAGAAAAGGCTAAAAGGCTGGAAGAGCATGTAC
 CAAGGGCTTTTTGGTCACTACCAGTACCCATTCAAATGCCACATGGTATAAGCTTGTATGGCACAGTTTGCAGATGAACCATAGC
 ACAAAGAGAAGGAGAAGCAGGCACATTGGTGCTAATCCTGCCTCCCTTTCCCTAGCTGAGTCCCTTGTACCAGGGCTTCTCCCCGTT
 GAAGGTGCAGACCTTACCTAGAGGATCTTGTACAGTGCAGAGTCTGACTCAACAGGTCTGGGGTGGAGCCTGAGACTCACA
 TTTCTAACAAAGCGCCTGGCAATGCTGAGGTTGCAGTGTCCAGGATCACAGGTGGAGGTGCGAGGCCTTCCACCATGACTTAGCT
 CCAAGTCAAGTGGGTCTATGGAAAAATCTTCCCCAATTTCTGTTTACGCCGCCAGAGCAGCTGACTCTTAATAGATGTCAGATGGATG
 GGTGAGAGCAAAGCCAAAGAAAAGTATTAGAGTTAGAGGCCCCCAAGAACCATTAAAATGCCTTTAACTAAAAGTCTTTCATA
 CATCTACTGAATTTAACTTTGTGCCAGGCACTGTTAGGATGACATAAGACAAGTAAAGATACGATGGCTACTTATGAGATGCTTA
 TGATCCAACAGAGACAGAGCAGCAAAAAGTATCAGAGCTCAGTTGCTGTGAGAGGGTTCATGCTGGGTGGGCTCAGAGCAGAG
 GACCTCCCCC[C/T]ATCTGGGCGAGTCAAGGCCAGTGGCCACATCTGCTCCTTCGCTCCAAATTTGTTTGAATATCTTCCACAGG
 CCACAGACCTCTCCACTGCAAAAGACTTCTTCAATATTTAAACAATTCCTTTGTTTTGTGTTCCCTTCATGAAAGATCGCAGAT
 CTAATAACCAGCTCCAGCCAAAACACTGACCTGCTGGAGAAGATGCAGTCTTCAAACGAGTGGGTAAGAAAGCAGCAAGATTATC
 CTCATCCTGAGAAACCTTGGAAATTTCTGCAGTTCAGCTGAGAGCTATTCGGATGAAGTAGCTGGGGCTCCCATGATTTGGTA
 GTTCTGGGCATCCCTCCTCTGGTCAAAAACCTGTCCACTGGGCACACAACCTTATGTTGTTCTCTATGAAGAACTAAAAGTATGA
 GCGTTAGGACACTATTTATCTTAAATTTATGATATTTAAATATGTGATTTTGGTAAATTTATATACATGATAAGTATTTATATTT
 TTAGAAGTGCCACTTGAATATTTTATGATTTGGTTTGAAGAAAGTAAACGTAATAATGGCTATGTGGCTTGAATGTCTTATTTGTT
 TTGGAGCAAAATCATTCTTGAATGTGTAGGCTTACCTCAAAAAATTTGCTAACTTATGCATATTTTTAAAGGCATATTTATATTG
 TATTATATAAATGTTTAGGCTGTTTTATAACAATAAACTCTTTTTTAAAGAAAAAAA

Table 5. Alignment of interspecies genomic regions with c. +472-152C>T SNP position in red

```

Cow -----gtgggctcagagcagagg-acctcc-----ccccatct
Dog -----gtgggctctgaggagaggtacctca-----ccctacct
Human taagaccagcctgggtcaacatggtaaaaccccatctctacttaaaaatacaaaaaagttagccagggcat
Mouse -----gtaagcttgggaacaaaagcatctcc-----cctggctt
Platypus =====

```

Genotype and allele frequencies

Cows were genotyped at c.+472-152C>T IL-6 SNP locus to determine genotype and allele frequency using generated Tetra-ARMS primers from Table2. The C allele had an allele frequency of 67.5% (27/40) and the T allele 32.5% (13/40). The CC genotype was observed at a genotype frequency of 35% (7/20), CT genotype 65% (13/20), and TT genotype not detected within the sample population as shown in Table 6. Shown in Figures 5 and 6 are allelic and genotypic frequencies of cows sequenced from the IL-6 c.+472-152C>T SNP of amplicon E with genotype determination by gel electrophoresis shown in Figure 7.

Table 6. Genotypes at the c. +472-152C>T SNP and pedigrees of sample population cows from ETREC

KES Cow Number	Sire		GS NAME	GGS Code	GGS NAME	MGSNmeCde	Genotype
	NmeCde	GS Code					
3781	7HO4937	009HO01289	OSCAR	009HO00719	NED BOY	7HO3948	CC
3788	11HO3276	001HO01464	CLEITUS	023HO00206	TRADITION	7HO3714	CC
3839	7HO4295	007HO01897	BLACKSTAR	009HO00584	CHAIRMAN	29HO6425	CC
3861	7HO5112	007HO03532	CHESAPEAKE	007HO00980	MARK	7HO4433	CT
3862	7HO4638	001HO00414	TESK	029HO02851	VALIANT	7HO4351	CT
3936	29HO8375	011HO03073	ALTALUKE	001HO01464	CLEITUS	7HO3994	CT
3944	1HO5021	009HO01401	MOUNTAIN	011HO01479	ROYALTY	7HO4226	CT
3992	1HO6149	007HO04161	RICK	100HO01903	LABAN	29HO8343	CC
4019	7HO5851	007HO03938	AIRLINER	029HO05730	MELWOOD	14HO2090	CT
4098	11HO5183	073HO01965	RUDOLPH	039HO00246	AEROSTAR	1HO6014	CT
4157	7HO6960	029HO08343	CONVINCER	023HO00206	TRADITION	1HO975	CT
4159	11HO6433	007HO05157	DURHAM	007HO02236	ELTON	1HO2993	CT
4160	11HO6433	007HO05157	DURHAM	007HO02236	ELTON	1HO2993	CT
4161	200HO44	011HO03562	ALTAFORMATION	008HO02024	LEADMAN	14HO2090	CT
4162	7HO5581	007HO03948	EMORY	007HO01897	BLACKSTAR	7HO5442	CT
4163	200HO44	011HO03562	ALTAFORMATION	008HO02024	LEADMAN	7HO5442	CT
4165	7HO6546	007HO04637	WINCHESTER	039HO00246	AEROSTAR	9HO2575	CC
4166	7HO6546	007HO04637	WINCHESTER	039HO00246	AEROSTAR	7HO5581	CC
4164	14HO3831	097HO00060	ADDISON	009HO01401	MOUNTAIN	14HO2090	CT
4167	7HO6546	007HO06546	REMINGTON	007HO04637	WINCHESTER	7HO8430	CC

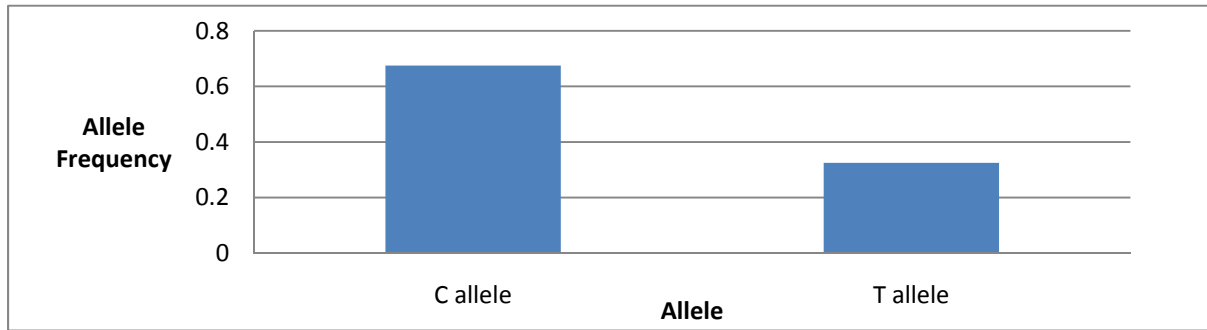


Figure 5. Allelic frequencies of cows sequenced from IL-6 c.+472-152 C>T SNP located on amplicon E intron 4 using Tetra-ARMS primers

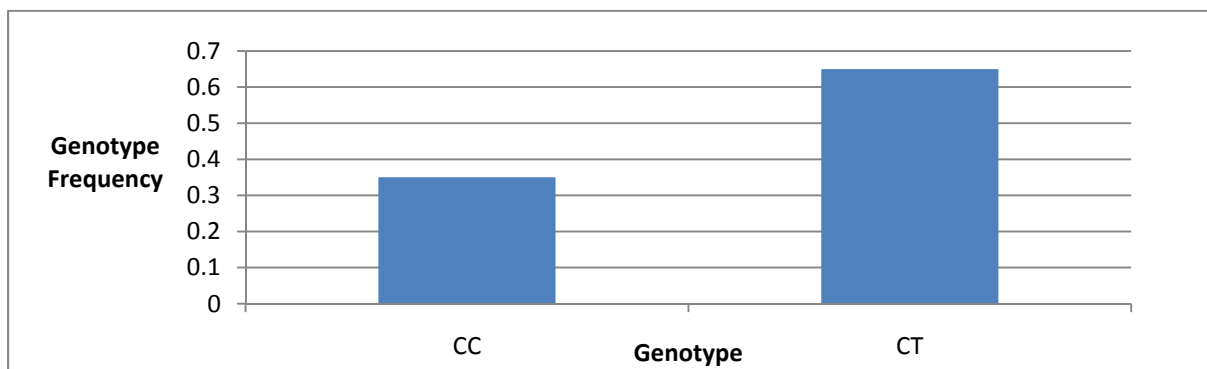


Figure 6. Genotype frequencies of cows sequenced from IL-6 c.+472-152 C>T SNP located on amplicon E intron 4 using Tetra-ARMS primers

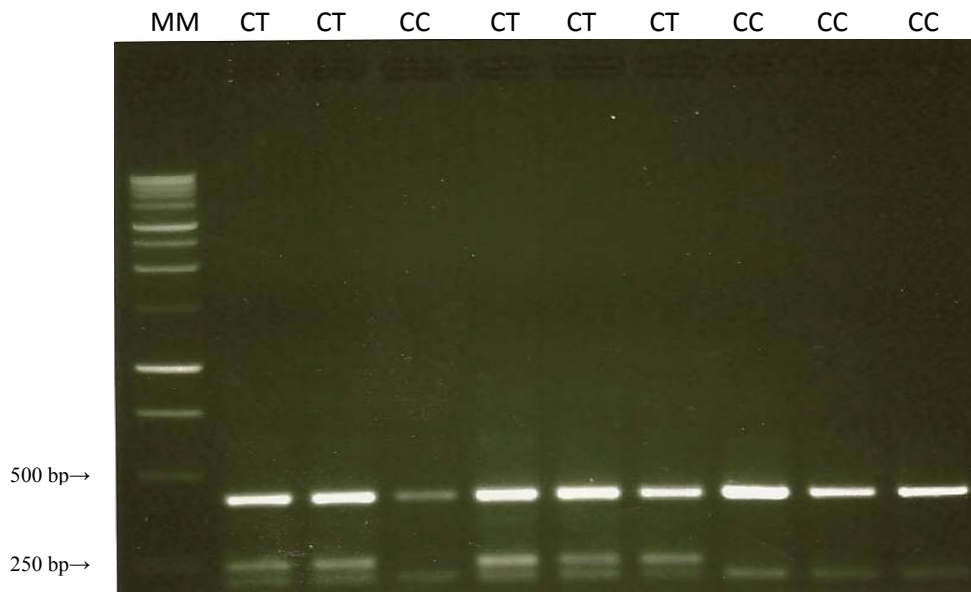


Figure 7. 2% gel electrophoresis IL-6 SNP genotyping of cows using specific Tetra-ARMS primers. MM represents molecular weight markers where CC and CT represents the cow's genotype at the IL-6c.+472-152C>T SNP locus.

Discussion

Further analysis of the IL-6 gene determined that the c. +472-152C>T SNP of amplicon E was located within intron four. Introns are genomic regions that are spliced out in pre-mRNA before incorporated into mature mRNA and translated into protein. Thus, introns are non-coding regions of the genome and by convention do not have an effect on the expression of a gene. However, recently discovered small RNA molecules about 20 nt in length called micro RNAs (miRNA) encoded by introns have been found to regulate gene expression by binding to complimentary mature mRNA sequences and suppressing protein translation (Ruby et al. 2007). Specifically, diseases have been associated with intronic polymorphisms or nucleotide repeats (Stangl et al. 2000, Sanghera et al. 2004, Kumar and Ghosh 2008).

Genetic diversity was maximized by selecting cows with unrelated pedigrees to the maternal and paternal grandsire generations representing 17 and 16 different families, respectively. Thirty percent of selected cows shared a common paternal grandsire with 35% sharing a common maternal grandsire. The C allele was found to be dominant in the sample population with a frequency of 67.5% and the T allele 32.5%. This may indicate that the SNP may have occurred recently within the sample Holstein population and the T allele is relatively new to the population or that the C allele has become more favorable. The genotype frequency for CC was found to be 35%, CT genotype 65%, and TT genotype not detected. The absent or relatively infrequent TT genotype in the sample cow population may indicate that this genotype is genetically unfavorable. According to Mendelian genetics, if two heterozygous individuals were crossed 25% CC, 50% CT, and 25% TT genotypes would be expected. If homozygote and heterozygote individuals were crossed 50% CC and 50% CT genotypes would be expected. The genotype frequencies for the sample cow population fail to follow Mendelian genetics. As a preliminary study with 20 cows, further experiments with a larger sample cow population may be warranted to study IL-6 gene expression in the CC, CT, and TT cow genotypes and determine if a certain genotype is more susceptible to mastitis or other diseases. Additionally, a larger sample size may determine an explanation for cow genotype frequencies failing to follow Mendelian genetic inheritance patterns.

Conclusion

The purpose of this study was to sequence the IL-6 gene and determine if SNPs were present that may serve as potential genetic markers. One SNP with a C to T transition was found within intron 4 of the IL-6 gene. Sequencing of Holstein dairy cattle at this locus depicted genotype frequencies not consistent with Mendelian genetics. By convention, polymorphisms within intron regions are not associated with disease; however, studies have linked intronic SNPs with disease. Further experiments with a larger sample cow population may be warranted to study IL-6 gene expression in the CC, CT, and TT cow genotypes and determine if a certain genotype is more susceptible to mastitis or other diseases. Additionally, a larger sample size may determine an explanation for cow genotype frequencies failing to follow Mendelian genetic inheritance patterns.

References

- den Dunnen, J.T. and S.E. Antonarakis 2000. *Hum.Mutat.* 15: 7-12.
- Dienz, O. and M. Rincon. 2008. The effects of IL-6 on CD4 T cell responses. *Clinical Immunology* 130: 27-33.
- Diez-Fraile, A., E. Meyer, C. Burvenich. 2003. Sympathoadrenal and immune system activation during the periparturient period and their association with bovine coliform mastitis. *The Veterinary Quarterly* 25: 31-44.
- Griffiths, A. J., R. C. Lewontin, S.B. Carroll, Sue Wessler, and Susan R. Wessler. *An Introduction to Genetic Analysis*. New York City: W.H. Freeman and Company, 2007.
- Kent WJ. BLAT - the BLAST-like alignment tool. *Genome Res.* 2002 Apr;12(4):656-64. <http://genome.ucsc.edu/>
- Kumar, A., and B. Ghosh. *Genes and Immunity*. 2008 A single nucleotide polymorphism (A -> G) in intron 3 of IFN gamma gene is associated with asthma. *Genes and Immunity* 9:294-301
- LeBlanc, S. J., K. D. Lissemore, D. F. Kelton, T. F. Duffield, and K. E. Leslie. 2006. Major Advances in Disease Prevention in Dairy Cattle. *J. Dairy Sci.* 89:1267–1279.
- Lievaart, J.J., H. W. Barkema, W. D. J. Kremer, J. van den Broek, J. H. M. Verheijden, and J. A. P. Heesterbeek. 2007. Effect of herd characteristics, management practices, and season on different categories of the herd somatic cell count. *J. Dairy Sci.* 90:4137–4144.
- Ruby, J.G., C.H. Jan, and D.P. Bartel. 2007. Intronic microRNA precursors that bypass Drosha processing. *Nature* 448: 83-86.
- Shim, E.H., R.D. Shanks, and D.E. Morlin. 2004. Milk loss and treatment costs associated with two treatment protocols for clinical mastitis in dairy cows. *J. Dairy Sci.* 87:2702–2708.
- Sanghera D.K., S. Manzi, F. Bontempo, C. Nestlerode, and M.Y. Kamboh. 2004. Role of an intronic polymorphism in the PDCD1 gene with the risk of sporadic systemic lupus erythematosus and the occurrence of antiphospholipid antibodies. *Human Gen.* 115: 393–398.
- Stangl K., I. Cascorbi, M. Laule, T. Klein, V. Stangl, S. Rost, K.D. Wernecke, S.B. Felix, A. Bindereif, G. Baumann, and I. Roots. 2000. High CA repeat numbers in intron 13 of the endothelial nitric oxide synthase gene and increased risk of coronary artery disease. *Pharmogenetics* 10: 133-140.
- 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.

Zhao, X., and P. Lacasse. 2008. Mammary tissue damage during bovine mastitis: Causes and control. *J. Anim Sci.* 2008. 86:57-65.