(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 May 2007 (10.05.2007)

PCT

(10) International Publication Number
WO 2007/053595 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/US2006/042425

(22) International Filing Date: 31 October 2006 (31.10.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/731,748 31 October 2005 (31.10.2005) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US Not furnished (CIP)
Filed on Not furnished

(71) Applicant *(for all designated States except US)*: **UNIVERSITY OF TENNESSEE RESEARCH FOUNDATION** [US/US]; 1534 WHITE AVENUE, SUITE 403, Knoxville, TN 37996-1527 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **PIGHETTI, Gina, M.** [US/US]; 2640 MORGAN CIRCLE, Knoxville, TN 37996 (US).

(74) Agents: **SCHICKLI, Warren, D.** et al.; KING & SCHICKLI, PLLC, 247 NORTH BROADWAY, Lexington, KY 40507 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US (patent), UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: METHODS FOR IDENTIFICATION OF SUSCEPTIBILITY TO INFLAMMATORY DISEASE

612
CTTCCGTGAGGCCTATCAACCACC**G/A**TACTCCGACCTAGTCTGCTACGAGGACCT

684
GGGTGCCAATACAACGAAATGGCGGATGATAATGCGTGTCCT**G/A**CCCCAGACCTT

TGGCTTCCTCCTGCCCCTGCTGGTCATGCTGTTCTGCTACGGATTCACCCTGCGC

777
ACGCTATTTTCAGCCCAAATGGGGCA**C/G**AAGCACCGGGCCATGCGGGTCATCTTT

858
GCTGTCGTGCTCGTCTTCCTGCTCTGCTGGCTGCCCTACAACCTGGTCCTGAT**C/A**

861
GC**G/A**GACACCCTCATGAGGGCCCATGTGATTGCTGAGACCT

(57) Abstract: Methods for determining predisposition to an inflammatory disease are provided, including obtaining a sample of nucleic acid from an animal predisposed to contracting an inflammatory disease such as mastitis, and determining the identity of the nucleotide in SEQ ID NO: 35 in at least one of positions 151970, 152084, 152164, or 152328, and in SEQ ID NO: 36 in position 546. The methods further include detecting the presence of a single nucleotide polymorphism in SEQ ID NO: 35 in at least one of positions 151970, 152084, 152164, or 152328. The methods further include detecting the presence of a single nucleotide polymorphism in SEQ ID NO: 36 in position 546. Still further, the invention provides probes for detecting single nucleotide polymorphisms in at least one of positions 151970, 152084, 152164, or 152328 of SEQ ID NO: 35, and for detecting a single nucleotide polymorphism in position 546 of SEQ ID NO: 36.

WO 2007/053595 A2

Published:

— *without international search report and to be republished upon receipt of that report*

— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

METHODS FOR IDENTIFICATION OF SUSCEPTIBILITY TO INFLAMMATORY DISEASE

This application is a continuation-in-part of U.S. Patent Application No. 11/198,504 filed on August 5, 2004, and also claims the benefit of priority in U.S. Provisional Patent Application Serial No. 60/731,748 filed on October 31, 2005, the disclosures of each of which are incorporated herein in their entirety by reference.

Technical Field

The present invention relates to single nucleotide polymorphisms in the bovine Interleukin-8 receptor gene (CXCR1) and the natural resistance-associated macrophage protein (NRAMP-1). In particular, the invention relates to use of the polymorphisms in CXCR1 and NRAMP-1 for evaluating susceptibility or predisposition of an animal to inflammatory diseases, including mastitis.

Reference to Sequence Listing Submitted on Compact Disc

This application incorporates by reference in its entirety the Sequence Listing that is provided in duplicate on compact discs that accompany the application. Each CD contains the following file: 1101009SEQLIST.txt, having a date of creation of Oct. 30, 2006 and a file size of 213 KB.

Background of the Invention

Mastitis, an inflammation of the mammary gland, not only impairs milk production and quality, but also is associated with other diseases that limit the productive life of the animal. A substantial amount of research has demonstrated that mastitis negatively impairs both milk quantity and quality. The inflammatory response activated by infection increases the level of white blood cells in the mammary gland, which contributes to destruction of milk-producing tissue, as well as alters the concentration of fat, protein, sugars, and minerals which subsequently interferes with the making of cheese and other milk products, shelf life, and palatability. However, the effects of mastitis are not limited to the mammary gland. Research conducted at the University of Tennessee and other institutions also has implicated mastitis with impaired reproductive ability. In particular, mastitis that occurred early in lactation increased the days to first service, days open, and services per conception. In a recent paper modeling the effects of disease on productivity, cattle with chronic mastitis exhibited a 1.5 fold greater chance of having cystic ovaries and a 3 fold greater chance of having metritis. A limited number of studies also have indicated a potential link with mastitis and diseases such as displaced abomasums, ketosis, and lameness.

The chemokine interleukin-8 (IL-8) plays a critical role in regulating the inflammatory response by inducing neutrophil migration, delaying neutrophil apoptosis (cell death) until the infection site is reached, and enhancing neutrophil killing ability (Glynn, P.C., Henney, E., Hall, I.P., 2002. The selective CXCR2 antagonist SB272844 blocks interleukin-8 and growth-related oncogene-α mediated inhibition of spontaneous neutrophil apoptosis. Pulmonary Pharmacology and Therapeutics 15, 103-110). Both human and bovine neutrophils recognize IL-8 by CXCR1 and/or CXCR2, which are members of the seven-transmembrane G-protein coupled receptor family (Sprenger, H., Lloyd, A.R., Lautens, L.L., Bonner, T.I., Kelvin, D.J., 1994. Structure, genomic organization, and expression of the human interleukin-8 receptor B gene. Journal of Biological Chemistry 269, 11065-11072). Human CXCR1 and CXCR2

receptors express 77% sequence homology and differ in their binding affinities for ELR^+CXC chemokines (Sprenger et al. 1994). CXCR1 binds only to IL-8 with high affinity, while CXCR2 binds to IL-8 and other ELR^+CXC chemokines such as growth related oncogene-α and neutrophil activating peptide-2 Murphy, P., 1994. The molecular biology of leukocyte chemoattractant receptors. Annual Review of Immunology 12, 593-633). Bovine IL-8 receptor loci (CXCR1 and CXCR2) have been mapped approximately 90.3 cM from the centromere of bovine chromosome (BTA) 2. These loci are approximately 1.3 cM from the natural resistance associated macrophage protein (NRAMP)-1, a polymorphic gene related to immune function in humans and mice, indicating that this region of BTA 2 may be associated with immune function and disease resistance (Grosse, W.M., S.M. Kappes, W.W. Laegreid, J.W. Keele, C.G. Chitko-McKown, M.P. Heaton, 1999. Single nucleotide polymorphism (SNP) discovery and linkage mapping of bovine cytokine genes. Mammalian Genome 10, 1062-1069).

Four single nucleotide polymorphisms (SNPs) with Mendelian inheritance patterns were identified within a 523 bp segment of the bovine IL-8 receptor locus in a commercial beef cattle population (Grosse et al. 1999). This fragment's position is predicted to reside in the coding region for the bovine CXCR2 gene. Bison and several different breeds of beef cattle exhibited varying frequencies of these four IL-8 receptor SNPs and the resulting five haplotypes (Heaton, M.P., Grosse, W.M., Kappes, S.M., Keele, J.W., Chitko-McKown, C.G., Cundiff, L.V., Braun, A., Little, D.P., Laegreid, W.W., 2001. Estimation of DNA sequence diversity in bovine cytokine genes. Mammalian Genome 12, 32-37). Of these SNPs, one results in an amino acid change within a region of the third intracellular loop and has the potential to influence signaling events that lead to neutrophil migration and reactive oxygen species generation. However, polymorphisms for the IL-8 receptor have not been identified previously in dairy cattle.

Neutrophils in the mammary gland constitute a primary defense mechanism against intramammary infection. As such, genetic markers within the bovine CXCR1 gene and the NRAMP1 gene have the potential to identify cattle susceptible to inflammatory diseases such as mastitis, as well as those more susceptible to related conditions such as reproductive diseases and performance, metabolic diseases, and lameness. As mastitis has such a strong impact on milk production, it also is highly probable that markers within these genes will be associated with the yield of milk and its components.

There is accordingly identified a need in the art for novel methods of predicting susceptibility of humans and animals, such as for example dairy cattle, to inflammatory diseases such as mastitis. To meet this identified need in the art, candidate SNPs and haplotypes resulting therefrom have been identified within a segment of the bovine IL-8 receptor locus, and frequencies of both SNPs and haplotypes within two different breeds of dairy cattle have been identified. The identified SNPs and resulting allelic haplotypes allowed evaluation of inflammatory disease phenotypes, including mastitis, with immune responses. The ability to select disease resistant cattle based on genetic markers associated with mastitis will not only enable the development of a more resistant population but also will permit closer examination of the mechanisms that contribute to differential disease resistance, leading to novel therapeutic strategies against mastitis and other inflammatory diseases.

Summary of the Invention

In one aspect of the present invention, a method is provided for determining predisposition to an inflammatory disease, comprising obtaining a sample of nucleic acid, which may be genomic DNA, from an animal predisposed to contracting such inflammatory disease, the sample comprising at least one of

SEQ ID NO: 35 and SEQ ID NO: 36. The present method further contemplates determining the identity of the nucleotide in at least one of positions 151970,

152084, 152164, or 152328 of SEQ ID NO: 35, and determining the identity of the nucleotide in position 546 of SEQ ID NO: 36. Typically, the method includes detecting the presence of a single nucleotide polymorphism in SEQ ID NO: 35 in at least one of positions 151970, 152084, 152164, or 152328. The polymorphism detected at position 151970 is G → A, the polymorphism detected at position 152084 is C → G, the polymorphism detected at position 152164 is A → C, and the polymorphism detected at position 152328 is G → A. The method may also include the step of detecting the presence of a single nucleotide polymorphism in SEQ ID NO: 36 in position 546. The polymorphism detected at position 546 is C → G.

The inflammatory disease may be an inflammation of the mammary gland. For example the inflammatory disease may be mastitis, and the animal predisposed to such inflammatory disease may be a ruminant animal such as a dairy cow.

In another aspect of the present invention, a method for determining predisposition to an inflammatory disease is provided. In one embodiment, the method includes obtaining a sample of nucleic acid (which may be genomic DNA) from an animal predisposed to contracting the inflammatory disease, and determining the identity of the nucleotide in at least one of positions 151970, 152084, 152164, or 152328 of SEQ ID NO: 35. Further, the method may include the step of detecting the presence of a single nucleotide polymorphism in SEQ ID NO: 35 in at least one of positions 151970, 152084, 152164, or 152328. The polymorphism detected at position 151970 is G → A, the polymorphism detected at position 152084 is C → G, the polymorphism detected at position 152164 is A → C, and the polymorphism detected at position 152328 is G → A.

In yet another aspect, the present invention provides a method for

determining predisposition to an inflammatory disease which includes obtaining a sample of nucleic acid from an animal predisposed to contracting the

inflammatory disease, and determining the identity of the nucleotide in position 546 of SEQ ID NO: 36. The method may include the further step of detecting the presence of a single nucleotide polymorphism in SEQ ID NO: 36 in position 546. The polymorphism detected at position 546 is C → G.

In still yet another aspect of the present invention, there is provided an isolated nucleic acid molecule comprising 400 contiguous nucleotides of SEQ ID NO: 35 or the complement thereof, including: (a) nucleotide 151970 wherein the G is replaced by an A; (b) nucleotide 152084 wherein the C is replaced by a G; (c) nucleotide 152164 wherein the A is replaced by a C, (d) nucleotide 152328 wherein the G is replaced by an A; or at least one of (a), (b), (c), and (d). Still yet further, there is provided an isolated nucleic acid molecule comprising 30 contiguous nucleotides of SEQ ID NO: 36 or the complement thereof, including nucleotide 546 wherein the C is replaced by a G.

The present invention provides also a probe for use in evaluating predisposition or susceptibility of an animal to an inflammatory disease, wherein the probe is selected from the group consisting of: (a) a probe that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO: 35 having A at position 151970 but not to a nucleic acid molecule consisting of SEQ ID NO: 35 having G at position 151970; (b) a probe that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO: 35 having G at position 152084 but not to a nucleic acid molecule consisting of SEQ ID NO: 35 having C at position 152084; (c)a probe that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO: 35 having C at position 152164 but not to a nucleic acid molecule consisting of SEQ ID NO: 35 having A at position 152164; (d) a probe that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO: 35 having A at position 152328 but not to a nucleic acid molecule consisting of SEQ ID NO: 35 having G at position 152328; and (e) a probe that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO: 36

having G at position 546 but not to a nucleic acid molecule consisting of SEQ ID NO: 36 having C at position 546. the probe may be detectably labeled, for use in an assay for determining predisposition or susceptibility of an animal to the inflammatory disease.

As should be appreciated, the embodiment shown and described in this patent application is an illustration of one of the modes best suited to carry out the invention. It will be realized that the invention is capable of other different embodiments and its several details are capable of modification in various, obvious aspects all without departing from the invention. Accordingly, the drawings and descriptions will be regarded as illustrative in nature and not as restrictive.

Brief Description of the Drawings

The accompanying drawings incorporated in and forming a part of the specification, illustrate several aspects of the present invention and together with the description serve to explain the principles of the invention. In the drawings:

Figure 1 shows SNPs identified in Holstein and Jersey cattle in the partial bovine Interleukin-8 receptor sequence (Genbank Accession No. U19947), with nucleotide location noted above each SNP and with nucleotide substitutions indicated in bold;

Figure 2 shows a phylogenetic tree of ten CXCR1 haplotypes present in dairy cattle, based on the partial coding sequence of the bovine Interleukin-8 receptor gene shown in Figure 1, and with individual haplotypes numbered (HAP1, HAP2, etc.) and showing specific polymorphisms at +612, +684, +777, +858, and +861;

Figure 3 graphically shows CXCR1 single nucleotide polymorphism +777 genotype by estimated observations positive for subclinical mastitis (%), with data

shown as mean ± SEM, and with letters in caps representing LSM separation ($P < 0.05$) and n = number of cows expressing CXCR1 SNP +777 genotype; and

Figure 4 graphically shows CXCR1 SNP +777 genotype by estimated observations positive for clinical mastitis (%), with data shown as mean ± SEM and n = number of cows expressing SNP +777 genotype.

Reference will now be made in detail to the present preferred embodiment of the invention, an example of which is illustrated in the accompanying drawings.

Detailed Description of the Invention

In accordance with the foregoing, methods are provided for predicting the susceptibility or predisposition of animals, for example ruminant animals such as dairy cattle, to inflammatory diseases such as mastitis. To meet this identified need in the art, candidate SNPs and haplotypes resulting therefrom have been identified within a segment of the bovine IL-8 receptor locus (CXCR1), and frequencies of both SNPs and haplotypes within two different breeds of dairy cattle have been identified. The identified SNPs and resulting allelic haplotypes allowed evaluation of inflammatory disease phenotypes, including mastitis, with immune responses. The ability to select disease resistant cattle based on genetic markers associated with mastitis will not only enable the development of a more resistant population but also will permit closer examination of the mechanisms that contribute to differential disease resistance, leading to novel therapeutic strategies against mastitis and other inflammatory diseases.

It is noted that in previous work (see Youngerman et al., 2004, Association of CXCR2 polymorphisms with subclinical and clinical mastitis in dairy cattle, J. Dairy Sci. 87: 2442-2448; see also Youngerman et al., 2004, Novel single nucleotide polymorphisms and haplotypes within the bovine CXCR2 gene, Immunogenetics 56: 355-359), the gene referenced herein as CXCR1 was identified as CXCR2. The present inventors have recently discovered (Pighetti et al., 2006, Genome Conservation Between the Bovine and Human Interleukin-8

Receptor Complex: Improper Annotation of Bovine Interleukin-8 Receptor B Identified, Vet. Immunol. Immunopathol. 114: 335-340) that the published Genbank sequence was erroneously identified as CXCR2, when in fact it corresponded to CXCR1. Specifically, bovine orthologs of human CXCR1 and CXCR2 were identified. Alignment of bovine CXCR2 reference mRNA to the bovine genome revealed two regions of similarity on BTA2 approximately 20 kb apart and on opposite strands. Comparison with the human genome suggested the more centromeric region to be CXCR2 and the more telomeric region to be CXCR1, which contradicts the current annotation of the bovine CXCR2 reference mRNA (Genbank Accession No. NM_174360.2). Accordingly, in the present disclosure, reference is made to the bovine CXCR1 gene, where previously the relevant gene was believed to be CXCR2 based on the published reference sequences.

Example 1

Jersey cows (n=42) from the Lewisburg Dairy Experiment Station, Lewisburg, TN, and Holstein cows (n=37) from the Middle Tennessee Experiment Station, Spring Hill, TN, that had completed at least two full lactations were selected at random. Blood was collected by puncture of the jugular vein in 10 ml Vacutainer tubes (Becton-Dickinson, Franklin Lakes, NJ) containing ethylenediamine tetracetic acid as anticoagulant. Blood was shipped overnight and upon arrival, 1.0 ml was aliquotted into 1.5 ml microcentrifuge tubes and stored at –80°C. Genomic DNA (gDNA) was isolated from whole blood using the UltraClean DNA Isolation Kit (MoBio Labs, Inc., Solana Beach, CA).

Forward and reverse primers were designed from the bovine IL-8 Receptor A (CXCR1) (GenBank Accession No. U19947SEQ ID NO: 1) sequence to amplify a 311 bp region that contained 4 SNPs as determined in a beef cattle population (Grosse et al. 1999). Each reaction was carried out in a final volume of 50 µl containing: 0.1 µM dNTP, 1.0 µM each of forward and reverse

oligonucleotide primers IL8Rec-SSCPFor (5'-CTTCCGTGAGGCCTATCAAC-3') (SEQ ID NO: 2) and IL8Rec-SSCPRev (5'-AGGTCTCAGCAATCACATGG-3') (SEQ ID NO: 3), 5 µl 10X magnesium-free thermophilic buffer [500 mM KCl, 100 mM Tris-HCl (pH 9.0), 1% Triton X-100], 3 µl of 25 mM $MgCl_2$, 20.9 µl nuclease free water, 3 U *Taq* DNA Polymerase (Promega, Madison, WI), and 200 ng gDNA suspended in nuclease free water (Sigma, St. Louis, MO). Thermocycler (Eppendorf, Westbury, NY) parameters consisted of an initial denaturing step at 94°C for 5 min, 30 cycles of a denaturing step at 94°C for 2 min, an annealing step at 60°C for 30 sec, and an extension period of 1 min at 72°C. Following cycling, a 20 min extension step at 72°C was incorporated before a final hold at 4°C. Following each reaction, PCR products were checked by electrophoresis on a 1% agarose gel at 100v for 40 min and stored at –20°C.

PCR products amplified from gDNA as previously described were purified using Wizard PCR Preps DNA Purification System (Promega). PCR products were transformed into plasmid vectors and cloned by utilizing TOPO TA Cloning Kit for Sequencing (Invitrogen, Carlsbad, CA). Prior to sequencing, plasmid DNA was purified with Wizard Plus Minipreps DNA Purification System (Promega). Purified plasmid DNA was sequenced by an ABI 3100 Genetic Analyzer at The University of Tennessee Molecular Biology Resource Facility, Knoxville, TN. Sequence data were analyzed using Chromas 2.23 (Technelysium Pty Ltd, Australia) and ClustalW (European Bioinformatics Institute, UK). After polymorphic loci were identified, sequences were then compared to previously available bovine CXCR1 sequences (GenBank) for homology.

All statistical calculations were carried out with SAS and SAS/Genetics version 9.0 (SAS Institute, Cary, NC). Observed allele and genotype frequencies were calculated using the ALLELE procedure. The HAPLOTYPE procedure used an expectation-maximization algorithm to obtain estimated haplotype frequencies. Independence between breed allelic, genotypic and haplotype frequencies were evaluated using chi square. Tests for allelic and genotypic

departures from Hardy-Weinberg equilibrium utilized chi-square goodness of fit. Tests for association between loci utilized chi square. All testing was done at $P = 0.05$ significance level.

Five single nucleotide polymorphisms were identified (see Figure 1) within a 311 bp segment (SEQ ID NO: 4) of the bovine IL-8 receptor locus at positions +612 (G→A), +684 (G→A), +777 (G→C), +858 (C→A) and +861 (A→G) relative to the available mRNA sequence (Genbank Accession No. U19947). Corresponding amino acid sequences revealed that a nonsynonymous substitution at +777 (G→C) results in histidine replacement of glutamine at amino acid residue 245, which occurs in the receptor's third intracellular loop. The four other polymorphisms result in synonymous nucleotide changes. Strong linkage disequilibrium was exhibited among all five polymorphic loci ($P<0.001$) as was expected due to the small size of the amplified gene segment and the close proximity of SNPs. All individual combinations of Holstein SNPs were in strong linkage disequilibrium ($P<0.001$), as were those in the Jersey ($P<0.01$) except between positions +684 and +858 ($P>0.05$).

Both Holstein and Jersey cattle exhibited all five single nucleotide polymorphisms, none of which departed from Hardy-Weinberg equilibrium ($P>0.05$; Table 1). Holsteins expressed all possible combinations of genotypes for each SNP, while none of the Jersey cows expressed the GG genotype for the polymorphism at +612 (data not shown). This was not unexpected, as the G allele at +612 represented less than 10% of the Jersey sample population. Even though the absolute frequency of allelic expression differed significantly between breeds ($P<0.00$) for all polymorphisms except for SNP +684 ($P>0.05$), the same alleles were dominant in both Holstein and Jersey cattle. Furthermore, Holsteins expressed increased heterozygosity for each SNP, ranging from 0.46 - 0.65, while Jersey heterozygosity was lower and more variable, ranging from 0.09 - 0.50 (data not shown).

Table 1. Estimated SNP allele frequencies within the CXCR1 locus of Holstein and Jersey dairy cattle.

	SNP	Jersey allele frequency (n=84)	Holstein allele frequency (n=74)
612*	A	0.92 (± 0.03)	0.58 (± 0.05)
	G	0.08 (± 0.03)	0.42 (± 0.05)
684	G	0.65 (± 0.05)	0.68 (± 0.05)
	A	0.35 (± 0.05)	0.32 (± 0.05)
777*	G	0.87 (± 0.04)	0.57 (± 0.05)
	C	0.13 (± 0.04)	0.43 (± 0.05)
858*	C	0.93 (± 0.03)	0.74 (± 0.05)
	A	0.07 (± 0.03)	0.26 (± 0.05)
861*	A	0.86 (±0 .04)	0.57 (± 0.05)
	G	0.14 (± 0.04)	0.43 (± 0.05)

Standard error in parentheses
* Allele frequencies affected by breed ($P<0.05$)

Greater specificity was obtained by combining single nucleotide polymorphisms together to form haplotypes. Ten haplotypes were derived from the five polymorphisms (Table 2). Both breeds expressed six common haplotypes, while two were unique to Jerseys and two were unique to Holsteins. In Jerseys, haplotypes AGGCA (1) and AAGCA (2) accounted for 83% of the total expressed haplotypes. In Holsteins, four haplotypes ((AGGCA (1), AAGCA (2), GGCAG (3), and GGCCG (8)) accounted for 95% of the total expressed haplotypes in Holsteins. Similar to individual SNPs, breed significantly affected haplotype frequencies ($P<0.0001$), potentially due to the greater allelic diversity within the Holstein population.

Table 2. SNP genotype frequencies within the CXCR1 locus of Holstein and Jersey dairy cattle.

SNP genotype	Jersey frequency	Holstein frequency

		(n=42)	(n=37)
612*	AA	0.83	0.27
	AG	0.17	0.62
	GG	0.00	0.11
684	GG	0.40	0.41
	AG	0.50	0.54
	AA	0.10	0.05
777*	GG	0.76	0.24
	GC	0.22	0.65
	CC	0.02	0.11
858*	CC	0.88	0.51
	AC	0.10	0.46
	AA	0.02	0.03
861*	AA	0.74	0.24
	AG	0.24	0.65
	GG	0.02	0.11

* Genotype frequencies affected by breed ($P<0.05$)

A phylogenetic analysis (Figure 2) of haplotypes resulting from all five SNP combinations present within the dairy population suggests three closely related groups: haplotype AAGGCA (1) alone; haplotypes AAGCA (2), GGGCA (7), AACCA (9), and GAGCA (10); and haplotypes GGCAG (3), AGCAG (4), AGGAG (5), AGCCG (6), and GGCCG (8). All haplotypes are tied together by single base mutations, except for the missing AGGCG connecting the upper two groups.

The polymorphic nature of the human and bovine IL-8R gene indicates that the gene may be an evolutionary hot spot for genetic mutation (Grosse et al. 1999; Heaton et al. 2001). Genetic variation may be accelerated in cattle when compared to humans due to increased selection pressure placed on specific traits such as milk production and carcass traits. Furthermore, the increased number of CXCR1 SNPs, haplotypes and haplotype genotypes expressed by both Holstein and Jersey cattle, when compared to beef cattle breeds, may be due to widespread use of artificial insemination in the dairy industry. As genotyping of different

breeds of cattle continues for the CXCR1 locus, more SNPs located within both the coding and noncoding regions may potentially be detected. Because of its highly polymorphic nature, CXCR1 may be an informative chromosomal region for association with disease phenotypes.

While no previous studies have associated bovine IL-8 receptor loci with immune responses during infection, two linked SNPs (+785 and +1208) within the human IL-8 receptor have been associated with susceptibility to fibrosing alveolitis and nonfibrosing alveolitis systemic sclerosis in humans (Renzoni, E., Lympany, P., Sestini, P., Pantelidis, P., Wells. A., et al.,, 2000. Distribution of novel polymorphisms of the interleukin-8 and CXC receptor 1 and 2 genes in systemic sclerosis and cryptogenic fibrosing alveolitis. Arthritis and Rheumatism 43, 1633-1640). While the polymorphism located at +785 did not result in an amino acid substitution and most likely does not affect receptor function, the +1208 SNP was located within the 3' untranslated region of the CXCR2 gene and may control mRNA processing and translation (Renzoni et al. 2000). While four of the five SNPs (+612, +648, +858, and +861) presently identified in dairy cattle are synonymous substitutions, the non-synonymous +777 (G→C) polymorphism results in a histidine substitution for glutamine at amino acid residue 245. This amino acid lies within the third intracellular loop when compared to the publicly available protein sequence for the bovine IL-8 receptor (Genbank No. AAA84996.1) and it is hypothesized that this could affect neutrophil activity during infection. Damaj et al. Damaj, B.B., McColl, S.R., Neote, K., Na, S.Q., Ogborn, K.T., Hebert, C.A., Naccache, P.A., 1996. Identification of G-protein binding sites of the human interleukin-8 receptors by functional mapping of the intracellular loops. FASEB Journal 10, 1426-1434) identified amino acid residues within the third intracellular loop of both human IL-8 receptors, CXCR1 and CXCR2, that were involved in mediating neutrophil calcium signaling and mobilization during IL-8 stimulation. Increased cytosolic free calcium is crucial for neutrophil respiratory burst and exocytosis during infection. Thus, while not

wishing to be bound by any particular theory, it is hypothesized that the amino acid substitution at residue 245 may affect CXCR1 signaling and neutrophil function in dairy cattle.

The bovine IL-8 receptor locus represents an excellent candidate gene marker for inflammation, such as during infection. In this study, we identified five SNPs within the bovine CXCR1 locus in two breeds of dairy cattle. A non-synonymous SNP at nucleotide +777 results in an amino acid substitution that was hypothesized to have functional implications.

Example 2

Animal selection was as described in Example 1. Each animal was genotyped as described in Example 1, by sequencing a 311-bp fragment (Figure 1; SEQ ID NO: 4) of the CXCR1 gene amplified from genomic DNA and inserted into a TOPO vector for cloning.

The animals were kept at dairy research facilities implementing conventional dairy cattle husbandry practices, including pre- and postmilking teat disinfection. All cows were dried off approximately 6 to 8 weeks before expected calving. An antibiotic preparation approved for use in nonlactating cows was infused into all cows' mammary glands following the last milking of lactation.

One foremilk sample from each quarter was collected from all lactating cows in both herds approximately every three months, prior to dry off, one week after calving and when cows exhibited clinical mastitis and were evaluated microbacteriologically. Prior to sample collection, teats of cows were cleaned and dried with individual disposable paper towels, and teat ends were sanitized with swabs containing 70% isopropyl alcohol. After sample collection, samples were stored frozen at –20°C until transported to the laboratory. Milk samples were examined following procedures recommended by the National Mastitis Council.

Samples of mammary secretion (10 μL) from each mammary gland were plated

onto one quadrant of a trypticase soy agar plate supplemented with 5% defibrinated sheep blood (Becton Dickinson and Company, Franklin Lakes, NJ).

Plates were incubated at 37°C, and bacterial growth was observed and recorded at 24-hour intervals for three days. Bacteria on primary culture medium were identified tentatively according to colony morphologic features, hemolytic characteristics, Gram stain reaction, and catalase test. Isolates identified presumptively as staphylococci were tested for coagulase production by the tube coagulase method. Isolates identified presumptively as streptococci were evaluated initially for growth in 6.5% NaCl, hydrolysis of esculin and sodium hippurate, and CAMP reaction. Streptococcal organisms were identified to the species level using the API 20 Strep system (bioMerieux Inc., Durham, NC) and a streptococcal agglutination system (Streptex, Remel, Lenexa, KS). Gram-negative isolates were identified to the species level using the following biochemical tests: triple sugar iron, urea, oxidase, motility, indole, and ornithine decarboxylase and by the API 20 E identification system (bioMerieux Inc.).

Clinical mastitis was defined as the presence of abnormal milk, and/or abnormal udder, and/or systemic signs of intramammary infection that warranted intramammary and/or anti-inflammatory therapies. Milking personnel identified cows with clinical mastitis and treatment was administered following the collection of samples from the infected quarter(s). Subclinical mastitis was defined as the detection of the same pathogen in the same mammary quarter in at least two out of three consecutive samples. Data for each cow were combined over time points and quarters to obtain percentages of subclinical and clinical mastitis cases. Final analyses were conducted using the percentages on a per cow basis. For example, at one collection time point a cow has four data points, one for each quarter. A cow with four lactations with five collection time points in each lactation had 80 total observations (4 quarters X 5 collection time points X 4 lactations = 80 total observations). If this cow had two subclinical events, she

would have a 2.5% percentage. The same would hold true if she experienced two clinical events.

Subclinical mastitis due to *Corynebacterium bovis*, coagulase-negative staphylococci (CNS) and *Staphylococcus aureus* were observed in the Holstein herd, while the majority of clinical mastitis was due to *Staph. aureus* and *Streptococcus uberis*. The majority of subclinical mastitis in the Jersey herd was due to CNS, *Strep. uberis* and *Staph. aureus*, while *Strep. uberis* and *Escherichia coli* caused the majority of clinical mastitis.

The number of somatic cells in milk was determined by Dairy Herd Improvement Association personnel at The University of Tennessee, Knoxville, TN with a Somacount 300 cell counter (Bentley Instruments, Chaska, MN). Somatic cell scores and projected 305-d mature equivalent milk yields were obtained from Dairy Records Management Systems (DRMS, Raleigh, NC).

All statistical calculations were carried out with SAS version 9.0 (SAS Institute, Cary, NC). A randomized block design blocking on lactation with covariates on somatic cell score and milk yield was used to determine effects of CXCR1 SNPs and haplotype genotypes on the percent incidence of subclinical and clinical mastitis over lifetime for each breed. Analysis of variance was used to detect significance differences between data. Data are presented as least squares means with associated standard error. Randomized block designs blocking on lactation were used to determine effects of CXCR1 SNPs on somatic cell scores (SCS) and milk yield. Correlations between percentage of subclinical and/or clinical mastitis incidence, SCS, and milk yield were determined. Statistical significance was declared at $P < 0.05$.

Holstein cows exhibited higher estimated means of percent subclinical mastitis for all SNPs when compared to Jersey cows. SNPs +612, +777 and +861 showed a significant association with subclinical mastitis ($P < 0.05$). Further attention focused on SNP +777, as SNPs +612 and +861 are synonymous polymorphisms. SNP +777 (G to C) results in a Gln^{245} to His^{245} replacement,

which may affect mastitis phenotype. Cows with genotype CC at SNP +777 expressed increased percentages of subclinical mastitis when compared to cows with genotype GG ($P < 0.05$), 37% versus 22%, respectively (Figure 3). This difference was not observed in Jerseys (Figure 3) ($P > 0.10$). Significant differences were not detected in either breed between SNP +777 genotypes and percentage observations positive for clinical mastitis (Figure 4). However, Holstein cows tended to have lower incidence of mastitis when the CC genotype was expressed compared to the GG phenotype.

The relationship between SNP +777 genotypes with SCS was evaluated also. SNP +777 had a significant association with SCS in Holsteins ($P < 0.001$); however, this polymorphism was not significant for Jersey SCS differences (Table 3). Cows expressing genotypes GG and GC had increased SCS relative to the CC genotype; however, only cows expressing the GC genotype had significantly greater SCS than those expressing genotype CC. No significant differences were observed for Jersey SNP +777 genotypes and SCS, which ranged from 3.09 to 3.40.

Table 3. Bovine CXCR1 single nucleotide polymorphism (SNP) +777 genotype by SCS and milk yield over all lactations.

Holstein SNP +777 genotype	Estimated SCS	Estimated 305d milk yield (kg)
GG	3.10 (± 0.15)[a]	9412 (± 151)[b]
GC	4.00 (± 0.11)[b]	10,231 (± 112)[a]
CC	2.72 (± 0.20)[a]	10,056 (± 220)[a]
Jersey SNP +777 genotype	Estimated SCS	Estimated 305d milk yield (kg)
GG	3.09 (± 0.07)	8419 (± 53)

GC	3.35 (± 0.13)	8585 (± 105)
CC	3.40 (± 0.40)	8629 (± 326)

[a,b] signify LSM separation within herd and column ($P < 0.05$).

We observed no significant differences between SNP +777 genotype and 305-day milk yield in Jerseys, as estimated means ranged from 8419 to 8629 kg (Table 3). However, Holsteins expressing genotype GG had significantly lower 305-day milk yields (9412 kg) versus Holsteins expressing genotypes GC and CC (10,231 and 10,056 kg), respectively.

We previously identified 10 CXCR1 haplotypes created by five SNPs located within the predicted coding region of the bovine CXCR1 gene (Example 1). CXCR1 SNPs +612, +777 and +861 were significantly ($P < 0.05$) associated with incidence of subclinical mastitis. We focused our attention on SNP +777 as it resulted in a change at amino acid 245 (Gln to His). As such this nonsynonymous polymorphism may be subject to natural selection because it may ultimately affect mastitis phenotype (Hannson, B., Westerberg, L., 2002. On the correlation between heterozygosity and fitness in natural populations. Molecular Ecology 11, 2467-2474).

Due to potential effects caused by the nonsynonymous SNP +777 (G to C) on CXCR1 signal transduction and function, we examined the association of SNP +777 genotypes with percentages of subclinical and clinical mastitis incidence. A non-significant decrease of clinical mastitis incidence was observed in cows that expressed genotype CC. A significant difference ($P < 0.05$) between subclinical mastitis incidence and SNP +777 genotype was identified also. Holsteins expressing the GG genotype had a lower incidence of subclinical mastitis than cows that expressed the CC genotype. Interestingly, greater 305d milk yield was observed in Holsteins that expressed genotype CC when compared to those

expressing GG and would support prior research that concluded cows selected for greater milk yields also have increased mastitis infections.

More surprising was the lack of significant SCS differences between SNP +777 genotypes GG or CC, as we expected cows with a CC genotype to have greater concentrations of somatic cells due to more bacterial infections. However bacterial presence does not always increase SCS. The correlation between SCS and clinical mastitis has typically ranged from 0.6 to 0.8 suggesting the relationship between SCS and clinical mastitis is not absolute. The absence of increased SCS in spite of bacterial presence suggests a potential defect in one or more of the mechanisms necessary for neutrophil migration to the mammary gland, including bacterial recognition, cell-to-cell communication, and/or intracellular communication.

The amino acid change resulting from the G to C substitution at SNP+777 offers a possible explanation for increased bacterial presence in Holstein cattle expressing the CC genotype. During the onset of mastitis, different pathogens invade the mammary gland and elicit cytokine and chemokine release from macrophages and surrounding epithelial cells. Inflammatory mediators, such as interleukin-8, growth related oncogene-α, and neutrophil activating peptide-2, bind to CXCR2 receptors and induce CXCR2 internalization. The subsequent cascade of signals result in the release of intracellular Ca^{2+} and granular enzymes, as well as initiating chemotaxis (Ahuja, S.K., Lee, J.C., Murphy, P.M., 1996. CXC chemokines bind to unique sets of selectivity determinants that can function independently and are broadly distributed on multiple domains of human Interleukin-8 receptor B. The Journal of Biological Chemistry 271, 225-232). While studies have identified binding sites in the N' terminus responsible for binding selectivity and in the C' terminus for receptor phosphorylation, few have focused on identifying functional roles of intracellular loops (Ahuja et al. 1996).

Damaj et al. (1996) determined that Met^{241} is important in G-protein binding and intracellular Ca^{2+} mobilization. Damaj et al. (1996) also determined other amino

acids in the third intracellular loop were moderately involved in mediating intracellular signaling. As the polymorphism at +777 results in an amino acid change at position 245 in the third intracellular loop, it has the potential to interfere with G-protein mediated signaling and subsequent neutrophil functions such as migration. This hypothesis is supported by the lack of increased SCS in cows expressing the CC genotype.

Another potential explanation for increased subclinical mastitis in Holstein cattle with the CC genotype at SNP+777 may be due to expression of other genes located near CXCR1 on BTA 2. NRAMP1 is a potential gene affecting mastitis and related immune responses as it is located near CXCR1 and has been shown to be polymorphic and associated with immune function. Single nucleotide polymorphisms located within other regions of the bovine CXCR1 gene may also affect CXCR1 function during mastitis infections. Several single nucleotide polymorphisms have been detected within different regions of the human CXCR1 gene. As such, other nonsynonymous SNPs located within segments of the CXCR1 coding sequence may affect receptor binding and function. Additionally, either synonymous or nonsynonymous SNPs in the promoter and 3' untranslated regions may affect initiation and/or termination of transcription or translation.

In conclusion, in these experiments we identified three polymorphisms within the CXCR1 gene that were associated with subclinical mastitis. This research is promising as it may represent an effective means of marker-assisted selection for mastitis resistance and potentially other inflammatory diseases. Moreover, the ability to genetically identify mastitis susceptible versus resistant animals provides an opportunity to determine what mechanisms are linked to disease susceptibility and may open the door to novel preventive and therapeutic measures for mastitis.

Example 3

The strength of the foregoing findings suggested that the identified polymorphisms in CXCR1 were within a locus informative for mastitis and represented quantitative trait nuclei (QTN) not located within an identified QTL for mastitis. Furthermore, the +777 SNP also has been associated with impaired neutrophil function – a key component of immunity against bacterial infection (Rambeaud, M. and G.M. Pighetti, Impaired neutrophil migration associated with specific bovine CXCR2 genotypes. Infect Immun, 2005. 73: p. 4955-4959; Rambeaud, M., R. Clift, and G.M. Pighetti, Association of a bovine CXCR2 gene polymorphism with neutrophil survival and killing ability. Vet Immunol Immunopathol, 2006. epub ahead of print). These effects included altered adhesion molecule expression, migration, reactive oxygen species generation, apoptosis, and calcium signaling. It was therefore considered useful to evaluate additional polymorphisms within CXCR1 and genes in the immediate vicinity, to determine whether additional informative markers could be identified. An evaluation of additional newly identified polymorphisms within the coding and untranslated regions of CXCR1, as well as NRAMP1, was undertaken with respect to frequency and strength of association with mastitis, milk, milk components, and reproductive measures.

Animal husbandry procedures, blood collection procedures, and milk collection and analysis procedures were as described in Examples 1 and 2. Identification of clinical and subclinical mastitis in milk samples was as in Examples 1 and 2. Bovine genomic DNA was isolated from blood samples of Holstein cows with known phenotypic information by the use of a commercial kit (MoBio, Carlsbad, CA). Amplification of selected areas of the CXCR1 gene was performed under the following conditions: 40ng of bovine genomic DNA was used as template in a 20ul reaction containing specific primers and Eppendorf HotMaster Mix (Eppendorf North America; Westbury, NY) according to manufacturer's guidelines. The primers used are set forth in Table 4.

Table 4. Primers developed for analysis of CXCR1 and NRAMP1 SNPs

Primer_ID	Primer_Location	Primer_seq	Range	SEQ ID NO
CXCR-1-F	AC_150887.154285	CCATATTGGGCCTGTGGA	AC_150887. 154285 - 155143	5
CXCR-1-R	AC_150887.155143	TCAGCAAGCATGGGGACT		6
CXCR-2-F	AC_150887.153621	ATCAGGGGTTGAGGTTGG	AC_150887. 153621 - 154423	7
CXCR-2-R	AC_150887.154423	CAGACGGAATCTTTGTCCTG		8
CXCR-3-F	AC_150887.152796	GGAAAGTCCCCTTCCCTTG	AC_150887.152796 - 153677	9
CXCR-3-R	AC_150887.153677	CATCCCATCCCTGGTTCC		10
CXCR-4-F	AC_150887.152103	TCATCTTTGCTGTCGTGCTC	AC_150887.152103 - 153008	11
CXCR-4-R	AC_150887.153008	CAGCCATGGTGGTAGCAT		12
CXCR-5-F	AC_150887.151393	GGAGGGGTTTGAGGATGA	AC_150887.151393 - 152293	13
CXCR-5-R	AC_150887.152293	ACGTAGATGAGGGGGTTGA		14
CXCR-6-F	AC_150887.150630	TTCAGCACTTTCGCTGCT	AC_150887.150630 - 151434	15
CXCR-6-R	AC_150887.151434	GCGTGCCGCTGTAATTT		16
CXCR-7-F	AC_150887.149921	ACAGCACGAGCACTGGAT	AC_150887.149921 - 150781	17
CXCR-7-R	AC_150887.150781	CTGGTGGGCTACTGTCCA		18
CXCR-8-F	AC_150887.149144	TGGGGCAGGGTGACTTA	AC_150887.149144 - 150014	19
CXCR-8-R	AC_150887.150014	ACCTCCAGGGCTCTTCC		20
CXCR-9-F	AC_150887.148419	CTCCTCCAATCACCACCA	AC_150887.148419 - 149238	21
CXCR-9-R	AC_150887.149238	AAGGAGGCAGGGAGTCAA		22
CXCR-10-F	AC_150887.147697	TCCAGGGACTACCCCATC	AC_150887.147697 - 148538	23
CXCR-10-R	AC_150887.148538	CCACCACCAACGAAGAAC		24
CXCR-11-F	AC_150887.146852	GTTCACACGGTCCTGCAT	AC_150887.146852 - 147714	25
CXCR-11-R	AC_150887.147714	ATGGGGTAGTCCCTGGAA		26
NRAMP1-K1-F	NW_301097.1.302	CAGCAGCCTCCACGACTA	NW_301097.302-319	27

NRAMP1-K1-R	NW_301097.1.626	AAACTGTCCCGCGTAGGT		28
NRAMP1-C1	AY438096.1_1223	GTGGCAGGAAGTTGTCCA		29
	NW_490962.1_579	GGCCCCACCAAATAAGAG		30
NRAMP1-F1	NW_490962.1_1517	GCTGCAGAGGGAGGAGTT	NW_490962.1_1517-1953	31
	NW_490962.1_1953	GCCTGAAGATCCGACTCA		32
NRAMP1-G1	NW_490962.1_1903	CATGAGCATCGCATTCCT	NW_490962.1_1903-2401	33
	NW_490962.1_2401	GGGTAGTAGAGATGGCAGACC		34

The conditions for amplification were as follows: an initial hot-start denaturation occurred at 94°C for 2min, followed by 37 cycles of 94°C denaturation for 30 sec, 61°C (58°C for some primer pairs) annealing for 30sec, and 68°C extension for 45sec. After the last cycle, a 10 minute final extension step at 68°C was added before reactions were chilled to 4°C. Amplified products were purified to remove primer and excess nucleotides by treatment with exonuclease-shrimp alkaline phosphatase (Exosap-it, USB Corporation; Cleveland, OH) as per the manufacturer's guidelines. The sequencing reaction was performed using Applied Biosystems Big Dye Terminator v3.1 Cycle Sequencing kit and the subsequent products ran on an ABI prism 3100 Genetic Analyzer 16 capillary array unit. Sequencing reactions were run in both directions (i.e. using both forward and reverse amplification primers). Detection and genotyping of SNP were carried out using Sequencher 4.2 software (Gene Codes Corp., Ann Arbor, MI). Genotypes were called independently in forward and reverse sequence traces and compared. If non-consensus occurred (most often because of poor sequence quality), the amplification, sequencing and genotyping were repeated in both directions.

Observed frequencies for each SNP were calculated using a commercial spreadsheet (Microsoft Excel). To identify linkage groups for association analysis a freeware program (LDSelect) was utilized (Carlson, C.S., M.A. Eberle, M.J. Rieder, Q. Yi, L. Kruglyak, and D.A. Nickerson, Selecting a Maximally Informative Set of Single-Nucleotide Polymorphisms for Association Analyses Using Linkage Disequilibrium. Am. J. Hum. Genet., 2004. 74: p. 106-120). This program determines the correlation among various SNP and places them into bins that group SNPs present at a specified frequency and correlated at a

requested level. In the present analysis, the minor allele frequency was set at 10% and correlation (r2) minimum was 80%. Tag SNP which represented the genetic variation in the region were partitioned into main effects by including single SNP terms, one at a time for individual SNP evaluation. For evaluating trait association with individual SNP, all statistical calculations were carried out with SAS version 9.0 (SAS Institute, Cary, NC). Individual SNP were evaluated using two types of models: genotype and allele substitution (or dose). The effect of allele substitution was estimated using linear regression on the number of X alleles in the SNP genotype. The X allele is represented by the polymorphic nucleotide that appears first alphabetically. The primary model included random effects of year and season, covariates of days in milk (DIM), SCS, linear effect of lactation number, and quadratic effects of lactation number, as well as an error term. Because sample sizes for each genotype are unbalanced, raw means also were computed to guard against unrealistic adjustments by the statistical model. All analyses were summarized using least squares means separated by Tukey's Significant Difference to reduce chances of Type I errors. Significance was declared at $P<0.05$ and a trend towards significance at $P<0.10$. To evaluate trait association with multiple SNP, initially, each of the five tag SNP were included in the mixed model and evaluated for their combined association with mastitis, reproduction, and milk traits. However, due to the relatively small sample size, the resulting adjusted means were not in agreement with the raw means. One SNP was removed from the model at a time in order to identify the most combinations that could be included in the model without gross alteration of the raw mean.

Production and reproductive data measured on a continuous scale (e.g. milk, fat, protein, somatic cell score, days open, and days to first service) were analyzed using mixed models (SAS Institute, Version 9.1) controlling for year and season effects. The covariates outlined previously were used in all models except in the case of days open where DIM was not included in as a covariate due to the conflict with a defined voluntary waiting period. Clinical mastitis was divided into 3 periods of 100 d each, and number of mastitic quarters during each of these periods, as well as during the entire lactation was analyzed using a Poisson linear model with log link (Proc GENMOD, SAS Institute). Days in milk during that period was used as a measure of risk exposure. Data were apparently too unbalanced for generalized mixed models to converge satisfactorily. Standard mixed models including random effects of year and season, covariates of SCS, linear effects of lactation number, and quadratic effects of lactation number, and correct error terms also were run to identify results that were model sensitive. In contrast, subclinical mastitis was evaluated for the entire lactation only due to the definition of subclinical mastitis which requires 2 of 3 consecutive samples to contain the same organism and prevents subdivision of the lactation by time. Days in milk were used as the measure of exposure, and linear and quadratic effects of lactation number included as covariates.

A total of 49 SNPs were identified in and near the bovine CXCR1 gene (4,767 base coverage) and the frequency of 43 SNP assessed in 96 Holstein cows. The position of each SNP relative to a published reference sequence for the bovine CXCR1 gene (Genbank Accession No. AC150887.4; SEQ ID NO: 35) is set forth in Table 5. Of these SNP, 13 were in the coding region, 4

introduced amino acid changes, and 1 introduced a stop codon. The remaining 30 SNP were located in the first intron, 3' untranslated region, and 3' to the gene. In NRAMP1, five SNP were identified. The SNP positions relative to the reference sequences, Genbank Accession Nos. NW_490962.1 and NW_301097 (SEQ ID NO. 36), are provided in Table 6. In NRAMP1, five SNP were identified in the 1,979 bases examined: 2 in the promoter region, 2 intronic, and 1 in the coding region which introduced an amino acid change. Only the non-synonymous NRAMP1 SNP (K1 245; position 546 of Genbank Accession No. NW_301097; C→G) was typed in all 96 cows.

Table 5: Identified polymorphisms in bovine CXCR1 gene on BTA2						
Primer set[4]	Position in Amplicon[1,2]	Position in Reference Sequence	SNP	SNP and Flanking sequence	RFLP[3]	SEQ ID NO.
CXCR_1	274	AC150887.4_154558	T/C	TACTTAAGTAACA[T/C]AATGTATGAAAAT		37
CXCR_1	392	AC150887.4_154676	T/C	GTATTCTTGCC[T/C]GGAAAATCCCAT		38
CXCR_3	118	AC150887.4_152913	G/A	CTTAAAAATAATG[G/A]TGTGCTGGTAAAT		39
CXCR_3	428	AC150887.4_153223	G/A	TATAGGATATGAT[G/A]TGTGGAGCTGAAG		40
CXCR_3	490	AC150887.4_153285	G/A	GGCTGATGGAGTCA[G/A]AGGATAATAAG	HPY188 I	41
CXCR_3	502	AC150887.4_153297	C/G	AGGATAATAAG[C/G]ACCTGAGAATAA	DRA II	42
CXCR_3	703	AC150887.4_153498	T/C	ATGGATCAAAGAC[T/C]TACAATTAGAG		43
CXCR_3	715	AC150887.4_153510	T/C	TACAATTAGAG[T/C]TAAAATTATAAAA		44
CXCR_3	736	AC150887.4_153531	G/A	TAAAACTCCTGG[G/A]AAAAAACTTAAG		45
CXCR_4	62	AC150887.4_152164	C/A	TCCTGAT[C/A]GCGGACACCCTC		46
CXCR_4	65	AC150887.4_152167	G/A	TCCTGAT[A/C]GC[G/A]GACACCCTC		47
CXCR_4	226	AC150887.4_152328	A/G	GACTCCTCA[A/G]GATCATGGCC		48
CXCR_4	241	AC150887.4_152343	G/A	ATCATGGCCATCC[G/A]TGGCCTGATCAG	NCO I	49
CXCR_4	254	AC150887.4_152356	C/T	GCCTGATCAG[C/T]AAGGAGTTCTTG		50
CXCR_4	314	AC150887.4_152416	G/A	TTCAGGGAACAC[G/A]TCTACTACCCT	ACC I	51
CXCR_4	**335** 336 339 344	AC150887.4_152437_152438_152441_152446	A/C	CTCTGAGACTC[A/C][A/C]CA[C/T]CGGG[C/T]TCCTCGCTTCC		52
CXCR_4	336	AC150887.4_152438	A/C			53
CXCR_4	339	AC150887.4_152441	C/T			54
CXCR_4	344	AC150887.4_152446	C/T		SDU I	55
CXCR_4	426	AC150887.4_152528	A/G	CATTGT[A/G]GTTACAG		56
CXCR_4	510	AC150887.4_152612	G/A	ACCTCTTGCT[G/A]TAATCACCAT		57
CXCR_4	528	AC150887.4_152630	G/A	CACCATGTCAACT[G/A]ATAGAGCTTGG		58
CXCR_4	593	AC150887.4_152695	A/C	GACATCCTTTGA[A/C]GATATTCTTT		59
CXCR_4	604	AC150887.4_152706	A/T	GATATTCTTT[A/T]CA[G/A]TGATTACAAAA		60
CXCR_4	607	AC150887.4_152709	G/A	GATATTCTTT[A/T]CA[G/A]TGATTACAAAA		61
CXCR_4	619 624 627 630	AC150887.4_152721	A/C	TGATTACAAAA[A/C]GTTC[C/T]CA[C/T]TA[A/T]TGGGAACCATT	ACL I	62
CXCR_4	624	AC150887.4_152726	C/T			63
CXCR_4	627	AC150887.4_152729	C/T			64
CXCR_4	630	AC150887.4_152732	A/T			65
CXCR_4	645	AC150887.4_152747	G/A	TGGGAACCATTAAT[G/A]TAAAACACCAT		66
CXCR_4	811	AC150887.4_152913	G/A	TAAAAATAATG[G/A]TGTGCTGGTAAAT		67
CXCR_5	158	AC150887.4_151551	A/C	GGGAAACTCC[A/C]TGGTGATGCT	NCO I	68

CXCR_5	247	AC150887.4_151640	C/T	GACCTGCTCTT[C/T]GCCATGACCTT	SAP I		69
CXCR_5	260	AC150887.4_151653	A/C	CCATGACCTTG[A/C]CTATCTGGGCCG			70
CXCR_5	321	AC150887.4_151714	C/T	TGTGCAAGGTGG[C/T]CTCACTCCTGAAG	Eco31 I		71
CXCR_5	522 **526**	AC150887.4_151915 AC150887.4_151919	A/C **G/A**	TGAGGCCTATCAAC[A/C]ACC[**G/A**]TACTCCGACCTAG			72
CXCR_5	577	AC150887.4_151970	G/A	GAAATG[G/A]CGGATGAT			73
CXCR_5	597 **598**	AC150887.4_151990 AC150887.4_151991	G/T **G/A**	GATAATGCGTGTCC[G/T][**G/A**]CCCCAGACCTTTGG			74
CXCR_5	**691**	AC150887.4_152084	**C/G**	CAAATGGGGCA[**C/G**]AAGCACCGGGCCAT	SDU I		75
CXCR_5	**772 775**	AC150887.4_152165 AC150887.4_152168	**A/C** **G/A**	CTGGTCCTGAT[**A/C**]GC[**G/A**]GACACCCTCA			76
CXCR_6	113	AC150887.4_150742	G/A	CATCTGAGAGCA[G/A]GTTTTGTCTTTG	BSPM I		77
CXCR_6	239	AC150887.4_150868	C/T	GATCAAACCCA[C/T]ATCTGCTGCATCAG			78
CXCR_6	255	AC150887.4_150884	A/T	TCTGCTGCATCAGT[A/T]GGCAGTTTCTTTA			79
CXCR_6	343	AC150887.4_150972	A/T	GGCCCTTCCCTGAA[A/T]GTATAGAACAACTT			80
CXCR_6	515	AC150887.4_151145	C/T	CCCACATTTAGGCCT[C/T]GGGGTTCTGGAAAG	AVA I STY I		81
CXCR_6	539	AC150887.4_151168	C/T	AAAGCTGACAACA[C/T]GATGAAGCTGATCCTT			82
CXCR_6	679	AC150887.4_151308	A/C	ATCTGAGTTACTAA[A/C]TCACCTTTCTTTC			83
CXCR_11	238	AC150887.4_147089	C/T	ATAAA[C/T]GGAAAG			84
CXCR_11	634	AC150887.4_147485	A/G	AGATGAG[A/G]CTTTTCTT			85

[1]The SNPs are denoted by their position within the amplified DNA fragment (amplicon), their position within the bovine BAC sequence (AC_150887), and within the sequences that flank the SNP or combination of SNPs.

[2]Those sequences in bold represent SNPs previously identified by Pighetti and are the equivalent of +612, 684, 777, 858, and 861, respectively

[3] RFLP indicates which restriction enzyme is capable of distinguishing the SNP by analysis of whether the DNA sequence can or cannot be cleaved by the enzyme.

[4]PCR conditions: Eppendorf HotMaster Mix. 94C x 2 min; 94C x 30 sec -> 58C x 30 sec -> 68C x 30 sec for 35 cycles; 68C x 10 min.

Table 6: Identified SNPs in the bovine NRAMP1 gene.

Primer set[3]	Position in Amplicon[1]	Position in Reference Sequence	SNP	SNP and Flanking sequence	RFLP[2]	SEQ ID NO
NRAMP1-C1	247	NW_490962.1_53	C/A	CAATCGCCACCGC[C/A]TAGAGAGG	no	86
NRAMP1-C1	584	NW_490962.1_390	C/T	TGGTGTCAGTAT[C/T]GCCTGCTCGGAA	no	87
NRAMP1-F1	251	NW_490962.1_1767	C/T	GGCCAGCCTGGG[C/T]CCCTGAGCCTGAG	APA I	88
NRAMP1-G1	85	NW_490962.1_1987	G/A	CATGGAGTC[G/A]GGGCACCATGGGG	no	89
NRAMP1-K1	245	NW_301097_546	C/G	GCCTCTTTGGT[C/G]CTCCAGCCCTGTAC	PST I	90

The SNPs are denoted by their position within the amplified DNA fragment (amplicon), their position within the bovine BAC sequence (NW_490962 or NW_301097), and within the sequences that flank the SNP or combination of SNPs.

RFLP indicates which restriction enzyme is capable of distinguishing the SNP by analysis of whether the DNA sequence can or cannot be cleaved by the enzyme.

PCR conditions: Eppendorf HotMaster Mix. 94C x 2 min; 94C x 30 sec -> 58C x 30 sec -> 68C x 30 sec for 35 cycles; 68C x 10 min.

Five linkage groups were identified for the two gene regions CXCR1 and NRAMP1 (Table 7). It was considered that choosing a SNP from each group would provide the most comprehensive coverage of the gene area with a minimum number of SNP. In addition, any of the SNP within that group or bin can be utilized since they are highly correlated (>0.8) with each other. This allows selection of a tag SNP within the bin that has functional significance, position significance, and/or is easier to identify. Those chosen for subsequent analysis were 4-226 (position 152328 of AC150887.4), 4-62 (position 152164 of AC150887.4), 5-577 (position 151970 of AC150887.4), 5-691 (position 152084 of AC150887.4), and NRAMP1 (position 546 of NW_301097). Four of the five previously identified SNPs in CXCR1 (See Examples 1 and 2) were identified. Three are present in bin 1: 5-526 is +612, 5-691 is +777, and 5-775 is +861. As such, 5-691 (+777) was chosen to represent bin 1 and maintain continuity with prior research. Bin 3 contains SNP 5-772, which was previously (Example 1) identified as SNP +858. This SNP was not utilized as the tag SNP due to the presence of nearby SNP which complicated detection. In its place, 4-62 was utilized. The other representative tag SNP chosen were 4-226 for bin 2, 5-577 for bin 4, and NRAMP1 for bin 5.

The 5-691 tag SNP, which as noted above is the +777 SNP in the CXCR1 gene (see Example 1), encodes an amino acid change (Q → H) at position 245 within the third intracellular loop of the CXCR1 protein. This amino acid is located near a G-protein binding site and has the potential to interfere with signaling. The G nucleotide is present in other species and encodes for glutamine (Q), while the C nucleotide encodes for histidine (H). The 4-226 tag SNP also encodes for an amino acid change (K → R) at position 327 in the C terminus.

The G nucleotide encodes for lysine which is conserved across species, while A encodes for arginine. This amino acid is located near a binding site for PP2A, a serine/threonine phosphatase. The 4-62 tag SNP is located in the coding region but does not induce an amino acid change. The 5-577 tag SNP (A nucleotide) introduces an early stop codon at amino acid 206 which would either allow expression through the second intracellular loop or may not allow expression. Either scenario would have the potential to significantly impact responses to its ligand interleukin-8. The final tag SNP in NRAMP1 induces an amino acid change (A → P) at position 356. The G nucleotide encodes for alanine which is conserved in other species, whereas the C nucleotide encodes for proline which represents the most frequent allele in our sample Holstein population. This amino acid is located in the 7th transmembrane domain. The frequency of each tag SNP is shown in Table 8.

Table 7. Linkage groups as determined by LDSelect.

Bin	TagSNP	Average minor allele frequency
1	1-274[1], 1-392, 3-118, 3-490, 3-502, 3-703, 3-736, 4-065, 4-314, 4-335, 4-336, 4-339, 4-344, 4-528, 4-604, 4-607, 4-619, 4-624, 4-630, 4-645, 4-811, 5-321, 5-526, **5-691**[2,3], 5-775, 11-238	29%
2	**4-226**, 4-241, 4-426, 6-239, 6-255, 6-539, 11-634	32%
3	3-428, 3-715, **4-62**, 4-254, 5-247, 5-772	18%
4	**5-577**, 6-515	12%
5	**NRAMP-1**	0%

[1]SNP are identified by amplicon and amplicon position.
[2]SNP in bold were selected for trait association analysis.
[3]Previously identified SNP and their equivalents: +612 = 5-526; +777 = 5-691; +861 = 5-775; +858 = 5-772

Table 8. Frequency of identified tag SNP.

Bin	SNP	Genotype		Genotype		Genotype	
1	5-691	CC	30%	CG	45%	GG	25%
2	4-226	AA	15%	AG	35%	GG	50%
3	4-62	AA	6%	AC	30%	CC	64%
4	5-577	AA	10%	AG	38%	GG	52%
5	NRAMP-1	CC	55%	CG	45%	GG	0%

Individual SNPs were analyzed with respect to phenotypic traits of interest. The analysis included an assessment of the association with relevant genotypes for each SNP, as well as using an allele substitution model for each SNP. With the relatively small sample size and variation present in samples, the allele substitution model was generally less informative than genotype models. Therefore, the allele dose models were presented when applicable. The data presented are for $P < 0.10$ or if the r2 value was greater than 0.01.

Two SNP were identified as being potential markers for clinical mastitis: 4-226 and NRAMP1, (Table 9). In the sample population, the AA (0.744 ± 0.2) and AG (0.830 ± 0.2) genotypes at position 4-226 tended to have more ($P<0.10$) cases of mastitis per year than the GG genotype (0.422 ± 0.15). Subsequent evaluation of allele specific effects with a substitution model revealed an increase of $\sim 0.18 \pm 0.1$ clinical cases ($P=0.11$) with each A allele. With each model, the GG genotype has the fewest number of clinical mastitis cases, while there is slight disagreement with respect to the affect of the A allele. In contrast, cows with CG genotypes for the NRAMP1 SNP experienced significantly more cases of clinical mastitis (0.842 ± 0.1) than cows with the CC genotype (0.454 ± 0.1). No cows with a GG genotype were present in our sample population, however examination of the allele substitution model indicates that addition of each C allele reduces clinical mastitis by 0.36 ± 0.2 cases per year ($P<0.025$).

Table 9. Tag SNP association with clinical mastitis.

SNP	Genotype	LS Mean	SE	Tukey	P<	SNP r^2	Total r^2
4-226	AA	0.7	0.2	A	0.08	0.019	0.213
	GG	0.4	0.2	A			
	AG	0.8	0.2	A			
NRAMP1	CC	0.4	0.1	B	0.05	0.021	0.208
	CG	0.8	0.1	A			

The 4-226 SNP genotypes also tended to be associated (P<0.10) with subclinical mastitis (Table 10): the AG genotype had the greatest level of subclinical mastitis (3.8 ± 1.0), the AA genotype was intermediate (2.7 ± 1.2), and the GG had the lowest (1.8 ± 0.9). Substitution of the G allele with the C allele resulted in an increase of 0.85 ± 0.5 cases of subclinical mastitis per year (P<0.10). These results correspond to the differences observed with respect to clinical mastitis for both the genotypic and allele substitution models, suggesting the 4-226 is associated with both clinical and chronic subclinical mastitis. Two additional SNP were observed to be associated with subclinical mastitis: 5-577 and 5-691. The 5-577 SNP introduces a stop codon into the CXCR1 gene. Cattle with an 'AA' genotype had 0.1 ± 1.5 average cases compared to 3.1 ± 0.9 and 3.0 ± 0.9 for the GG and AG genotypes, respectively. This SNP accounted for ~ 3% of the variation in the sample means and with a larger sample population the differences between genotypes may prove to be significant (currently P<0.15). The genotypes representing 5-691 appeared to have similar means as the 4-226 SNP, cows with the CC genotype had more clinical cases of mastitis (3.9 ± 1.1), those with the CG genotype were intermediate (2.6 ± 0.9), and those with the GG genotype were lowest (1.6 ± 1.1). With the aid of the allele substitution model it is estimated that an additional 1.2 ± 0.5 cases of subclinical mastitis occur with

each C allele (P<0.025). This was expected, as 5-691 is the same as +777 (see Example 1).

Table 10. Tag SNP association with subclinical mastitis.

SNP	Genotype	LS Mean	SE	Group	P<	SNP r^2	Total r^2
4-226	AA	2.7	1.1	AB	0.10	0.043	0.268
	GG	1.8	0.9	B			
	AG	3.8	1.0	A			
5-577	AA	0.1	1.5	A	0.15	0.028	0.265
	GG	3.1	0.9	A			
	AG	3.0	1.0	A			
5-691	CC	3.9	1.0	A	0.15	0.022	0.274
	GG	1.6	1.1	A			
	CG	2.6	0.9	A			

A slightly different pattern of SNPs were associated with the average lactation somatic cell score: 5-577 and NRAMP1 (Table 11). For the 5-577 SNP, cows with the GG genotype had the greatest mean SCS (2.7 ± 0.3) compared to cows with either an AA (2.2 ± 0.5) or AG genotypes (2.3 ± 0.10). Part of the basis for greater SCS in cows with the GG genotype may be related to the greater number of subclinical cases that also were observed in these cows, relative to the lower rates of subclinical mastitis and SCS in cows with the AA genotype. The presence of fewer somatic cells in cows with an AA or AG genotype also may be directly related to the introduction of a stop codon with the A allele which has the potential to either prevent expression of the receptor or impair function of the interleukin-8 receptor, CXCR1. Furthermore, based upon the allele substitution model, each increase in the A allele reduces SCS by 0.3 ± 0.2 (P<0.10). A similar set of SCS were observed with the SNP in NRAMP1, with cows with the CC

genotype having a mean score of 2.8 ± 0.3 and cows with the CG genotype having a slightly lower mean score of 2.3 ± 0.3.

With respect to both 5-577 and NRAMP1 SNP, it should be noted that SCS represent somatic cell counts under 100,000 cells/ml and the gland is considered to be free of infection. This suggests that the dairy facility performed well with respect to maintaining udder health. Another interesting note is that the most common genotypes for each SNP (GG for 5-577 and CC for NRAMP1) coincide with the Holstein average SCS of 2.74 in 2004.

Table 11. Tag SNP association with SCS.

SNP	Genotype	LS Mean	SE	Group	P<	SNP r^2	Total r^2
5-577	AA	2.2	0.5	A	0.17	0.015	0.187
	GG	2.7	0.3	A			
	AG	2.3	0.4	A			
NRAMP1	CC	2.8	0.3	A	0.10	0.012	0.182
	CG	2.3	0.3	A			

Two traits were evaluated with respect to reproduction – days open and days to first service. The number of days open tended to be associated with SNP at positions 4-226, 5-577, and NRAMP1, similar to that observed with clinical and subclinical mastitis (Table 12). For 4-226, cows with the AA genotype were open for the fewest days (118 ± 18) and those with the AG genotype were open for the most days (165 ± 14), while cows with a GG genotype were intermediate between the two (143 ± 11). Polymorphisms at position 5-577 were not significantly associated with days open (P>0.5), however, it accounted for ~1% of the mean variation in the sample population and also was associated with subclinical mastitis. Those cows with an AA genotype at this position were open for ~ 165 days relative to 148 and 138 days for cows with the GG and AG

genotypes, respectively. As with 4-226, cows with the 5-577 AA genotype had the fewest clinical cases of mastitis but also had the greatest number of days open. For the NRAMP1 SNP, cows heterozygous at this position tended to have fewer days open (138 ± 8) relative to cows homozygous for the C allele (155 ± 7).

Table 12. Tag SNP association with days open.

SNP	Genotype	LS Mean	SE	Group	P<	SNP r^2	Total r^2
4-226	AA	118	18	A	0.10	0.031	0.202
	GG	143	11	A			
	AG	165	14	A			
5-577	AA	165	24	A	0.6	0.010	0.176
	GG	148	11	A			
	AG	138	13	A			
NRAMP1	CC	154	11	A	0.2	0.014	0.179
	CG	135	12	A			

The projected 305 milk, fat, and protein yields also were evaluated for their association with each of the tag SNP. Of the five tag SNP, one in particular was associated with milk and protein yield, 4-62 (Tables 13 and 15). Those cows with an AA genotype had significantly greater ($P<0.05$) greater milk and protein yields than cows with an AC genotype, while those cows with a CC genotype were intermediate between the two. This SNP also explained ~ 2.7 – 3.4% of the variation in these two traits. However, this SNP was not associated with any other traits evaluated. The 5-577 SNP also tended to explain ~ 2.0 - 2.5% of the observed variation in milk and protein yield, as well as more than 1% of fat yield, but was not significantly associated with any of these traits. Similar observations were made with respect to SNP 4-226 and 5-691, in that greater than 1% of the

mean sample variation was explained by inclusion of these SNP in the models for fat/protein and fat respectively.

Table 13. Tag SNP association with milk yield.

SNP	Genotype	LS Mean	SE	Group	P<	SNP r^2	Total r^2
4-62	AA	23,281	1041	A	0.05	0.027	0.411
	CC	21,550	392	A			
	AC	20,731	488	A			
5-577	AA	21,264	890	A	0.30	0.025	0.403
	GG	21,072	414	A			
	AG	21,870	474	A			

Table 14. Tag SNP associated with fat yield.

SNP	Genotype	LS Mean	SE	Group	P<	SNP r^2	Total r^2
4-226	AA	854	32	A	0.5	0.018	0.379
	GG	834	21	A			
	AG	874	23	A			
5-577	AA	799	46	A	0.5	0.014	0.379
	GG	845	19	A			
	AG	874	23	A			
5-691	CC	881	23	A	0.5	0.017	0.374
	GG	835	28	A			
	CG	839	20	A			

Table 15. Tag SNP associated with protein yield.

SNP	Genotype	LS Mean	SE	Group	P<	SNP r^2	Total r^2
4-62	AA	699	30	A	0.025	0.034	0.386
	CC	646	11	AB			
	AC	618	14	A			
5-577	AA	628	26	A	0.40	0.020	0.371
	GG	633	12	A			
	AG	652	14	A			

4-226	AA	628	18	A	0.50	0.014	0.379
	GG	631	13	A			
	AG	655	14	A			

It is accordingly shown that the evaluated region of BTA2 is a relevant mastitis resistance locus, and potentially a reproductive trait locus. Of the five SNP evaluated, 4 were associated with clinical mastitis, subclinical mastitis, and/or somatic cell score – namely 4-226, 5-577, 5-691 (+777 in examples 1 and 2), and NRAMP1. Further work is contemplated to analyse haplotypes that result from the 5 identified tag SNP: (A) AGGCC, (B) CAGCC, (C) CGAGC, (D) CGGGC, and (E) CAGCG (order 4-62, 4-226, 5-577, 5-691, NRAMP1) which represent the majority of haplotypes observed. Of these haplotypes, A is the most distinct, while haplotypes B and E only differ at NRAMP1, and haplotypes C and D only differ at 5-577. Based upon this and phenotypic information, an informative tag SNP for mastitis and reproductive related traits was 4-226, 5-577, and NRAMP1.

Of these, 5-577 was the distinctive in that it introduces a stop codon and appears to influence disease resistance and reproductive efficiency. Lower levels of clinical mastitis and somatic cell scores were evident which would be expected if there is an impaired ability of neutrophils to reach the site of infection due to either altered expression or function of CXCR1. However there also is a substantial increase in days open (17-27 days) compared to cows with either a GG or AG genotype. The lower migration of neutrophils to the mammary gland with the AA genotype, suggests that neutrophil migration to the uterus also may be impaired and influence the development of metritis and other diseases post-calving which increase days open. Similar trends were observed with the 4-226

and NRAMP1 tag SNP, in that when high levels of clinical and subclinical mastitis were observed, days open was reduced and vice versa. Interestingly, SCS trends among genotypes for each tag SNP were similar to the trends for days open.

In conjunction with the individual allele models, models that included all tag SNP were evaluated for their association with mastitis, reproductive, and milk related traits. However, due to the relatively small sample size it was not possible to include all tag SNP in the final evaluations as it caused the adjusted means to differ from raw means. As this analysis was conducted in a similar time frame as the individual models, tag SNP and covariates were removed based upon their relative contribution to explaining model variation. In general, we observed that the largest imbalances between adjusted and raw means occurred when both 4-62 and 4-226 were both in the model due to the lack of animals in certain combinations. In particular, the AA genotype for 4-62 was present in only 5 cows and heterozygotes representing only 30% of cows. Therefore, we used 4-62 in place of 4-226 for models with milk, fat, and protein yields.

In general, the inclusion of at least three SNP in the model increased the level of known variation in the sample population relative to having only individual SNP. The degree of change was relative to the trait being selected (Table 9). With respect to clinical mastitis, the SNP accounted for ~ 3.25% of the variation in the sample mean. It is noted that when 5 SNP (4-226, 4-62, 5-577, 5-691, and NRAMP1) were included in the model for clinical mastitis ~ 10% of the sample variation was explainable. The greatest loss in known variation occurred when 4-62 was removed from the model and reduced the r2 value to ~ 5%.

Therefore, it would be beneficial to evaluate this SNP in combination with 4-226, 5-577, and NRAMP1 in a larger sample population.

In contrast, for subclinical mastitis the greatest reduction in the r2 value occurred when 5-691(+777) SNP was removed from the model leaving SNP 4-226, 5.577, and NRAMP1 (9.6% vs 5.5%). Days to first service also experienced a 2-3% absolute reduction in r2 values when 4-62 and 5-691 were removed from the models. A less dramatic difference was observed for days open (1%) when SNP 4-226, 5-577, and NRAMP1 were the only SNP in the model. The SCS was evaluated with a minimum combination of 5-577, NRAMP1, and 4-226 or 4-62 in the model. Of these two variable SNP (4-226 and 4-62), 4-62 was better at helping explain variation in the sample population as a loss of ~ 2% was observed when 4-226 was included in its place. Overall, the loss of known variation most likely is due to a loss in power to resolve cows with different genetic backgrounds since each of the 5 tag SNP would be necessary to best define genetic variation of the sample population. In the models for milk, fat, and protein, less overall variation was observed in the sample and allowed inclusion of 4 of SNP in the model – 4-62, 5-577, 5-691, and NRAMP1. However, it was not possible to include both 4-62 and 4-226 in the model as this lead to large shifts in raw versus adjusted means. Therefore, for milk traits, 4-62 was included due to its greater association with these traits than that demonstrated by 4-226. In comparison to when individual SNP analyses were run, a 4-5% increase in the absolute r2 value was observed for milk, fat, and protein. Overall, these data suggest that for traits such as clinical mastitis, subclinical mastitis, and days to first service the best model most likely will

include a combination of multiple SNP, which is reflective of the multiple factors that can influence these traits.

Table 16. Variation attributed to tag SNP in models with different tag SNP included.

Tag SNP included in model	Clin Mast[1]	Sub Mast	SCS	DO	DFS	Milk	Fat	Protein
	r2							
4-62, 4-226, 5-577, 5-691, NRAMP1	0.100	0.097	0.097	0.067	0.14	0.063	0.059	0.082
4-62, 5-577, 5-691, NRAMP1	ND[2]	ND	0.055	ND	ND	0.072	0.052	0.082
4-226, 5-577, 5-691, NRAMP1	0.051	0.096	0.096	0.060	0.123	ND	ND	ND
4-226, 5-577, NRAMP1	0.032	0.055	0.037	0.061	0.091	ND	ND	ND

[1] Clin mast = Clinical mastitis; Sub mast = Subclinical mastitis; SCS = somatic cell score; DO = Days open; DFS = days to first service
[2] ND = not determined

The ability to genetically identify cattle more susceptible to mastitis represents a key factor in helping to produce a healthier cattle population. Within the CXCR1 and NRAMP1 genes, we have now identified SNP that are associated with mastitis and other dairy related traits. With respect to clinical mastitis, the combination of 5 tag SNP that best captured the genetic variation in the tested region of BTA2 (4-226, 4-62, 5-577, 5-691, and NRAMP1), also best explained the variation in the sample population (~10%). In contrast, less genetic

information was required for subclinical mastitis, as the loss of 4-62 in the model did not impair the statistical relevance of the remaining 4 SNP. The presence of 4-62 also was better at explaining the sample variation in SCS than was the 4-226 tag SNP. The reduced model with only 4-226, 5-577, and NRAMP1, also were sufficient to explain 6-9% of the sample variation for both days open and days to first service, suggesting these SNP also may be relevant for determining reproductive efficiency. Similar observations were made with respect to milk, protein, and fat yields when all but 4-226 were included within the model.

The foregoing description of a preferred embodiment of the invention has been presented for purposes of illustration and description. It is not intended to be exhaustive or to limit the invention to the precise form disclosed. Obvious modifications or variations are possible in light of the above teachings. The embodiment was chosen and described to provide the best illustration of the principles of the invention and its practical application to thereby enable one of ordinary skill in the art to utilize the invention in various embodiments and with various modifications as are suited to the particular use contemplated. All such modifications and variations are within the scope of the invention as determined by the appended claims when interpreted in accordance with the breadth to which they are fairly, legally and equitably entitled.

What is claimed is:

1. A method for determining predisposition to an inflammatory disease, comprising:

obtaining a sample of nucleic acid from an animal predisposed to contracting the inflammatory disease, the sample comprising at least one of SEQ ID NO: 35, and SEQ ID NO: 36;

determining the identity of the nucleotide in at least one of positions 151970, 152084, 152164, or 152328 of SEQ ID NO: 35; and

determining the identity of the nucleotide in position 546 in SEQ ID NO: 36.

2. The method of claim 1, further including the step of detecting the presence of a single nucleotide polymorphism in SEQ ID NO: 35 in at least one of positions 151970, 152084, 152164, or 152328.

3. The method of claim 2, wherein the polymorphism detected at position 151970 is G → A, the polymorphism detected at position 152084 is C → G, the polymorphism detected at position 152164 is A → C, and the polymorphism detected at position 152328 is G → A.

4. The method of claim 1, further including the step of detecting the presence of a single nucleotide polymorphism in SEQ ID NO: 36 in position 546.

5. The method of claim 4, wherein the polymorphism detected at position 546 is C → G.

6. The method of claim 1, wherein the nucleic acid is genomic DNA.

7. The method of claim 1, wherein the inflammatory disease is an inflammation of the mammary gland.

8. The method of claim 7, wherein the inflammatory disease is mastitis.

9. A method for determining predisposition to an inflammatory disease, comprising:

obtaining a sample of nucleic acid from an animal predisposed to contracting the inflammatory disease; and

determining the identity of the nucleotide in at least one of positions 151970, 152084, 152164, or 152328 of SEQ ID NO: 35.

10. The method of claim 9, further including the step of detecting the presence of a single nucleotide polymorphism in SEQ ID NO: 35 in at least one of positions 151970, 152084, 152164, or 152328.

11. The method of claim 10, wherein the polymorphism detected at position 151970 is G → A, the polymorphism detected at position 152084 is C → G, the polymorphism detected at position 152164 is A → C, and the polymorphism detected at position 152328 is G → A.

12. A method for determining predisposition to an inflammatory disease, comprising:

obtaining a sample of nucleic acid from an animal predisposed to contracting the inflammatory disease; and

determining the identity of the nucleotide in position 546 of SEQ ID NO: 36.

13. The method of claim 12, further including the step of detecting the presence of a single nucleotide polymorphism in SEQ ID NO: 36 in position 546.

14. The method of claim 13, wherein the polymorphism detected at position 546 is C → G.

15. An isolated nucleic acid molecule comprising 400 contiguous nucleotides of SEQ ID NO: 35 or the complement thereof, including: (a) nucleotide 151970 wherein the G is replaced by an A; (b) nucleotide 152084 wherein the C is replaced by a G; (c) nucleotide 152164 wherein the A is replaced by a C, (d) nucleotide 152328 wherein the G is replaced by an A; or at least one of (a), (b), (c), and (d).

16. An isolated nucleic acid molecule comprising 30 contiguous nucleotides of SEQ ID NO: 36 or the complement thereof, including nucleotide 546 wherein the C is replaced by a G.

17. A probe selected from the group consisting of:

a probe that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO: 35 having A at position 151970 but not to a nucleic acid molecule consisting of SEQ ID NO: 35 having G at position 151970;

a probe that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO: 35 having G at position 152084 but not to a nucleic acid molecule consisting of SEQ ID NO: 35 having C at position 152084;

a probe that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO: 35 having C at position 152164 but not to a nucleic acid molecule consisting of SEQ ID NO: 35 having A at position 152164;

a probe that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO: 35 having A at position 152328 but not to a nucleic acid molecule consisting of SEQ ID NO: 35 having G at position 152328; and

a probe that hybridizes under high stringency conditions to a nucleic acid molecule comprising SEQ ID NO: 36 having G at position 546 but not to a nucleic acid molecule consisting of SEQ ID NO: 36 having C at position 546.

18. The probe of claim 17, wherein the probe is detectably labeled.

612

CTTCCGTGAGGCCTATCAACCACC**G/A**TACTCCGACCTAGTCTGCTACGAGGACCT

684

GGGTGCCAATACAACGAAATGGCGGATGATAATGCGTGTCCT**G/A**CCCCAGACCTT

TGGCTTCCTCCTGCCCCTGCTGGTCATGCTGTTCTGCTACGGATTCACCCTGCGC

777

ACGCTATTTTCAGCCCAAATGGGGCA**C/G**AAGCACCGGGCCATGCGGGTCATCTTT

858

GCTGTCGTGCTCGTCTTCCTGCTCTGCTGGCTGCCCTACAACCTGGTCCTGAT**C/A**

861

GC**G/A**GACACCCTCATGAGGGCCCATGTGATTGCTGAGACCT

Fig. 1

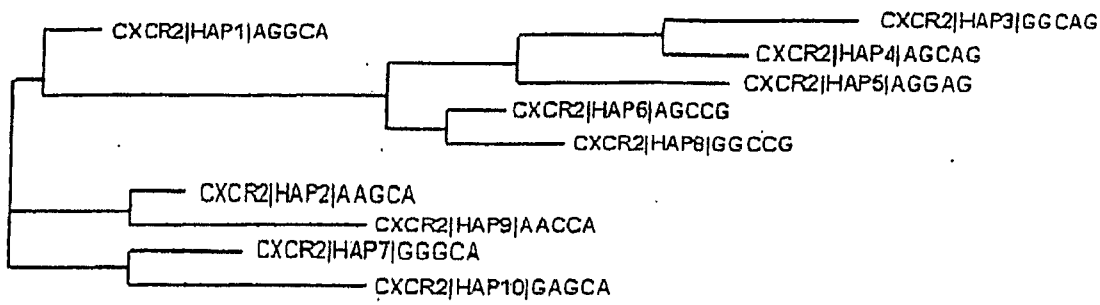

Fig. 2

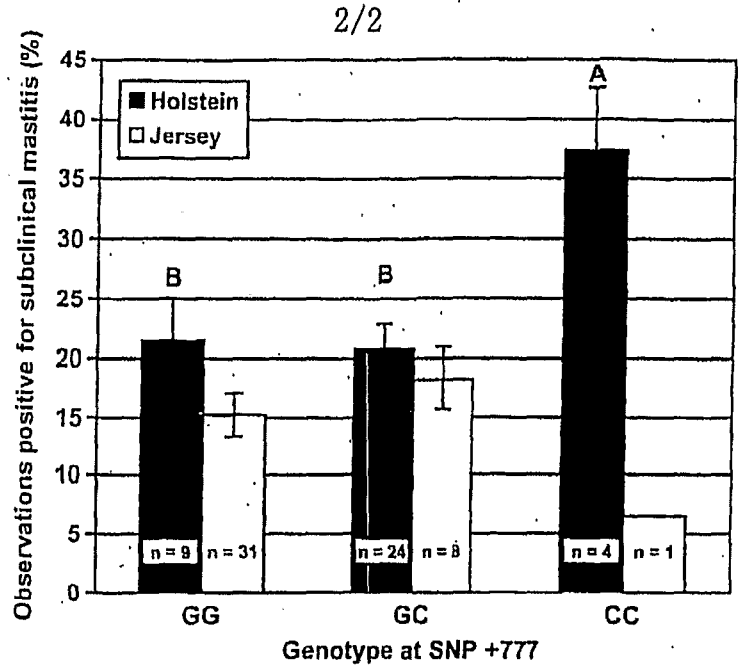
Observations positive for subclinical mastitis (%)
45
40
35
30
25
20
15
10
5
0
Holstein
Jersey
A
B
B
n = 9
n = 31
n = 24
n = 8
n = 4
n = 1
GG
GC
CC
Genotype at SNP +777

Fig. 3

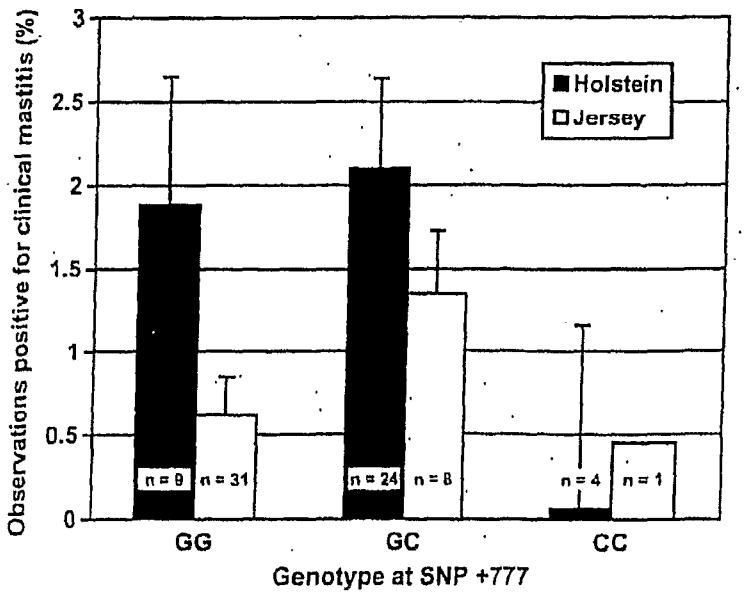
Observations positive for clinical mastitis (%)
3
2.5
2
1.5
1
0.5
0
Holstein
Jersey
n = 9
n = 31
n = 24
n = 8
n = 4
n = 1
GG
GC
CC
Genotype at SNP +777

Fig. 4