BOVINE *SLC11A1*: GENOMIC SEQUENCE VARIATION AND FUNCTIONAL ANALYSIS IN CATTLE NATURALLY RESISTANT AND SUSCEPTIBLE TO BOVINE BRUCELLOSIS

A Dissertation

by

CHRISTOPHER JOHN SCHUTTA

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2006

Major Subject: Veterinary Microbiology

BOVINE *SLC11A1*: GENOMIC SEQUENCE VARIATION AND FUNCTIONAL ANALYSIS IN CATTLE NATURALLY RESISTANT AND SUSCEPTIBLE TO BOVINE BRUCELLOSIS

A Dissertation

by

CHRISTOPHER JOHN SCHUTTA

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Joe W. Templeton
Committee Members,	L. Garry Adams
	Donald S. Davis
	James N. Derr
Head of Department,	Gerald R. Bratton

August 2006

Major Subject: Veterinary Microbiology

ABSTRACT

Bovine *SLC11A1*: Genomic Sequence Variation and Functional Analysis in Cattle Naturally Resistant and Susceptible to Bovine Brucellosis.

(August 2006)

Christopher John Schutta, B.S., Binghamton University,

State University of New York; M.S., Texas A&M University

Chair of Advisory Committee: Dr. J. W. Templeton

Previous analysis of the bovine *SLC11A1* complementary DNA (cDNA) failed to identify any nucleotide variations other than a microsatellite length variation within the 3' untranslated region functionally associated with bovine brucellosis. In this study I set out to identify mutations in the genomic complement of the gene that may be associated with resistance or susceptibility to bovine brucellosis, and to determine if the microsatellite length polymorphism in the 3'UTR of bovine *SLC11A1* modulates gene expression and subsequent disease resistance in a phase dependent manner. The results of this study demonstrate that there are seventy-five total single nucleotide polymorphic (SNP) sites (excluding indels) located within the bovine genomic *SLC11A1* sequence of a *Brucella abortus* resistant bull and a susceptible cow. Twenty of these SNPs segregated between resistant and susceptible populations, with 3 non-synonymous SNPs significantly associating with resistance or susceptibility to *B. abortus* infection. An A695G within exon 2 resulted in a histidine (resistant allele) to arginine (susceptible allele)

amino acid substitution and was in significant linkage disequilibrium with the previously described 3' untranslated region (UTR) microsatellite length variation associated with brucellosis resistance. A transcriptional element search in the 3' UTR revealed a ETS-domain PU.1 site, an IFN-γ activation site (GAS), an Interferon Consensus Sequence Binding Protein site (ICSBP) and several Initiation Response sites (Inr), suggesting a possible function for this region in regulation of the expression of *SLC11A1*. A mobility shift assay confirmed sequence-specific DNA-protein interaction within this region. A luciferase reporter assay indicated that the 3'UTR of *SLC11A1* could act as a downstream enhancer for expression. Macrophage killing assays with RAW264.7 cells expressing bovine *SLC11A1* demonstrated that the microsatellite repeat is functionally associated with the macrophage killing efficiency, but not in a phase-dependent manner, suggesting that these length polymorphisms do not affect the angular orientation between cooperatively binding transcription factors, and leaves the possibility that the 3'UTR microsatellites regulate SLC11A1 transcription through some alternate mechanism, possibly mRNA stability.

DEDICATION

Dedicated to the memory of

Lt. Colonel John Curatelli, USAF

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Joe Templeton, for his guidance throughout my graduate career, and for teaching me how to ask the right questions.

Thanks to my committee members, Dr. Garry Adams, Dr. Donald Davis, and Dr. James Derr for their advice, support and infinite patience during the course of this research.

Finally, I would like to thank Leigh Ann Clark and my parents, Edward and Carmella Schutta, for keeping the faith and believing in me.

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

BRUCELLOSIS

Brucellosis is a major worldwide bacterial zoonosis. It has remained a disease of global importance since its discovery by Bruce in 1887 (15, 23). The causative agent of brucellosis is a group of gram-negative facultative intracellular bacteria belonging to the genus *Brucella.* The members of this genus are subdivided into seven species categorized by antigenic variation and primary host: *B. melitensis* (sheep and goats), *B. abortus* (cattle), *B. canis* (dogs), *B. suis* (hogs), *B. ovis* (sheep), *B. neotomae* (rats) and *B. maris* (marine mammals) (51). The primary host preference exhibited by the individual *Brucella* species reflects a co-evolutionary balance between the pathogen's genome, which has evolved strategies enabling survival within the preferred host, and the primary host genome, which has evolved strategies of immune response which suppress pathogen expansion within the host biological system (2).

In most cases of *B. abortus* infection, exposure occurs through inhalation or ingestion of the organism. In the case of exposure in dogs, sheep, and swine, *B. abortus* is sexually transmitted. Like other intracellular pathogens, *brucellae* require four steps for successful infection: adherence, invasion, establishment, and dissemination (34). The intracellular niche of infecting *brucellae* are specialized

---

This dissertation follows the style and format of Infection and Immunity.

acidified phagosomes within the macrophages of the host, called brucellasomes (20, 67). The exact mechanism of bacterial entry into the host cell is not known (51), however, there is research that supports the existence of a *B. abortus* specific cell surface receptor on macrophages that mediates binding and internalization of the pathogen (17).

Once inside the phagosome, the internalized bacteria that survive the initial oxidative burst and the nutrient/oxygen deprived environment subvert the host cellular machinery through the interaction of the pathogen's Type IV secretion system (*virB*) with the host's endoplasmic reticulum (20). The outcome of this host-pathogen interaction is prevention of lysosomal fusion, neutralization of phagosomal pH, utilization of nitrate ions for anaerobic respiration, and regulated multiplication of the bacteria within the phagosome. The accumulated bacteria eventually escape the phagosome and are disseminated to other host cells. In female ruminant hosts, *B. abortus* interacts with the placental trophoblasts through the tissue tropism activity of erythritol produced by these cells. Within these cells, *brucellae* enter their acute replicative stage, with placental disruption resulting in fetal loss or the birth of weak and infected offspring (27). In humans, *B. abortus* spreads to the lymph nodes, spleen, liver and bone marrow. The pathological manifestations of untreated infection may be meningitis, endocarditis, spondylitis, and arthritis (51). The clinical presentation of human brucellosis can include fever, malaise, arthralgias, hepatomegaly, splenomegaly, lymphadenopathy, peripheral arthritis, sacroiliitis, epididymoorchitis, vomiting and diarrhea (64).

Worldwide brucellosis continues to be an economic and public health concern. Geographical regions listed as high risk are the Mediterranean Basin, South and Central America, Eastern Europe, Asia, Africa, the Caribbean, and the Middle East (80). Of even greater concern is the fact that brucellosis is considered a re-emerging problem for countries such as Israel, Kuwait, Saudi Arabia, Brazil and Colombia (23). World Health Organization data documents approximately 500,000 human cases of brucellosis per year (80). The absence of a human vaccine means that along with pasteurization and proper sanitation, the prevention of human infection requires control of the disease in the animal hosts. The epidemic potential of *Brucella* species, the efficiency of aerosol infection, and the lack of an efficient human vaccine also makes this airborne pathogen a potential agent of bioterrorism (51). In fact successful attempts at weaponizing *B. suis* occurred as early as 1954 (36). In 1999, a suspected case of human brucellosis with an atypical clinical presentation prompted an investigation of possible bioterrorism in New Hampshire and Massachusetts (21). Because of this, members of the genus *Brucella* are classified as biosafety level 3 pathogens and have been placed on the Select Agent list of Bioterrorism as part of the 2001 Patriot Act (84).

NATURAL DISEASE RESISTANCE AND THE INNATE IMMUNE RESPONSE

The ability of a naive host to resist primary infection by a pathogen is known as natural disease resistance. F. B. Hutt specifically defined natural disease

resistance as “the inherent capacity of an animal to resist disease when exposed to pathogens, without prior exposure or immunization, of which the major component is heritable and, therefore, stably passed from parent to offspring” (47). Natural disease resistance involves the nonspecific immune response in vertebrates. Nonspecific immunity includes pattern-recognition receptors (PRRs) that bind to ancient and conserved bacterial pathogen-associated molecular patterns (PAMPs) as well as single-stranded viral ribonucleic acid (ssRNA) (59). Toll-like receptors (TLRs) (79), nucleotide-binding oligomerization domain proteins (NODs) (22), lipopolysaccharide binding protein (LBP) (88), and mannose-binding protein (75) are eukaryotic receptors that all recognize PAMPs. Binding of these PRRs with their target PAMPs results in opsinization or phagocytosis of the bacteria or virus, and initiates the transcription of nonspecific immune response genes. The products of these genes include antimicrobial peptides (35) and inflammatory cytokines (9). This innate immune response, which occurs immediately after infection, serves to sequester the pathogen, limit its replication, and in some instances, clear the pathogen altogether. It is only when the innate immune response is bypassed, evaded or overwhelmed that a specific or adaptive immune response is required.

Phylogenetically the non-specific innate immune response predates the specific adaptive immune response. Innate immunity is found in some form or another in all multicellular organisms, while the adaptive immune response, appearing around 400 million years ago, is only found in vertebrates (82). It was believed that because the innate response is more primitive than the adaptive

immune response, the innate response was also less advanced and therefore not as important as the adaptive response. The innate immune response was thought to function merely as a stopgap measure during primary infection until the adaptive immune response could be activated. Additionally, because the innate immune response does not lead to immunological memory, it was considered less critical to the overall immunity in higher organisms than was the adaptive immune response. However, there has been a paradigm shift in the scientific community with regards to the importance of the innate immune response (25), and the prevailing thought now is innate immunity not only functions as the first line of defense against an invading pathogen, but it also plays an instructive role in coordinating an effective adaptive immune response for short and long-term defense (31).

Natural disease resistance and the nonspecific immune response both play a significant role in the outcome of *Brucella* infection in cattle. Studies to determine the genetic basis for resistance to *B. abortus* in cattle indicate that natural resistance to brucellosis in cattle is a complex phenotypic trait determined by two or more interacting genes resulting in complex genetic types (3). Classical breeding studies involving previously unvaccinated and unexposed mature bulls and heifers phenotyped resistant or susceptible by *in vivo* challenge with *B. abortus* 2308 demonstrated that the population frequency of natural resistance to *Brucella* could be significantly increased with one round of selective breeding (81). In these experiments the frequency of natural resistance was increased from 20% to 58.6% in one generation when mating a resistant bull to a resistant heifer. Analysis of

immune correlates related to natural resistance to *Brucella* revealed a differential response in macrophage activation between resistant and susceptible cattle, along with the associated segregation of specific alleles of bovine solute carrier gene 11A1 (SLC11A1), a gene involved in macrophage activation (3).

THE SOLUTE CARRIER GENE SUPERFAMILY AND SOLUTE CARRIER 11A1

The solute carrier gene superfamily (SLC) encodes a group of integral membrane proteins that includes passive transporters, ion coupled transporters, and exchangers that traffic crucial compounds into and out of cells and organelles. Currently the human SLC superfamily consists of 43 gene families encoding 298 individual transporter proteins. (41). A number of human disease syndromes are linked to SLC transporter gene defects, including hypertension (40), deafness (61), and hemochromatosis (58), and these integral membrane transporters proteins are prime candidates to be exploited as drug delivery systems or drug targets in the treatment of disease.

Solute carrier family number eleven (SLC11, formerly known as the NRAMP family) is made up of two member genes that encode proton coupled metal ion transporters, *SLC11A1* (formerly known as *NRAMP1*) and *SLC11A2* (formerly known as *DMT1*). Originally characterized in mice, *SLC11A1* mapped to the mouse *Bcg*/*Lsh*/*Ity* locus, a region conferring resistance to infection by the live attenuated Bacillus Calmette-Guerin (BCG) strain of *Mycobacterium bovis*,

Leishmania donovani, and *Salmonella typhimurium* in the murine host (85). *SLC11A1* is a highly conserved gene, with orthologs found in both prokaryotes and eukaryotes, while homologs have been identified in mycobacteria (4), yeast (54), *Drosophila* (68), chickens (45), swine (83), dogs (7), horse (43), deer (57), cattle (33), and humans (49). *SLC11A1* expression occurs mainly in macrophages and is interferon-γ-inducible. *SLC11A1* expression in macrophages is also modulated by exposure to bacterial lipopolysaccharide (LPS) (8). The gene product is a polytopic integral membrane protein made up of 10-12 transmembrane domains. It is a 548 amino acid protein containing three putative phosphorylation sites, two SH3 binding motifs, and a single exofacial glycosylation site (33). Bovine *SLC11A1* also contains a highly conserved binding-protein-dependent transport system inner membrane component signature that alludes to the putative divalent cation transporter function of SLC11A1. Localization studies demonstrate that the protein product is recruited to the late endocytic compartment in a Lamp1 (lysosomal-associated membrane protein-1)-positive compartment and remains associated with this compartment as it matures to a phagolysosome (37), supporting the theory that SLC11A1 restricts replication of infecting pathogens by altering the intracellular environment. Research into the function of SLC11A1 suggests that the protein may modulate phagosomal pH. (38), as well as alter divalent cation concentrations, including iron, within the phagosome (11, 89).

The disease resistance association of *SLC11A1* inferred by the mapping of *SLC11A1* to the *Bcg/Lsh/Ity* locus was confirmed in experiments in which

resistance to intracellular pathogens was abrogated in *SLC11A1* knockout mice (86). In addition, the phosphoglycoprotein gene product of *SLC11A1* was determined to be absent in macrophages from *Bcg* susceptible mouse strains, with anti-SLC11A1 antiserum failing to detect mature SLC11A1 protein in macrophages isolated from C57BL6/J and BALB/c Bcg^s mice (87). A single, non-conserved glycine-to-aspartic acid amino acid substitution at position 169 results in a nonfunctional gene product and susceptibility of murine macrophages to *BCG*. In contrast to macrophages from $Nramp1^{G169}$ mice, in which a 100-kDa mature Nramp1 protein is detected, Nramp1 specific antiserum fails to detect an immunoreactive protein in macrophage lysates from susceptible mouse strains bearing the $Nramp1^{D169}$ allele (87). This disrupting amino acid substitution is not found in SLC11A1 proteins sequenced from other species, including deer, bison, cattle, dogs, and humans. However, DNA sequence analysis and disease association studies in humans have identified a functional microsatellite polymorphism in the human *SLC11A1*. In the human *SLC11A1* gene, polymorphisms in a GT dinucleotide microsatellite in the 5' promoter of the gene correlate with both autoimmunity and infectious disease susceptibility (74). A similar microsatellite length polymorphism is found in the 3' untranslated region of the bovine *SLC11A1*gene (32). Single Stranded Conformation Analysis (SSCA) indicates that length polymorphisms within this region of the gene are associated with disease resistance to *B. abortus* in cattle. *In-vivo* challenge studies revealed that microsatellite lengths of 13 GT dinucleotide repeats were found in cattle

resistant to challenge by *B. abortus*, while microsatellite lengths of 14, 15, and 16 GT dinucleotide repeats were found in cattle susceptible to *B. abortus* challenge. *In-vitro* assays demonstrated that macrophages from animals with the resistant phenotype had a greater ability to kill *B. abortus* than did those from animals with the susceptible phenotype (65). When an expression vector carrying the bovine *SLC11A1* resistant allele (GT13) was transfected into a susceptible mouse macrophage cell line, the phenotype switched from susceptible to resistant in an *in-vitro* macrophage-killing assay with *B. abortus* challenge (10). Alternately, when the expression vector carrying a susceptible allele from bovine *SLC11A1* (GT16) was transfected, there was no change in the susceptible phenotype. The same study showed that there was a greater level of *SLC11A1* mRNA present in transfected cells carrying the resistant allele than in the transfected cells carrying the susceptible allele. Population Analyses in Holstein and Zebu cattle breeds demonstrated breed specific differences in the allelic frequencies of these microsatellite length polymorphisms (63).

NUCLEOTIDE REPEAT INSTABILITY AND DISEASE ASSOCIATION

In humans, instability in DNA dinucleotide and trinucleotide repeats is associated with several neurodegenerative and tumorigenic diseases. Genome wide instability of dinucleotide repeats is observed in tumors isolated from patients suffering from Hodgkin's disease and hereditary nonpolyposis colon cancer (1, 55). The archetypal triplet repeat disease, Huntington's disease, is categorized as a

translated polyglutamine (polyQ) repeat disease, with its functionally associated unstable trinucleotide repeat (CAG), located in the first exon of the huntingtin (Htt) gene, laying down an expanded tract of glutamines within the expressed sequence of the gene (46). Repeat lengths of eight to 39 are found in normal individuals, while repeat lengths of 37 or more are found in affected individuals (70). It is thought that the expressed polyQ tract results in a toxic gain of function of the mutant protein, such as formation of neuronal intranuclear inclusions, interference with cytoskeletal and vesicular transport, and impaired gene expression (30). Other polyQ diseases include Kennedy's disease, Spinobulbar Muscular Atrophy, and Machado-Joseph disease. These diseases all exhibit genetic anticipation, in which subsequent generations have earlier onset and increased severity of disease due to enlargement of the expansion over generations.

Microsatellite repeat diseases are not restricted solely to translated expansions within mutant protein. Repeat diseases with untranslated repeat expansions found in the 5' and 3' untranslated regions of the gene, as well as within the introns of the gene, have also been identified. Friedreich's ataxia (FRDA) is one such disease, containing an expanded GAA trinucleotide repeat in the first intron of the frataxin gene and inhibiting expression of the gene (18). The mechanism of this altered gene regulation is unknown at this time. One possibility is that changes in polymorphic microsatellite repeat can hinder binding of transcriptional factors through the localized formation of tightly packed heterochromatin or through

disruption of synergistic protein-protein interactions and cooperative binding of the proteins to the DNA.

PROTEIN-PROTEIN INTERACTIONS AND COOPERATIVE DNA BINDING OF REGULATORY FACTORS

Cooperative binding is a phenomenon whereby protein-protein interactions between two or more DNA binding proteins reduce the free energy of binding to DNA so that together proteins bind more tightly to DNA than each factor alone. This is a key mechanism for generating specificity within multicomponent nucleoprotein complexes (62).

DNA bending facilitates the protein-protein interactions between factors whose binding elements are separated by long distances across the DNA molecule (26). Over the persistence length of DNA, about 140 base pairs, DNA bending extracts an energetic penalty that must be paid by the strength of the protein-protein interaction or by a protein that stabilizes the bend. Studies on cooperatively binding lambda repressor proteins have shown that changes in the angular orientation of the protein-binding site along the DNA alpha helix affects the strength of this protein-protein interaction (42). Normally the two binding sites for lambda repressor have a center-to-center distance of 25 bp or 2.4 helical turns. Insertion of 10 or 11 bp (1 helical increment) has no effect on cooperativity up to a distance of 80 bases or 8 turns of the α-helix. However, the insertion of non-helical increments abolishes

cooperativity. This is due to the energetic penalty of twisting the DNA in order to place the proteins into a favorable angular orientation to interact cooperatively.

RATIONALE AND SPECIFIC AIMS

Nucleotide variations found in bovine *SLC11A1* cDNA do not appear to be associated with natural disease resistance to *B. abortus* in cattle. However, functionally associated nucleotide variants may be overlooked in under-represented alleles due to reduced expression levels of these alleles. This is particularly important since it has been reported that individuals carrying the susceptible allele for *SLC11A1* have reduced mRNA expression in both humans and cattle (32, 74). Screening of the genomic complement of DNA for the sequence variation should provide a more accurate accounting of allelic variation in each individual since the DNA template would not be affected by allelic differences in expression.

To date, the only bovine *SLC11A1* polymorphisms identified to be significantly associated with *B. abortus* natural disease resistance in cattle is the dinucleotide microsatellite length polymorphism found in the 3' untranslated region of the gene. While it is suggested from previous studies that variable microsatellite lengths affect macrophage bactericidal activity, the mechanism of this effect is not known. It is possible that changes in the 3'UTR microsatellite length may alter the angular orientation of two or more flanking transcription regulation sites specific for cooperatively binding transcription factors and reduce the efficiency of the synergistic binding of these proteins to the DNA molecule. In this case,

polymorphisms in microsatellite length would alter transcription in a helical-phase dependent manner.

The research described in this dissertation had two specific aims. The first aim was to determine the complete genomic sequence for the bovine *SLC11A1* gene and use this sequencing data to implement a search for informative SNPs that segregate with disease resistance to *B. abortus* in resistant and susceptible cattle populations. The second aim was to determine if the microsatellite length polymorphism in the 3'UTR of bovine *SLC11A1* modulates gene expression and subsequent disease resistance in a phase dependent manner.

CHAPTER II

GENETIC VARIATION IN THE GENOMIC DNA SEQUENCE OF BOVINE *SLC11A1* IN CATTLE NATURALLY RESISTANT AND SUSCEPTIBLE TO BOVINE BRUCELLOSIS

INTRODUCTION

Nucleotide variations in *SLC11A1* have long been known to be associated with host natural disease resistance to bacterial infection. In mice a guanine (G) to adenine (A) transition results in a glycine to aspartic acid amino acid substitution at position 169 in the SLC11A1 protein and abrogates resistance to infection by *Mycobacterium bovis BCG*, *Leishmania donovani*, and *Salmonella typhimurium* (85). In chickens, a G to A transition at nucleotide position 696 of expressed *Nramp1* sequence was identified in *Salmonella* susceptible strains of chickens and not observed in resistant strains (44). Variation in a GC-rich region of canine *SLC11A1* and complete deletion of exon 11 was detected in dogs susceptible to *Leishmania infantum* (7), while several polymorphisms in human *SLC11A1*, including a (GT)n microsatellite length variation in the 5' promoter, a single nucleotide substitution in intron 4, a non-conserved single base change at codon 543 resulting in an apartic acid to asparagine amino acid substitution, and a four base deletion in the 3' UTR, were associated with tuberculosis, sarcoidosis, and rheumatoid arthiritis (12, 28, 48, 71, 72).

In cattle, sequencing of *SLC11A1* cDNA isolated from animals phenotyped for resistance and susceptibility to bovine brucellosis failed to identify any

nucleotide variations in the expressed regions of the gene associated with natural disease resistance *Brucella abortus* (33). The only associated polymorphism identified using cDNA as a template, was a polymorphic (GT)n microsatellite located in the 3' untranslated region of the gene. However, since cDNA isolation is dependent upon gene expression levels and the concentration of gene specific mRNA, it is possible that polymorphisms occurring in low expressing alleles could be overlooked. Stable transfection studies in the mouse macrophage cell line RAW264.7 demonstrated that *SLC11A1* mRNA is nearly undetectable in cells expressing the susceptible bovine *SLC11A1* allele, while *SLC11A1* mRNA is readily detected in cells expressing the resistant allele (10). A search for nucleotide polymorphisms solely in expressed sequences also fails to detect functional polymorphisms within the unexpressed introns that may affect gene regulation. Using genomic DNA as the template in a screen for SNPs provides a more complete accounting of any putative disease associated variants in the gene sequence.

In this study, in order to determine if there are any polymorphisms associated with natural disease resistance to *B. abortus* that may have been overlooked using cDNA as a template, a search for associated nucleotide variants within the exons of bovine *SLC11A1* was performed using genomic DNA isolated from resistant and susceptible animals. To facilitate this search, the complete genomic sequence was determined for bovine *SLC11A1* isolated from a bacterial artificial chromosome (BAC) library, and this sequence information was used to

develop a strategy for amplifying and sequencing across the intron-exon junctions and through the complete exons of genomic *SLC11A1* from phenotyped cattle.

MATERIALS AND METHODS

BAC Clone Sequencing. A bovine genomic BAC library was kindly provided by Dr. Scott Davis, Department of Animal Science, Texas A&M University (16). A PCR based screening strategy was previously employed to isolate clones containing the complete bovine *SLC11A1* gene (32). The primer pair 1F and 1R from a murine *SLC11A1* genomic sequence (5'-tctctggctgaaggctctcc-3' and 5'-ccaagctcaccttagggtag-3' respectively) was used to screen the BAC clone library. Single-plate DNA pools were created by combining the 100 ml overnight cell cultures of 96 BAC clones. Two separate clones positive for SLC11A1 were isolated and the bovine DNA insert was removed from the pBeloBACII vector by NotI restriction enzyme digestion. The digested DNA was run on a 1% agarose gel using a CHEF apparatus (Bio-RAD).

PCR amplification products from the BAC clone insert were generated in a 50μL reaction mix containing 1X PCR Buffer, 2.5mM $MgCl_2$, 400μM dNTPs, 0.2μM primers, and 2.5U LA Taq Polymerase (Panvera). PCR conditions consisted of an initial denaturation step of 94ºC for 1min., followed by 30 cycles of 98ºC for 20sec and 68ºC for 2min. One final extension step of 72ºC for 10 min. was added after the last cycle to ensure complete extension of any partial product. Sequencing was accomplished by primer walking. Each sequencing reaction included 3μL Big

Dye (Applied Biosystems), 3μL of 6.33mM $MgCl_2$, 1μL DMSO, 40ng of PCR product DNA and 50μM primer. Cycle sequencing conditions consisted of an initial denaturation step of 96ºC for 5min., followed by 47 cycles of 95ºC for 30sec., 55ºC for 20sec., and 60ºC for 4min. Sequence alignments and analysis were performed using Macvector v.8.0 software (Accelrys).

Confirmation of Naturally Occurring Bovine SLC11A1 Genomic Sequence. Three cattle of interest were selected for complete sequencing of the naturally occurring SLC11A1 gene; a prototypical *B. abortus* resistant bull designated Bull 86, a prototypical *B. abortus* susceptible cow designated Cow 292, and the clone of Bull 86 designated 86^2. Genomic DNA was isolated from peripheral blood monocytes (PBMC) or fibroblasts from each animal by phenol-chloroform extraction and standardized to a concentration of 100ng/μL in TE buffer. Bovine genomic *SLC11A1* intron and exon specific primers were designed using the *SLC11A1* sequence isolated from the BAC clone. PCR products were generated in a 20μL reaction containing 1X PCR Buffer, 2.5mM $MgCl_2$, 50μM dNTPs, 8.5μM primers, and 1U Taq Polymerase (Sigma). Cycling parameters were 35 cycles of 94ºC for 15sec., 58ºC for 15sec., and 70ºC for 20 sec. Cycle sequencing of amplified products was conducted in a reaction mix containing 3μL Big Dye (Applied Biosystems), 3μL of 6.33mM $MgCl_2$, 1μL DMSO, 40ng of PCR product DNA and 50μM primer. Cycling conditions were 96ºC for 10 sec., 50ºC for 5 sec., and 60ºC for 4min. for 26 cycles. Overlapping contigs were constructed using AssmblyLIGN v1.0.9 (Oxford Molecular). Sequence alignment and analysis were

performed using MacVector v8.0 (Acclerys) and DNAsp v4.0 (69). SNP analysis was performed using Seqscape v2.0 (Applied Biosystems Incorporated).

Genomic DNA for Single Nucleotide Polymorphism (SNP) Analysis. DNA from 18 genetically unrelated cattle originating from certified brucellosis free herds in 3 different U.S. states (Montana, Kansas, and Texas) were isolated by phenol-chloroform extraction. The test animals included one purebred Angus (*Bos taurus*) bull and 17 crossbred cows resulting from a three-way cross – F1 [Jersey (*Bos taurus*) X Hereford (*Bos Taurus*)] X American Brahman (*Bos indicus*) (39). All test animals were ELISA and card tested monthly to verify continued exposure-free status to *B. abortus* prior to experimental challenge. Natural resistant and susceptible phenotypes were based on *in vivo* bilateral conjunctival challenge with $1X10^7$ CFU of virulent *B. abortus* strain 2308. Cows were challenged at mid-gestation (180 ± 30 days), the time of maximal susceptibility to infection with *B. abortus*. The bull and cows phenotyped resistant (R) were all *B. abortus* culture and serologically negative, with R cows also carrying full-term, live, *B. abortus* culture negative calves and the bull maintaining *B. abortus* culture negative semen post challenge. Phenotyped susceptible (S) cows failed to carry to term and none of the fetuses were live births. Susceptible cows and their aborted fetuses were also serologically and culture positive. All test animals were housed in USDA-approved facilities with daily supplemental feeding. All experimental protocols were University Animal Use Committee reviewed and approved. DNA from a clone generated by nuclear transfer from the donor purebred Angus bull was included in

the sample group (52). The disease resistance phenotype for the cloned bull has yet to be determined by experimental challenge. However, data from *in vitro* macrophage killing assay indicate that the bactericidal activity of the macrophages from the cloned bull is comparable to the donor bull (bull 86), and the cloned bull has been categorized as *B. abortus* resistant (M. Westhusin, submitted for publication).

SNP Analysis. PCR primers positioned within the *SLC11A1* introns were designed using the complete genomic sequence of *SLC11A1* isolated from the BAC clone. Products were generated in a 20μL reaction containing 1X PCR Buffer, 2.5mM $MgCl_2$, 50μM dNTPs, 8.5μM primers, and 1U Taq Polymerase (Sigma). Cycling parameters were 35 cycles of 94ºC for 15sec., 58ºC for 15sec., and 70ºC for 20 sec. Cycle sequencing of amplified products was conducted in a reaction mix containing 3μL Big Dye (Applied Biosystems), 3μL of 6.33mM $MgCl_2$, 1μL DMSO, 40ng of PCR product DNA and 50μM primer. Cycling conditions were 96ºC for 10 sec., 50ºC for 5 sec., and 60ºC for 4min. for 26 cycles. Sequences were scanned for SNP differences using Seqscape v2.0 (Applied Biosystems Incorporated). To facilitate population analyses, individual sequences were concatenated and aligned using Clustal W. Haplotype reconstruction was accomplished using the Bayesian Algorithm model found in the Phase v.2.1.1 software package (76, 77) and population analyses were performed using DNAsp 4.0 (69).

RESULTS

Initial PCR and sequencing primers specific for the *SLC11A1* insert isolated from a bovine genomic BAC library were designed from the published bovine *SLC11A1* cDNA sequence (GENEBANK Accession #U12862) and from homologous conserved regions of mouse (GENEBANK Accession #NM013612) and human (GENEBANK Accession #D50403) *SLC11A1* cDNA (Table 1). The complete genomic sequence consisted of 10,816 nucleotide bases from the start codon to the end of the poly-A tail signal (AAATTAA). A number of important sequence motifs previously identified in the cDNA sequence of bovine SLC11A1 (33) were confirmed in the BAC sequence. Two predicted SH3 binding motifs (CCGGCACGCCAGCCA and CCACCCAGTCCAGAGCCA) were located at nucleotide positions 26 through 40 of the genomic sequence within the first exon and at nucleotide positions 675 through 692 within exon 2 of the genomic sequence respectively. Three predicted phosphorylation sites of consensus sequence C(S/TXR/K) were found at nucleotide positions 722, 1794, and 5620, within exons 2, 3 and 9 respectively, and a single exofacial glycosylation site (AACCTGACC) was located within the 10th exon from nucleotide position 7163 through 7171. In addition, a consensus signal transport motif was found within exon 11 at nucleotide position 7408. A Pustell Matrix comparison between the genomic sequence and the

TABLE 1. BAC specific primer pairs

Primer	Sequence	Position	Ta
NRAMP1_AF	AGCCACTCGCACAGAGAG	36-53	
NRAMP1_AR	GGTCAGCCAGAGGAGAATG	3479-3461	57.1
NRAMP1_AF1	GATGGGAATGAAAGGCTCTTATTTG	479-503	
NRAMP1_AR1	GGCTACAGACCCTGAATTATCATTC	3055-3031	58.9
NRAMP1_AF2	TGCTTCTGTCCTCAAGTGCTTAGCTG	911-936	
NRAMP1_AR2	TCAGCCTTATCTATCACGTGGACT	2477-2454	59.1
NRAMP1_3F	CAAGAATACTGGAATGGATAGCCAT	1214-1239	
NRAMP1_3R	TCTCCACCCACTTAGCCTGGTCACCA	1970-1945	59.6
NRAMP1_5F	ACGTGAAACCACCATGGCACCATGAA	3298-3323	
NRAMP1_5R	AGTCAGAGATGACTGCAGGATGAAGT	3669-3644	58.8
NRAMP1_NCFA	AGCTATTGCATTCAGTCTGCTCT	3527-3549	
NRAMP1_NCRA	CTTCAATCAGGAAGTACATGTTGG	5684-5661	58.2
NRAMP1_NCFA1	ATCTACAACACGCAGAGACTAGCA	4198-4221	
NRAMP1_NCRA1	TCCTCAATTCAGAGGCAGGAGTC	5112-5090	58.9
NRAMP1_BF1	ACAGGGGAGAGGACCTTAGCTAT	4534-4556	
NRAMP1_BR1	TGGAGTGGGGGAGAGGGACTGGT	4642-4620	58.2
NRAMP1_DF	CTTTGGGCAAGCCTTCTACA	5737-5756	
NRAMP1_DR	ATGCTGGTGAAGGTGAGGAT	9549-9530	57.4
NRAMP1_DF2	CGAAATATGTCGGGCTGGATTCAGA	6995-7019	
NRAMP1_DRS2	TACCAGAACCTTGTAGGTAAAGTGT	8164-8140	58.7
NRAMP1_11F	ACATTTACCAAGGAGTAAGCAAG	7199-7221	
NRAMP1_11R	TCATCTGTATAGTAGGGCTTATA	7593-7571	57.0
NRAMP1_12F	CGTGATCCTGGGCTGCCTCTTTGGT	7356-7380	
NRAMP1_12R	ATCACAAACTGTCCCGCGTAGGTGC	7469-7445	60.2
NRAMP1_EF	CCTGTCAGGCCTCAACGA	7987-8004	
NRAMP1_ER	ACTGTGCCCACTCATTCCA	10418-10400	57.6
NRAMP1_EF1	TGGACTACAGTTTATGGGGTCACAA	8415-8439	
NRAMP1_ER1	TGATAACTTTGCTCACCCTTGGGAA	9989-9965	58.5
NRAMP1_ENDF	CTACAGTTTATGGGGTCACAAAGAG	8419-8443	
NRAMP1_ENDR	ACTAGTCATATGGATTCGGGGACTA	10930-10906	57.9
NRAMP1_ENDF1	TTGGAAGGATTATTGCTGAAGCTGA	8988-9012	
NRAMP1_ENDR1	ATCCTCTTCAGGAAGCCCATACAG	10347-10324	59.7
NRAMP1_ENDF2	ATGCAGGAGTTTGCCAATGGCCT	9560-9582	
NRAMP1_ENDR2	CTCAGGATTCCAGTCCTGGCT	9756-9736	55.5
NRAMP1_ENDFA	TACCTGGGCCTCACCACTTA	10103-10122	
NRAMP1_ENDRA	ACAGGTGCCCTCTGGAACT	10530-10512	58.6
NRAMP1_ENDFB	GAATGAGTGGGCACAGTGG	10402-10420	
NRAMP1_ENDRB	GGTGGGTCTGTCAGGTCACT	10829-10810	55.4

cDNA sequence of bovine *SLC11A1* (Fig. 1) displays the 15 exons that make up the organization of this gene and is in agreement with the GT-AG donor-acceptor splice rule for intron-exon junctions (Table 2). Comparison of the bovine genomic *SLC11A1* sequence with human, mouse, and canine genomic *SLC11A1* indicated a high degree of sequence homology. The bovine SLC11A1 isolated from the BAC library has a 60.0% conserved nucleotide identity with human *SLC11A1* (GENEBANK Accession #AF229163), 50.5% with canine *SLC11A1* (GENEBANK Accession #AF091049) and 40.3% with murine *SLC11A1* (GENEBANK Accession #AC098570) genomic sequences. A Neighbor-Joining tree confirms that the bovine *SLC11A1* genomic sequence closely aligns with the human *SLC11A1* sequence (Fig. 2).

The bovine *SLC11A1* genomic sequence from the BAC library was compared to the 6-fold coverage whole genome shotgun (WGS) assembly of the completed bovine genome sequence provided by The Human Genome Sequencing Center at Baylor College of Medicine (www.hgsc.bcm.tmc.edu/projects/bovine) using BLAST (Fig 3). It aligned with the sequence fragment gn1/Btaurus 2.0/Chr2.147 of the bovine genome assembly from nucleotide position 274547 to 284428. This region of the assembly was referenced against the Bovine Genome Database provided by the Bovine Genome Database Consortium in support of the Bovine Genome Sequencing Project (www.BovineGenome.org). The aligned region of the genome coincided with 3 contigs of the BOMC EST cluster, 5474_CL1Contig1, 5474_CL4Contig1, and 5474_CL1Contig2, and corresponded

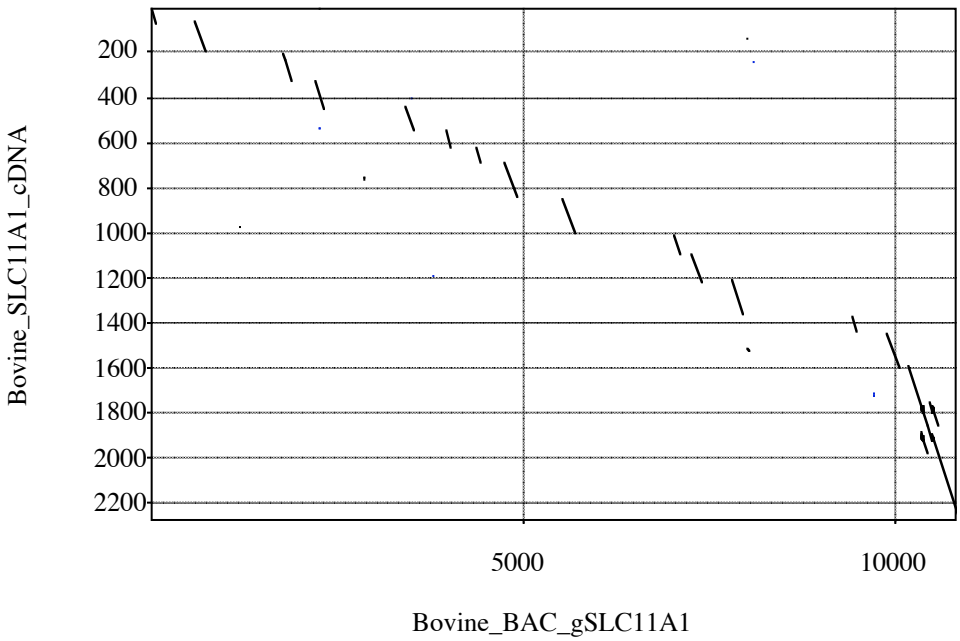

FIG 1. Pustell matrix analysis. The diagonal lines indicate the one-to-one alignment of the 15 exon sequences found in both the genomic and cDNA sequences. The gaps indicate the intronic sequences found within the genomic sequence but spliced out of the mRNA-derived cDNA sequence.

TABLE 2. Bovine *SLC11A1* intron/exon boundaries

Intron/Exon#	5' end	3'end
Exon 1	1	73
Intron 1	74	595
Exon 2	596	749
Intron 2	750	1784
Exon 3	1785	1908
Intron 3	1909	2218
Exon 4	2219	2347
Intron 4	2348	3445
Exon 5	3446	3552
Intron 5	3553	3975
Exon 6	3976	4046
Intron 6	4047	4386
Exon 7	4387	4454
Intron 7	4455	4771
Exon 8	4772	4927
Intron 8	4928	5540
Exon 9	5541	5699
Intron 9	5700	7045
Exon 10	7046	7135
Intron 10	7136	7276
Exon 11	7277	7396
Intron 11	7397	7805
Exon 12	7806	7955
Intron 12	7956	9431
Exon 13	9432	9505
Intron 13	9506	9895
Exon 14	9896	10049
Intron 14	10050	10183
Exon 15	10184	10297
3' UTR	10298	10816

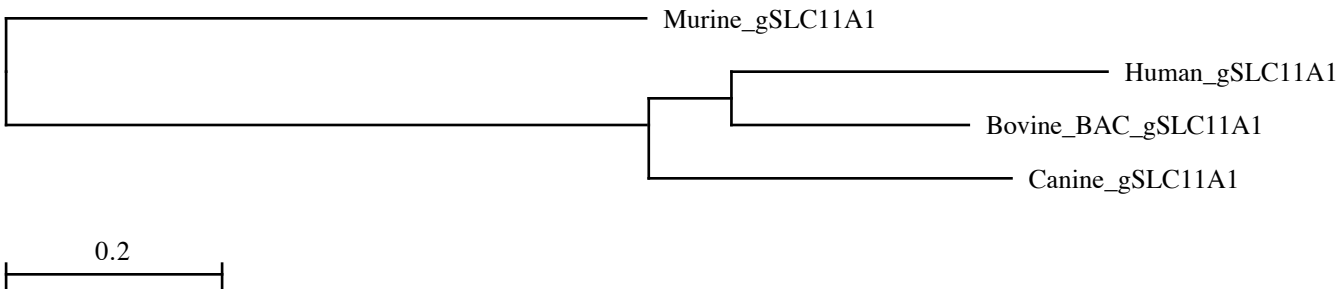

FIG 2. Neighbor-joining relationship tree of mammalian Slc11a1 genomic sequences. The distance indicates the proportional differences between sequences and a scale bar for branch length is shown.

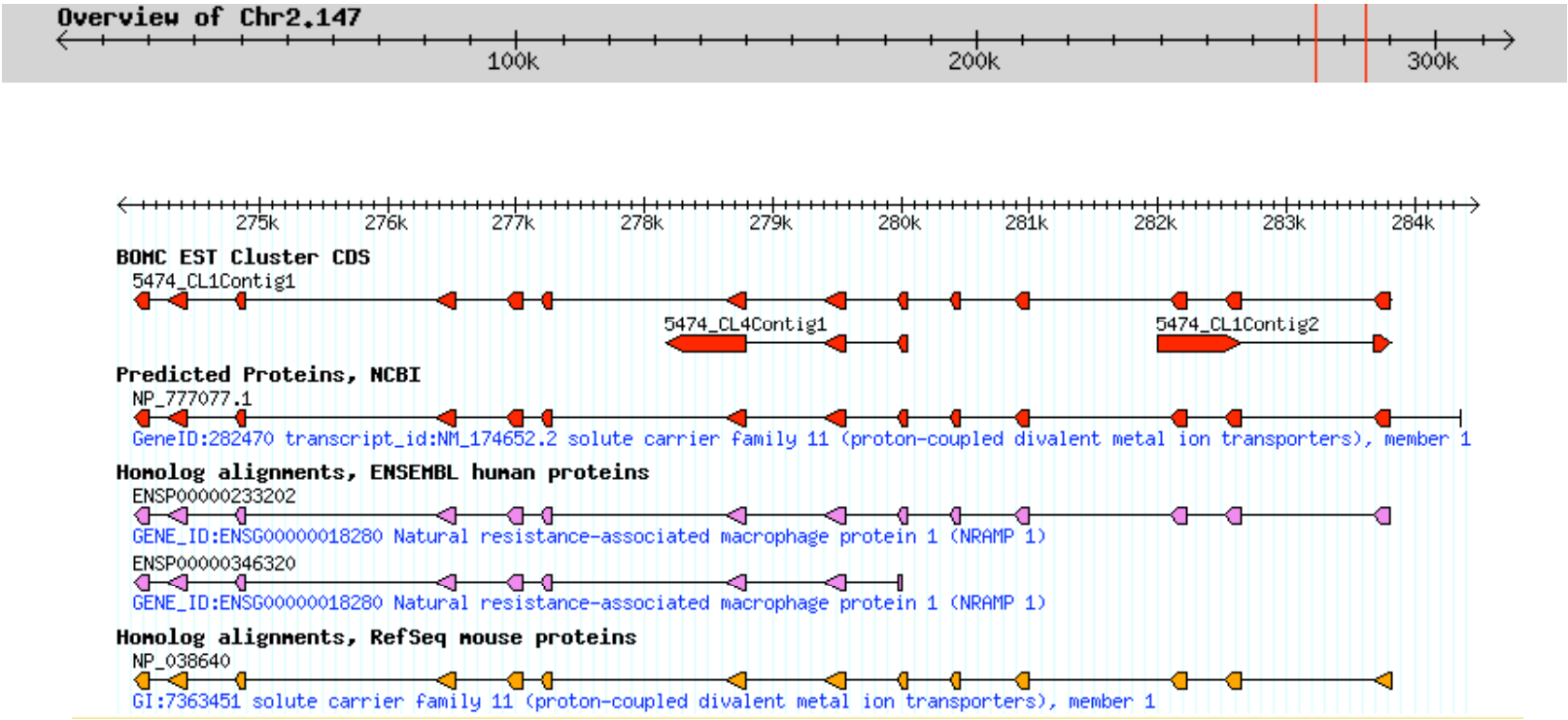

Fig 3. Bovine genome alignment. Bovine *SLC11A1* BAC sequence mapped to Chromosome 2.147 from the Bovine Genome Project assembly from 274547 to 284428. Partial *SLC11A1* sequence was contained in three contigs (5474_CL1Contig1, 5474_CL1Contig2, and 5474CL4Contig1). Alignments to bovine, human and mouse proteins confirmed *SLC11A1* identity.

to ENSEMBLE homolog alignments of the human and mouse NRAMP1/SLC11A1 proteins.

In order to confirm the sequence isolated from the bovine BAC library, three animals phenotyped for resistance and susceptibility to infection by *B. abortus* were chosen for complete genomic sequencing of their naturally occurring *SLC11A1* gene: a susceptible cow (#292), a resistant bull (#86), and the clone of the resistant bull (#86[2]). *SLC11A1* specific PCR and sequencing primers were designed using the SLC11A1 genomic sequence from the BAC clone (Table 3). The sequence covered the complete gene extending from the start codon to the poly-A tail signal, and included all 15 exons, the intervening introns, and 3'UTR.

Seventy-five total segregating SNP sites (S), excluding indels, were identified for these three animals (Table 4), and the Watterson's Estimator of Diversity (θ) calculated from the number of segregating sites was 0.00382. There was no significant linkage disequilibrium found within comparisons of all parsimony informative sites (excluding gaps) by Fisher's exact test using the Bonferroni correction. Among parsimony informative sites, the rate of silent substitution (Ks) was 0.00316 and the rate of non-synonymous substitution (Ka) was 0.00648 between these three cattle. Alignment with the BAC clone sequence and the Bovine Genome Project database confirmed the close nucleotide identity of the *SLC11A1* sequence isolated from these three cattle (Fig. 4)

In order to make some comparisons between *B. abortus* resistant and susceptible cattle populations, and to scan for functional nucleotide variations,

TABLE 3. *SLC11A1* intron/exon specific primers

Primer	Sequence	Product
5CAT_23F	GCTGCAGCAACCTGAGGAAACATCAC	5'UTR+EXON_1
5CAT_161R	ATGTCGACTTACCCGCCCTGACTCTC	
NRAMP1_23F	CACGCCAGCCACTCGCACAG	INTRON_1
NRAMP1_731R	TGGGGATCTTCTCACTTAGGTAGG	
NRAMP1_2F	AGGTGGGGAAGAGGACAGAT	EXON_2
NRAMP1_2R	AGCACTTGAGGACAGAAGCAG	
NRAMP1_705F	GGGACCTACCTAAGTGAGAAGAT	INTRON_2
NRAMP1_1573R	TGGGTCCAGGAATGCGATGCT	
NRAMP1_3FA	CACATCATCCTGCATTTCCA	EXON_3
NRAMP1_3RA	ACCTGTTCCCAGCAACACTG	
NRAMP1_1566F	TGGACCCAGGAAACATTGAGT	INTRON_3
NRAMP1_1954R	GCCCACAGCAGCACCCAGAGC	
NRAMP1_4F	CAGTGTTGCTGGGAACAGGT	EXON_4
NRAMP1_4R	AGCTTTCAATGGCTGTTTGG	
NRAMP1_1992F	CGGCTGGGCGTGGTGACA	INTRON_4
NRAMP1_3251R	CCGGCGGAGAGCAGACT	
NRAMP1_5FA	GCAGGCTTCCAGGCATATAG	EXON_5
NRAMP1_5RA	AAGCCTGCAAGATGACCAAC	
NRAMP1_3204F	GCAGGAAGTCATTGGCACAGC	INTRON_5
NRAMP1_3741R	ATCGAGGAAGAGGAAGAAGAAAGT	
NRAMP1_6F	TGAGGATGGGAGATCCAGAC	EXON_6
NRAMP1_6R	AGCACTTTGTTCCAGGCACT	
NRAMP1_3724F	TCTTCCTCTTCCTCGATAACTACG	INTRON_6
NRAMP1_4106R	CTTCCAGCTTCCGCAACC	
NRAMP1_3795F	GGAGGGGGCTGGCAATGAATGTAA	EXON_7
NRAMP1_4217R	GGGCAGAGTAGGGTGGCTTTGGTG	
NRAMP1_4065F	TGTTGTTCAGCCCCTTTGCTC	INTRON_7
NRAMP1_5048R	AGCCCCCTCCTCAGTCATCCTC	
NRAMP1_4491F	TACAGTCCACGGGGTCACAAAGAA	EXON_8
NRAMP1_5048R	AGCCCCCTCCTCAGTCATCCTC	
NRAMP1_4678F	CCTGCATTCCTCCCTGGTCA	INTRON_8
NRAMP1_5362R	ATGTTGGCCTCTCGGATGTC	
NRAMP1_9F	GCAAGTGCAAAGGTCCTGAG	EXON_9
NRAMP1_9R	CTCAAAGGGCTTAGCTGCAC	
NRAMP1_5411F	AACCTGTTTGTCATGGCTGTCTTT	INTRON_9
NRAMP1_6855R	CGCGTAGTCGTGGAGGCTGCTGTT	

TABLE 3. Continued

Primer	Sequence	Product
NRAMP1_10F	CACCAGTTTTCAACCCTTGG	EXON_10
NRAMP1_10R	AGATGTACAGGGCTGCAGGA	
NRAMP1_6830F	GCCAACAGCAGCCTCCACGACTAC	INTRON_10
NRAMP1_7089R	TACAGGGCTGCAGGACCAAAGAGG	
NRAMP1_11FA	AAGGAGTAAGCAAGCCAACG	EXON_11
NRAMP1_11RA	GGCCTCACCTCCCTAAATCT	
NRAMP1_7139F	GGCACCTACGCGGGACAGTTTG	INTRON_11
NRAMP1_7658	CAGGAGCACAGTGGGCAGGATG	
NRAMP1_12F	CTGAGGGAGGATGGAGGATT	EXON_12
NRAMP1_12R	AATGGGAGTAATGGTGGTTCC	
NRAMP1_7582F	TCCTGAAGCTGCGGTGGTC	INTRON_12
NRAMP1_9251R	GGCGGGCATGCTGGTGAAG	
NRAMP1_13F	GAAGCAGAGGGTGACTGGAG	EXON_13
NRAMP1_13R	TTGGCAAGAGTACAGGCTGA	
NRAMP1_9218F	GCTGCCCATCCTCACCTTCACC	INTRON_13
NRAMP1_9715R	ACGGCGCAGACCAGCACCAT	
NRAMP1_14F	CCAGGGAAGCCCTAGTTACC	EXON_14
NRAMP1_14R	AACTGTGGGCCAGAAGAGTG	
NRAMP1_9708F	TGCGCCGTCAACCTTTACTTCGTG	INTRON_14
NRAMP1_10060R	GGTCCTCCCCTTCTCCTGAT	
NRAMP1_15F	GCAAGGAGAGGAGACCACAG	EXON_15
NRAMP1_15R	ACACACACACCCTTGTCTGG	
NRAMP1_10036F	AGAGGATCAGGAGAAGGGGAGGAC	3'UTR
NRAMP1_10613R	GATTCGGGGACTAAGCACAACAGC	

TABLE 4. Variant positions in genomic *SLC11A1*

Position	BovGenome	BAC	86	86^2	292
21	G	G	T	T	G
22	G	G	T	T	G
119	G	G	G	C	G
120	G	G	G	C	G
2372	T	T	T	A	T
2373	A	A	A	G	A
2391	A	A	A	A	C
2392	G	G	G	G	C
2394	C	C	C	C	A
2395	A	A	A	A	T
3744	T	T	T	T	C
3936	C	C	C	C	T
3938	G	G	G	G	A
3941	C	C	C	C	G
3942	T	T	T	T	G
3946	C	C	C	C	G
4367	C	C	C	C	G
4448	C	C	C	C	A
4732	A	A	A	A	G
4786	G	G	G	G	T
4807	T	T	T	T	G
4849	C	C	C	C	A
4853	C	C	T	C	C
4862	G	G	G	G	T
4903	C	C	C	C	T
4938	T	C	T	T	T
4939	G	T	G	G	G
5165	C	C	T	C	C
5323	G	A	A	G	G
5476	T	C	T	T	T
5734	T	G	T	T	T
5765	G	G	G	A	G
6191	T	C	T	T	C
6396	T	T	T	T	A
6553	G	G	G	G	A
6918	T	C	T	T	C
6997	G	G	G	A	G
7025	C	C	T	C	C
7027	T	T	C	T	T
7056	G	G	G	C	G
7061	G	A	A	A	A
7075	C	C	G	G	G
7307	C	G	G	G	G
7363	G	G	A	T	G
7364	A	A	C	C	A
7366	C	C	C	T	C
7427	C	C	A	C	C
7428	A	A	C	A	A

TABLE 4. Continued

Position	BovGenome	BAC	86	86^2	292
7434	G	G	A	G	G
7474	G	G	A	G	G
7512	G	G	A	G	G
7531	A	A	A	A	G
7816	A	A	C	A	A
7896	A	A	A	A	G
8109	T	T	G	T	T
8116	A	A	C	A	A
8827	T	T	T	G	T
9107	G	A	G	G	G
9322	T	C	T	T	C
9333	A	A	A	T	A
9334	G	G	G	T	G
9335	T	T	T	C	T
9336	G	G	G	T	G
9337	A	A	A	G	A
9347	T	T	T	G	T
9352	G	G	G	T	G
9414	C	C	C	T	C
9444	G	G	T	G	G
9914	T	T	T	C	T
10292	G	G	C	G	G
10416	C	C	C	C	T
10447	A	C	A	A	A
10463	A	A	A	A	T
10474	T	T	T	T	C

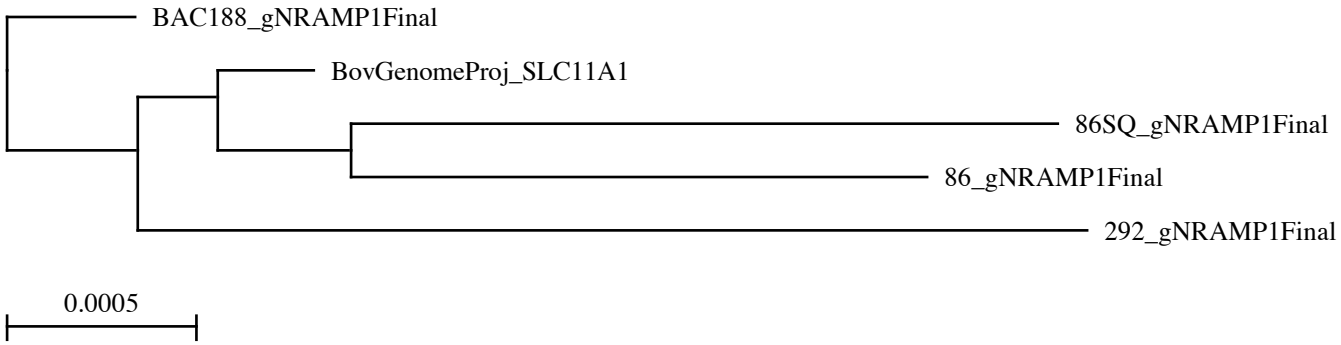

FIG 4. Neighbor-joining relationship tree for resistant and susceptible cattle genomic *SLC11A1*. The distance indicates the proportional differences between sequences and a scale bar for branch length is shown.

sequences of genomic SLC11A1 across the intron-exon junctions containing the complete exons and a short length of flanking intron sequence were concatenated and aligned using Clustal-W. The complete 3'UTR from the stop codon to the poly-A tail signal was included in the alignment for each animal. Haplotypes were reconstructed using the Bayesian algorithm within the Phase v.2.1.1 software and analysis was conducted using DNAsp 4.0.

A total of twenty segregating SNPs were detected in a panel of 18 individuals (8 resistant and 10 susceptible), and 8 of these were in coding regions. A polymorphic GT microsatellite previously described (33) was also detected (Table 5). A total of 17 different haplotypes were predicted and these haplotypes were unevenly distributed between resistant and susceptible populations, with four separate haplotypes predicted in resistant individuals and thirteen predicted haplotypes in susceptible animals (Table 6). Although overall nucleotide diversity of the bovine *SLC11A1* gene was low, there was a differential rate of variability between susceptible and resistant populations. The nucleotide diversity (π) for the eight resistant animals examined was 1.7×10^{-4}, with an average number of nucleotide differences (k) equal to 1.133. For the susceptible animals, $\pi = 8.0 \times 10^{-4}$ and k = 5.264. The rate of nucleotide divergence (K) between the two populations of cattle was $K = 5.7 \times 10^{-4}$. Analyses of the pattern of non-synonymous and synonymous substitutions in resistant and susceptible populations gave results consistent with the deletion of mutations between the two populations, with the average rate of synonymous substitution, $Ka = 4.4 \times 10^{-4}$, lower than the average

TABLE 5. SNPs in short context

Exon	Position	Context	Amino Acid Change
SLC11A1_2	623	GACAC**R**GGCCC	Synonymous
SLC11A1_2	695	CAAGC**R**CCTCC	(His/Arg)
SLC11A1_3	1635	CTCCA**Y**TAACC	Intron
SLC11A1_3	1696	CTGGG**Y**CCCT	Intron
SLC11A1_3	1731	GCTGG**Y**CAAGG	Intron
SLC11A1_4	2048	AGAAA**R**GGGTC	Intron
SLC11A1_5	3415	GACAG**Y**CTTGC	Intron
SLC11A1_5	3497	GCTCA**S**ACATT	(Asp/His)
SLC11A1_5	3654	AGTCA**Y**CTCTG	Intron
SLC11A1_5	3714	GTAAA**K**CGACT	Intron
SLC11A1_5	3716	AAAGCRACTTG	Intron
SLC11A1_6	4135	AGAAG**R**GCTCT	Intron
SLC11A1_8	4791	GCCTG**Y**TCAGG	Synonomous
SLC11A1_9	5476	GACTG**Y**GAGGG	Intron
SLC11A1_9	5569	GCGGG**Y**GGACA	(Ala/Val)
SLC11A1_10	7061	GTGCC**R**ACAGC	(Asp/Asn)
SLC11A1_11	7307	GGTCCT**S**CAGC	(Pro/Ala)
SLC11A1_13	9419	CCTTC**S**GACTCT	Intron
SLC11A1_15	10225	TCTGGC**S**CACAG	(Arg/Pro)
SLC11A1_15	10326	GGGTGG**R**ATGAG	Intron
SLC11A1_3'	10486	ACAAGG**(GT)n**GCAT	untranslated region

TABLE 6. Haplotype reconstruction and population frequencies

Phenotype	Haplotypes	Frequency	S.E.
Resistant	GACCCGCCTGGACTCACCGA 13	0.240633	0.023951
Resistant	GACCCGCCTGGACTCGCCGA 13	0.542554	0.048811
Resistant	GACCCGCCTGGATTCGGCGG 12	0.024197	0.030444
Resistant	GACTCGCCTGAACTCGGCGA 14	0.024624	0.030539
Susceptible	GACCCGCCTGGACTCACCGA 13	0.223895	0.057592
Susceptible	GACCCGCCTGGACTCAGCGA 13	0.023032	0.022725
Susceptible	GACCCGCCTGGACTCGCCGA 13	0.129106	0.016650
Susceptible	GACCCGCCTGGATTCGCCGA 13	0.045455	0.000000
Susceptible	GACCCGCCTGGATTCGGCGG 13	0.069335	0.022698
Susceptible	GACCCGCGTGGACTTACGGA 13	0.037395	0.017360
Susceptible	GACTCGCCTGAACTCGGCGG 13	0.021624	0.022700
Susceptible	GGCCCGTCTTGGCTCACCGA 16	0.068481	0.022730
Susceptible	GGCCTACCTGGGCCCAGCGA 13	0.026895	0.022342
Susceptible	GGTCCACCTGGGCCCAGCGA 14	0.013601	0.020814
Susceptible	AGCCCACCCGGGCCCAGCGA 13	0.032266	0.020671
Susceptible	AGCCCACCCGGGCCTAGGCG 16	0.045406	0.001483
Susceptible	AGCCCACGCGGGCCTAGGCG 16	0.040214	0.014517

rate of non-synonymous substitution, Ks = 11.4 x 10^{-4}. Overall evaluation of nonrandom association of SNPs resulted in a total of eight pair wise comparisons statistically significant by Fisher's Exact test (P=0.001) with the Bonferroni correction, and five significant comparisons by Fisher's test alone (Fig. 5). None of these pair wise comparisons was significant when examined within the resistant population. However, two of the significant comparisons, 695/4135 and 2048/5476 were determined to be significantly associated within the susceptible population.

Analyses using Tajima's (D_{Taj}) test for neutrality along the concatenated exon sequences resulted in an overall lack of significance for both the resistant and susceptible populations. A sliding-window analysis of this parameter did indicate a negative trend in the D_{Taj} index centered between the second and third exons for all sequences analyzed (Fig. 6). Analysis by Fisher's 2-tailed test and χ^2 showed a highly significant association between the A695G SNP within Exon 2 upstream to this region of the gene and disease associated phenotype (Table 7). Associations of lower significance were also calculated for positions G4135A within intron 6, and for position A7061G within exon 10.

DISCUSSION

To date, the only published bovine *SLC11A1* nucleotide variation associated with natural disease resistance to *B. abortus* in cattle is a dinucleotide repeat length polymorphism identified in the 3'UTR of *SLC11A1* cDNA (33). However, some functionally informative SNPs may be overlooked in low expressing alleles due to

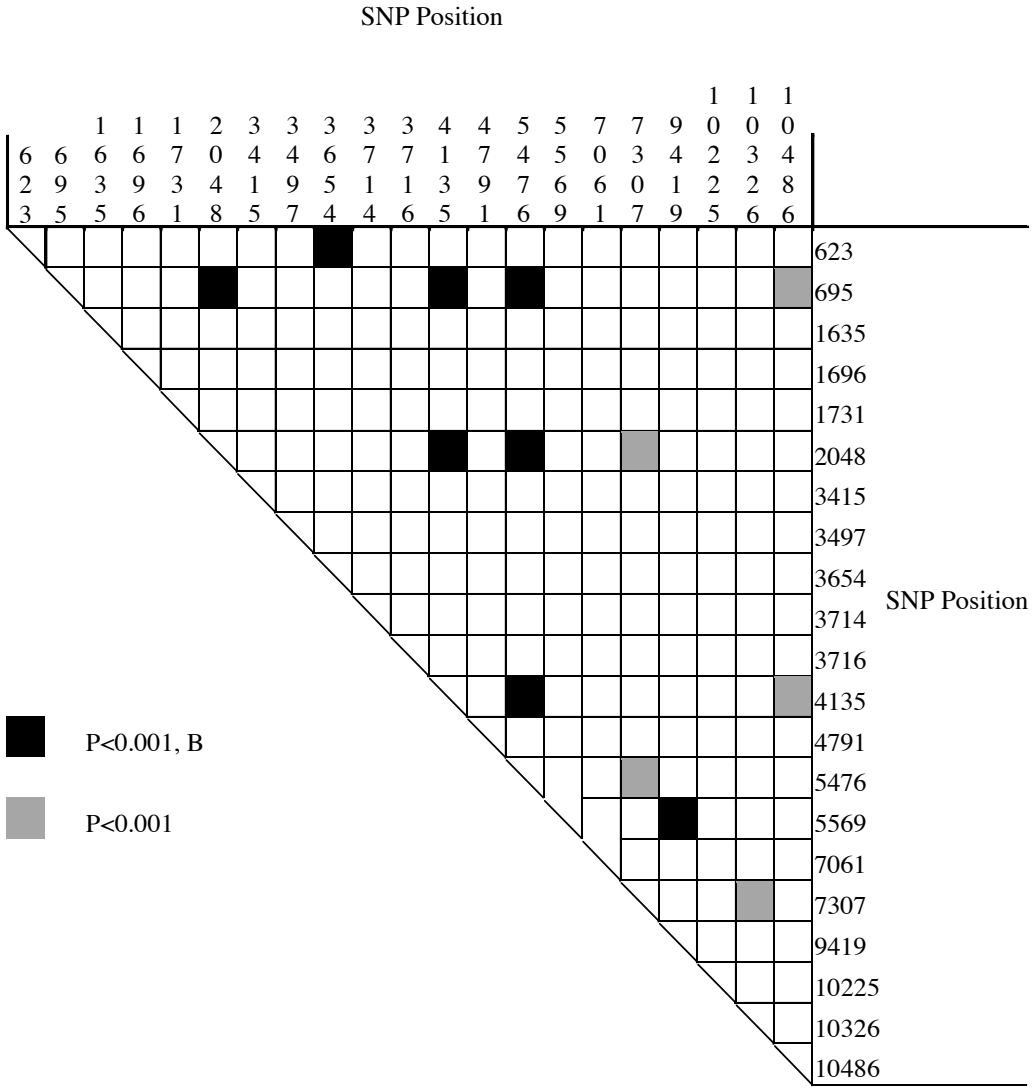

FIG 5. Linkage disequilibrium of informative sites in *B. abortus* resitant and susceptible cattle. Pairwise comparisons were performed using Fisher's Exact test and incorporating the Bonferrini correction procedure. Those nonrandom associations significant at P<0.001 are designated with the darker shading indicating those pairwise comparisons significant with the Bonferrini correction as well.

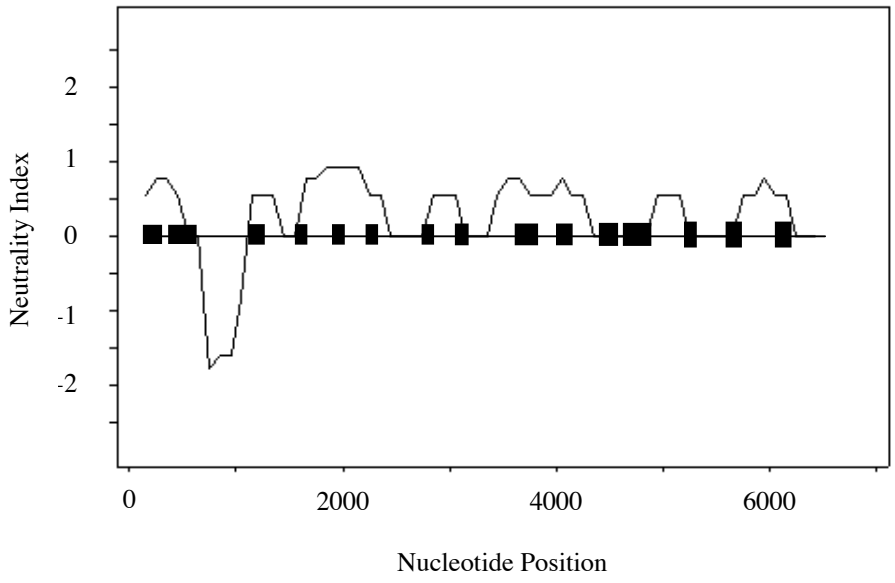

Fig. 6 Neutrality index for exons of genomic bovine *SLC11A1* sequence. Exon sequencing data, including partial flanking intron sequences, were concatenated for analysis. The horizontal axis is the nucleotide position within the concatenated sequence. Filled-in squares indicate the position of exon sequence.

TABLE 7. Variant association with *B. abortus* resistance and susceptibility

SNP Position	Fischer Exact (2-tail)	X^2 (w/Yates correction)
623	0.248933	0.865
695	**0.000002**	**15.283**
1635	1.000000	0.026
1696	1.000000	0.253
1731	1.000000	0.026
2048	0.061166	2.435
3415	0.499289	0.253
3497	0.499289	0.253
3654	0.248933	0.865
3714	0.499289	0.253
3716	1.000000	0.253
4135	**0.014420**	**4.303**
4791	0.624467	0.039
5476	0.061166	2.435
5569	0.248933	0.865
7061	**0.020113**	**5.290**
7307	0.077606	2.385
9419	0.248933	0.865
10225	0.499289	0.253
10326	0.217360	0.855
104860.	0.217360	0.855

the dependence of cDNA isolation on overall allele specific mRNA concentration and isolation. In this study genomic DNA was used as a template for SNP analysis in order to avoid any variation in allele isolation due to differential expression.

In order to search for functionally informative variants encoded within genomic *SLC11A1*, it was necessary to first determine the complete genomic sequence of the gene in order to insure 100% coverage in our search. Previous attempts in other laboratories resulted in only a partial consensus sequence for the gene lacking sequence for intron 7 (24). For this study, access to a bovine bacterial artificial chromosome library facilitated the assessment of the complete genomic sequence from the start codon to the poly-A tail signal. In addition, the recent completion of the second draft of the complete bovine genome sequence and its deposition into free public databases provided a tool for comparison and confirmation of the *SLC11A1* sequence generated in this study.

The resulting sequencing data from the BAC clone was consistent with other published sequences from bovine *SLC11A1*. Intron and exon sizes determined from the BAC sequence were in close agreement with those determined from naturally occurring sequences and the nucleotide sequence aligned with those partial sequences deposited in GENEBANK. The results of this study show that the previously unpublished sequence for intron 7 is 330 bases long within the sequence isolated from the BAC library, with a length and nucleotide sequence that closely aligns with human *SLC11A1* intron 7 (300 bases). This coincides with the Neighbor-Joining relationship analysis showing bovine and human *SLC11A1* in the

same clade, separate from the murine *SLC11A1* and canine *SLC11A1*. This result is not wholly unexpected. At least one feature of human genomic *SLC11A1* sequence is known to have an apparently conserved counterpart in the bovine genomic sequence. Human *SLC11A1* has 4 alleles carrying a distinct series of 3 microsatellite repeats in the 5' promoter region (74). Length polymorphisms within this Z-DNA forming dinucleotide repeat microsatellite affect the rate of transcription of the human *SLC11A1* in response to γ–IFN and bacterial LPS. Bovine *SLC11A1* has a similar microsatellite profile made up of 3 GT repeats, albeit of lower "purity", having accumulated a greater number of interruptions between the repeat sequences, with a disease associated length polymorphism (33).

When the entire genomic sequence was examined for 2 phenotyped cattle and 1 clone, the overall conservation of the gene was again in evidence. Out of 11kb for the complete *SLC11A1* sequence, approximately 1% of the sequence is divergent between the *B. abortus* resistant and susceptible animals. Several nucleotide differences between Bull 86 and its genetic clone, Bull 86^2 were also identified, but these differences are probably due to artifacts and limitations inherent within the sequencing method used. Multiple sequence coverage of these regions of variation, along with targeted genotyping by fluorescent SNP analysis should reconcile these differences between the genetic donor and his clone. If however, these nucleotide variations between Bull86 and his clone cannot be reconciled within SLC11A1, then the next step would be to sequence other genes and search for differences between the two. Sequencing another well-characterized

and relatively short immune response gene, such as Toll-like Receptor 4, a highly conserved gene like the Histone H2 gene, as well as one of the constitutively active housekeeping genes, like GAPDH, should help confirm induced mutations within a genetic clone by comparing the rate of nucleotide variation between these genes. If the rate of mutation is the same for each gene, then this would suggest that induced mutations within a clone is a true phenomenon.

Comparative frequency distribution of polymorphisms within the intron/exon junctions and across the exons indicated an excess of rare alleles in the *SLC11A1* gene of *B. abortus* susceptible cattle as compared to resistant cattle. The average nucleotide difference in the susceptible animals in this study was 5 times greater than in the resistant animals. Haplotype reconstruction and the observed uneven distribution of haplotypes within phenotyped populations indicate greater nucleotide variability within the *SLC11A1* gene in susceptible populations. Thirteen different haplotypes are present in the group of susceptible animals studied, while only 4 occur in the resistant group. This pattern of nucleotide variation indicates an excess of rare haplotypes within the *B. abortus* susceptible cattle population as compared to the resistant population and suggests that these haplotypes may carry mildly deleterious mutations.

A total of twenty informative SNPs were found between resistant and susceptible cattle; 12 were silent, 2 were synonymous, and 6 resulted in amino acid substitutions. However, only 3 SNPs were found to be statistically associated with natural disease resistance to *B. abortus* by Fisher Exact test and Chi-square

analysis, and only the nucleotide transition at position 695 within the second exon was highly significant ($P<0.001$). The adenine (resistant) to guanine (susceptible) transversion at this position resulted in an amino acid substitution of histidine for arginine. Although histidine and arginine both have a basic side chain, the nitrogens in the histidine side chain have a weak affinity for H^+ and therefore have a much weaker positive charge at neutral pH than do the nitrogens in the arginine side chain. Linkage disequilibrium analysis also indicated that this informative site was significantly associated with the susceptible dinucleotide repeat allele found in the 3' untranslated region ($P<0.001$).

The SNP within the second exon is the second disease-associated polymorphism, along with the 3' untranslated dinucleotide repeat length polymorphism, to be found within the bovine *SLC11A1* gene. The discovery of multiple functionally informative nucleotide variants, within *SLC11A1* is not unprecedented. In human *SLC11A1*, four separate polymorphisms were associated with susceptibility to tuberculosis (12). Haplotype marker analysis made up of three intragenic single nucleotide polymorphisms within human *SLC11A1* has demonstrated a significant association of marker level with juvenile idiopathic arthritis (71). It is possible that all or a fraction of these mutations, rather than simply being markers, may each contribute to the reduced functionality of the protein product, either by affecting the rate of gene expression or by expression of a protein product that is functionally less efficient. In the case of the bovine *SLC11A1* gene, the amino acid substitution resulting from the A695G mutation in

exon 2 may lead to a functionally less efficient, yet still viable protein product, while alterations in the 3' untranslated microsatellite length may affect gene expression. In fact, the increased rate of nucleotide substitution seen in susceptible animals when compared to resistant animals may indicate that it is the accumulation of mildly deleterious point mutations that play a large role in the functionality of a gene and its gene product and that only a subset of nucleotide variants show a significant association with function when examined individually. These individually significant mutations may have a greater contribution to the altered function of the gene/protein. The ease with which we can now resolve genetic variation at the nucleotide sequence level provides an opportunity to observe and study such an effect.

CHAPTER III

FUNCTIONAL ANALYSIS OF BOVINE *SLC11A1* 3'UTR MICROSATELLITE LENGTH VARIATION IN CATTLE NATURALLY RESISTANT AND SUSCEPTIBLE TO BOVINE BRUCELLOSIS

INTRODUCTION

Previous studies have shown that the bovine *SLC11A1* gene does not contain a canonical TATA or CAAT box in its promoter region (32). The absence of these conserved elements of transcription initiation is also seen in the promoter of the human homologue of *SLC11A1* (13). However, both the human and bovine *SLC11A1* promoter regions contain tissue-specific transcription factor binding motifs that could facilitate transcription without the need for a TATA or CAAT box. In addition to these protein-binding motifs, the human *SLC11A1* promoter also contains a polymorphic microsatellite made up of multiple guanine-thymine (GT) dinucleotide repeats (13). *In vitro* reporter assays have indicated that length variations within this microsatellite are associated with differential gene expression of the human *SLC11A1* gene (74). Searle *et al.* identified four alleles in this putative enhancer region containing the Z-DNA forming dinucleotide repeat: (1) $t(gt)_5ac(gt)_5ac(gt)_{11}g$; (2) $t(gt)_5ac(gt)_5ac(gt)_{10}g$; (3) $t(gt)_5ac(gt)_5ac(gt)_9g$; (4) $t(gt)_5ac(gt)_9g$. In a luciferase reporter gene construct, alleles 1, 2, and 4 drive low constitutive expression of the reporter gene, while allele 3 drives high levels of gene expression. All four alleles have a similar percentage increase in expression in the presence of interferon-γ. However, the addition of lipopolysaccharide (LPS) has

no effect on alleles 1 and 4, reduces expression driven by allele 2 and enhances expression driven by allele 3. In addition, the high expression allele is functionally linked to infectious disease resistance, while the low expression allele 2 is associated with infectious disease susceptibility (5, 28, 56, 74).

The bovine *SLC11A1* gene has a similar Z-DNA forming dinucleotide repeat microsatellite profile made up of 3 distinct GT repeats separated by intervening sequences. However the bovine microsatellite is found in the 3' untranslated region of the gene rather than in the 5' promoter region (32). The bovine SLC11A1 microsatellite profile also has lower nucleotide purity than the human microsatellite, exhibiting a greater number of intervening sequences between dinucleotide repeats. Four alleles in the third microsatellite containing dinucleotide repeat lengths of GT13, GT14, GT15, and GT16 were found naturally occurring in a population of unrelated crossbred (Jersey X Hereford X Brahaman) cows. Data from *in vivo* challenge studies and *in vitro* macrophage killing assays with *Brucella abortus* in these animals indicated a significant association between natural disease resistance and SLC11A1 genotype (10, 33, 65). The *SLC11A1* allele containing a GT13 3' untranslated region repeat was associated with the resistant phenotype, while *SLC11A1* alleles containing GT14, GT15, and GT16 repeat lengths were associated with the susceptible phenotype. Macrophages expressing the resistant allele also exhibited greater constitutive and inducible levels of SLC11A1 protein in response to LPS/γ-IFN in Western Blot analysis when compared to macrophages expressing the GT16 susceptible allele, suggesting that

resistance to *B. abortus* may be related to the quantity of SLC11A1 protein generated within host macrophages.

It is possible that the region surrounding the polymorphic dinucleotide repeats in the 3'UTR may function as a downstream enhancer region for regulating gene expression. An enhancer is a control region found at a greater distance from the transcription start site than the core promoter, either upstream or downstream of the gene or within an intron (19). Enhancers bind activators and other sequence-specific proteins in a cooperative and combinatorial manner to form an enhancesome. Once constructed, the enhancesome loops out the intervening DNA to interact with proteins on the core promoter and stabilize the transcription complex assembly (66). This looping model facilitates the necessary long-range interactions between promoters and distal enhancers. The immunoglobulin kappa locus is one example of an immune response gene containing a long-range downstream enhancer (60). In this case, an enhancer 9kb downstream of the locus has a profound effect on the induced transcription rate at the distal 5' promoter. One hypothesis is that the 3'UTR dinucleotide repeat in bovine *SLC11A1* links two or more flanking protein binding domains for cooperatively binding activator proteins that make up a distal enhancesome for the *SLC11A1* promoter. Dinucleotide length polymorphisms within this microsatellite structure would change the angular orientation of these interacting proteins by a nonhelical increment and affect the synergistic binding of these proteins and the formation of the enhancesome. This in turn would alter the rate of transcription of the *SLC11A1* gene.

The research described in this chapter set out to identify any putative protein-binding domains within the 3'UTR specific for *SLC11A1* transcription enhancement and *B. abortus* resistance association in the 3'UTR and determine if there was an observable sequence-specific DNA-protein interaction within the 3'UTR of the bovine *SLC11A1* gene. This study also attempted to determine if the 3'UTR of bovine SLC11A1 can function as a downstream enhancer region for gene transcription, and if length polymorphisms in the 3'UTR dinucleotide microsatellite affect the disease resistance phenotype of SLC11A1-expressing macrophages in a phase-dependant manner, indicating a possible effect on the angular orientation of flanking cooperatively binding transcription enhancers within this region of the gene.

MATERIALS AND METHODS

Survey of 3'UTR for Nuclear Protein Binding Motifs. The online software package, Transcriptional Element Search Software (TESS), was used to query the TRANSFAC (v 4.0), IMD (v1.1), and CBIL/GibbsMat (v1.1) online database for putative nuclear protein binding motifs in the 3' UTR sequence of bovine *SLC11A1* (73). The URL for this site is http://www.cbil.upenn.edu/tess.

Nuclear and Cytoplasmic Protein Extraction. Nuclear and cytoplasmic extracts were isolated from peripheral blood derived macrophages unchallenged and challenged with 100 *Brucella* per macrophage. Twenty-four hours post-challenge $5x10^7$ macrophages were collected for protein extraction according to the

protocol described for extraction from a fragile B-lymphocyte cell line (29). Briefly, harvested cells were resuspended in 1mL PBS and microcentrifuged at 200xg for 10 minutes at 4ºC. The PBS was aspirated and the cell pellet was resuspended in 500μL of sucrose buffer containing 0.32M Sucrose, 3mM $CaCl_2$, 2mM Magnesium Acetate, 0.1 mM EDTA, 10mM Tris-HCL (pH 8.0), 1mM DTT, 0.5mM PMSF and 0.5% NP-40 in order to gently disrupt the cell membrane. The lysate was microcentrifuged at 500xg for 10 minutes at 4ºC. The supernatant was transferred to a clean tube and a 0.22x volume aliquot of cold 5x cytoplasmic extraction buffer made up of 0.15M TIME (pH 7.9), 0.7M KCL and 0.015M $MgCl_2$ was added. This cytoplasmic fraction was microcentrifuged 12000xg for 15 minutes at 4ºC. The supernatant containing the cytoplasmic protein extract was transferred to a fresh tube with 25% glycerol and stored at -80ºC. The nuclear pellet was resuspended in sucrose buffer without NP-40 and microcentrifuged 500xg for 5 minutes at 4ºC. The supernatant was aspirated and the intact nuclei were resuspended in 40μL of 4ºC low salt buffer containing 20mM TIME (pH 7.9), 25% glycerol, 1.5mM MgCl2, 0.02M KCl, 0.2mM EDTA, 0.5mM DTT, and 0.5mM PMSF. The nuclei were extracted by titrating with high salt buffer of 0.8M KCl and 1% NP-40 in 1/4 aliquots of the low salt buffer volume, up to 1x volume. The lysed nuclei were mixed at 4ºC for 20 minutes, diluted 1:2.5 with diluent (25mM TIME, 25% glycerol, 0.1mM EDTA, 0.5mM DTT, and 0.5mM PMSF) and microcentrifuged at 13000xg for 15 minutes at 4ºC. The supernatant containing the

nuclear protein extract was stored in 25μL aliquots at -80ºC until used in mobility shift assays.

Electrophoretic Mobility Shift Assay. Two single-stranded DNA probes consisting of 60 bases centered around the first 3' UTR microsatellite repeat were designed and synthesized from the sequence data of bovine *SLC11A1* resistant and susceptible alleles from nucleotide position 10345 to nucleotide position 10405. Position 10362 contained either a T10362 (R allele) or a G10362 (S allele). The oligonucleotides were 3'-end labeled with digoxigenin-11-ddUTP using terminal transferase (Roche). Each binding reaction contained 5μg of protein extract (nuclear or cytoplasmic) from *Brucella* challenged or unchallenged macrophages, labeled oligo probe (0.4ng/μL), poly-d(I-C) (1μg/μL) and poly L-lysine (0.1μg/μL) in 4μL protein binding buffer (Roche) and brought to a final volume of 20μL with distilled water. The reactions were incubated on ice for 30 minutes and electrophoresed on a 6% nondenaturing polyacrylamide gel at 4ºC at 10V/cm in TBE buffer for 6 hrs. The gel was electroblotted to positively charged nylon membrane and UV-crosslinked (120mJ for 3minutes). The DIG labeled probes were visualized by an enzyme-linked immunoassay using anti-Digoxigenin-AP, Fab-fragments, and the chemiluminescent substrate CSPD (Roche).

Production of Reporter Vector Construct. The complete 3' untranslated region from bovine *SLC11A1* was inserted into a PGL3 luciferase reporter vector (Promega) by recombination using the In-Fusion Dry-Down PCR Cloning Kit (BD Biosciences). The insert was amplified from a previously described cloning vector

containing the complete cDNA from the resistance-associated allele of bovine *SLC11A1* (10). The two *SLC11A1*-specific primers, forward primer PGL3-ProV-Fmicro (5'-aaatcgataaggatgcatcaggagaaggggaggac-3') and reverse primer PGL3-ProV-3end (5'-aagggcatcggtcgaccggggactaagcacaacagc-3'), were designed with 5' BamH1 and Sal1 sequences respectively to facilitate insertion of the resulting amplicon and putative enhancer into the reporter vector downstream from the luciferase coding region. The parameters for PCR amplification were 95ºC (10 min) followed by 35 cycles of 94ºC (15 sec), 57ºC (15 sec), and 72ºC (2 min) with 1mM $MgCl_2$, 2μL 10x PCR buffer, 1 unit AmpliTaq Gold (ABI), and 10ng p269-25 DNA in a final volume of 20μL. The resulting PCR product was gel purified. Both the purified PCR product and the PGL3 vector were digested with BamH1 and Sal1 restriction enzymes to generate complimentary ends. The digested PCR product and vector were mixed together at a 2:1 molar ratio in 10μL of distilled water and added to a single tube of lyophilized In-Fusion PCR Cloning reaction mix and incubated at 42ºC for 60 min according to the kit protocol. The In-Fusion reaction product was used to transform competent *E. coli*. Colonies were screened for the *SLC11A1* insert by PCR and plasmid preps from positive clones were sequenced to confirm PGL3 vector and *SLC11A1* 3'UTR fusion.

Luciferase Reporter Assay. 5x10^6 cells from the murine macrophage cell line RAW 264.7 were electroporated with 1μg of the *SLC11A1*-reporter vector fusion product. The electroporation conditions were 200 volts, 1600 faradays, and 129 ohms. The electroporated cells were resuspended in 10mL fetal bovine serum.

The cells were incubated with or without *E. coli* LPS (Sigma) and recombinant mouse IFN-γ(Sigma) at 37ºC and 5% CO_2 for 24hrs. Cells were pelleted at 2500rpms for 30m and washed 1x with 1mL PBS. 100μL of Glo-Lysis buffer (Promega) was added and the cells were incubated for 10 min at room temperature. The cellular extract was transferred to a Dynatech Lumiplate and 100μL of Bright-Glo luciferase substrate (Promega) was added to each well. Luminescence was read on a Dynatech 500 Luminometer. Results were reported as Relative Light Units and calculated as a percentage of observed luciferase activity for a positive control reporter vector containing both the SV40 promoter and an SV40 specific downstream enhancer (PGL3-Control Vector, Invitrogen).

Production of SLC11A1 Expression Constructs. An expression vector carrying the resistance-associated allele of the bovine *SLC11A1* gene under the control of the bovine *SLC11A1* promoter was constructed in the pcDNA3.1 mammalian expression vector (Invitrogen) and subjected to PCR based site-directed mutagenesis to generate susceptible-associated bovine *SLC11A1* expression vectors. The insert was amplified from a previously described cloning vector containing the promoter and complete cDNA from the resistance-associated allele of bovine *SLC11A1* (10). The primers, pcDNA-Bgl2-R20 (5'-gacggatcgggagatccagctatgaccatgattacg-3') and pcDNA-BamH1-U19 (5'-ccgttactagtggatcgttttcccagtcacgacgt-3'), were designed with 5' Bgl2 and BamH1 sequences respectively to facilitate insertion of the resulting amplicon into the expression vector. The parameters for PCR amplification were 94ºC (2 min)

followed by 35 cycles of 94ºC (30 sec), 55ºC (30 sec), 68ºC (4 min), and a final extension step of 68ºC (7 min) with 1mM $MgCl_2$, 2μL 10x PCR buffer, 1 unit Platinum Taq HF (Invitrogen), and 10ng pCR-BoR DNA in a final volume of 20μL. The resulting PCR product was gel purified. Both the purified PCR product and the pcDNA3.1 vector were digested with BamH1 and Bgl2 restriction enzymes to generate complimentary ends. The digested PCR product and vector were mixed together at a 2:1 ratio molar ratio in 10μL of distilled water and added to a single tube of lyophilized In-Fusion PCR Cloning reaction mix and incubated at 42ºC for 60 min according to the kit protocol. The In-Fusion reaction product was used to transform competent *E. coli*. Colonies were screened for *SLC11A1* insert by PCR and plasmid preps from positive clones were sequenced to confirm pcDNA3.1 vector and complete *SLC11A1* gene fusion.

PCR-mediated Site-directed Mutagenesis was used to introduce length variations in the 3'UTR microsatellite of the bovine *SLC11A1* insert. The parental expression vector containing the insert was methylated according to the protocol described in *Short Protocols in Molecular Biology* (78). The reaction mix consisted of 50mM Tris-Cl (pH 7.5), 10mM EDTA, 5mM 2-mercaptoethanol, 16mM s-adenosylmethionine (SAM), 4 units of *dam* DNA methylase and 100ng plasmid in a total volume of 20μL. The reaction was incubated for 1hr. at 37ºC. The complete plasmid was then amplified using two 5' overlapping primers, a plasmid specific primer, 269-25SelectP (5'-ccttgtctggcaggccagtgtgtcacacac-3'), and a mutagenic primer 269-25MUT (5'-cactggcctgccagacaagg$(gt)_n$gcatgcaca-3'), with the position

and number of dinucleotide repeats incorporated into the mutated 3' microsatellite of the *SLC11A1* insert by the mutagenic prime denoted by $(gt)_n$. For this study, the parental GT13 3'UTR microsatellite was mutated to lengths of GT14, GT15, GT16, and GT18. The mutagenesis reaction consisted of 1x PCR buffer, 0.3 mM dNTPs, 1mM $MgSO_4$, 0.3μM primers, 25ng of methylated parental plasmid, and 2.5 units of Platinum *Taq* High Fidelity (Invitrogen) in a 50μL reaction mix. The cycling parameters for this PCR-mediated Site-directed mutagenesis reaction was 94ºC for 2 minutes, followed by 20 cycles of 94ºC for 1 minute, 55ºC for 1m and 68ºC for 6 minutes. A final extension step of 68ºC for 10 minutes was added after the last cycle to ensure complete amplification of the 5kb plasmid construct. The mutagenesis product was cloned in *E. coli* DH5a-T1 competent cells where the methylated parental plasmid was digested by native McrBC endonuclease activity, leaving the unmethylated vector with the mutated insert. Colonies were screened by PCR and positive colonies were minipreped using the High Purity Plasmid Miniprep Kit (Marligen). Plasmid preps from positive clones were sequenced to confirm mutagenesis and the presence of complete vector.

Macrophage Killing Assay. $5x10^6$ cells from the murine macrophage cell line RAW 264.7 were electroporated with 1μg of the expression vector/*SLC11A1* fusion product. The electroporation conditions were 200 volts, 1600 faradays, and 129 ohms. The electroporated cells were suspended at a concentration of 4×10^5 cells per mL in DMEM without antibiotic. 500μL of resuspended cells were loaded to the wells of a 24 well microtiter plate (a total of $2x10^5$ cells per well) and the

cells allowed to adhere to the plate overnight. After replacing the old media with fresh DMEM (no antibiotic), the cells were inoculated with 10 μL of a $5x10^9$ colony forming units per mL stock of *B. abortus* strain 2308 in triplicate wells. The plate was centrifuged for 5 min at 1000 RPMs to synchronize the infection and the cells were incubated 2h at 37ºC. Unbound bacteria were removed by washing the cells with 1 mL of peptone saline three times. 500μL of DMEM with 20μg/mL of gentamicin was added to each well. At time points T=0 and T=24hrs, the cells were lysed with 500μL of 0.5% Tween-20 and the lysate was plated on TSA plates in tenfold serial dilutions for bacterial counts. Bacterial survival was calculated as the ratio of the total number of viable intracellular bacteria at the end of the assay (T=24hr) to the total number of viable intracellular bacteria at the start of the assay (T=0hr). A cutoff of %Survival $\leq$ 80 was used for the resistant phenotype (65).

RESULTS

An initial TESS search of the bovine *SLC11A1* 3'UTR resulted in an overabundance of possible transcription factor binding sites found in the 519 base sequence scanned, many of them apparently unrelated to regulation of macrophage-restricted immune functioning genes similar to *SLC11A1*. When the search was restricted to known transcription factors of immune response genes, four different associated regulatory binding factors were found (Fig. 7). All four factors have previously been shown to be associated with the rate of *SLC11A1* transcription (6), and a deficiency in one of these factors, interferon consensus sequence binding

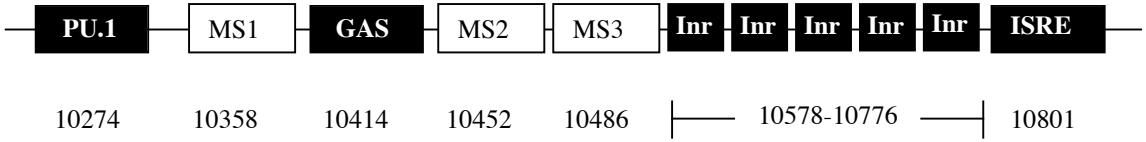

FIG 7. Organization of bovine *SLC11A1* 3'UTR. MS1, MS2, and MS3 denote the genomic nucleotide position of the three distinct dinucleotide repeats. MS3 is the variable microsatellite observed naturally occurring in cattle *SLC11A1*.

protein (ICSBP), has been shown to be associated with susceptibility to brucellosis (50).

A Mobility Shift Assay was conducted to confirm that the 3'UTR region of the bovine *SLC11A1* gene can form a sequence-specific nucleoprotein complex (Fig. 8). Surprisingly, the results of the assay show that little or no DNA-protein complexes are formed when labeled bovine SLC11A1 3'UTR probe is combined with the isolated nuclear protein fraction from both *B. abortus* resistant and susceptible macrophages. However, two discrete complexes were evident when resistant or susceptible probes were mixed with cytoplasmic protein fraction from resistant and susceptible macrophages. Both nucleoprotein complexes disappear in the presence of excess unlabeled probe, indicating that the DNA-protein interaction is sequence specific. The same sequence specific complexes are evident in the nuclear protein fraction isolated from resistant macrophages challenged with *B. abortus* (Fig. 9).

In order to investigate a possible regulatory function for the bovine *SLC11A1* 3'UTR, a luciferase reporter construct was designed with the complete sequence of the 3' flanking region inserted downstream from the enzymes open reading frame. The luciferase activity from the vector construct containing the *SLC11A1* 3'UTR resistant allele (PGL3-GT13) was greater than the levels from the original luciferase construct (PGL3) containing only the SV40 promoter and lacking the enhancer insert when electroporated into unstimulated RAW264.7 macrophages (Fig. 10). Transfected cells stimulated with LPS and IFN-γ

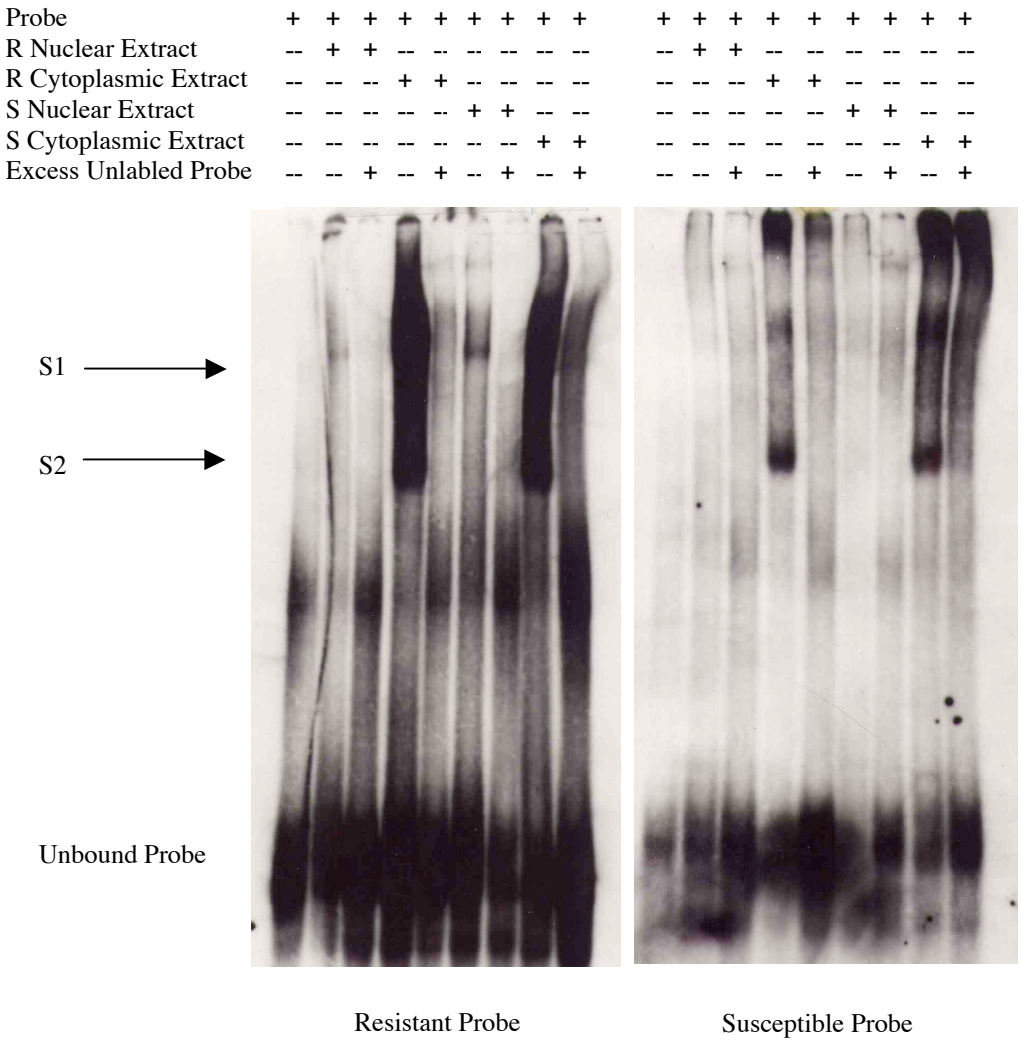

Fig. 8. Electrophoretic mobility shift assay on unchallenged bovine macrophage protein extracts. S1 and S2 are sequence specific nucleoprotein complexes. DNA probe is 60bp fragment of the 3'UTR from bovine *SLC11A1* resistant and susceptible alleles.

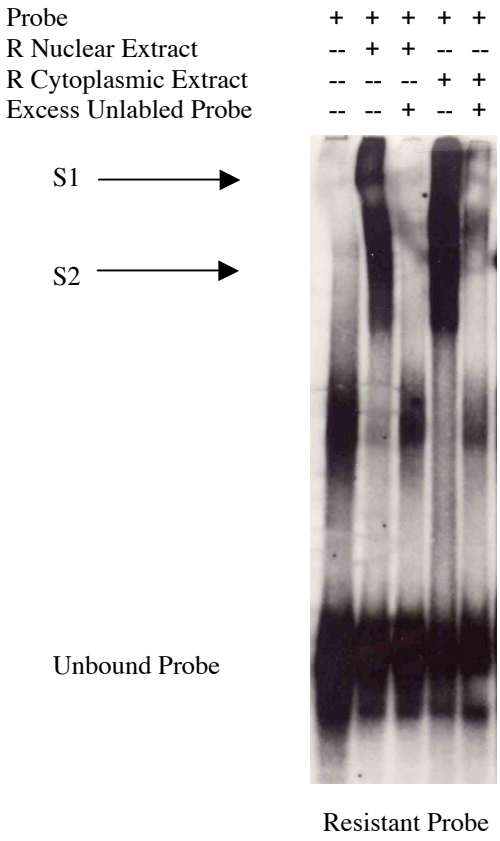

Fig. 9. Electrophoretic mobility shift assay on *B. abortus* challenged resistant bovine macrophage protein extracts. S1 and S2 are sequence specific nucleoprotein complexes. DNA probe is 60bp fragment of the 3'UTR from bovine *SLC11A1* resistant allele.

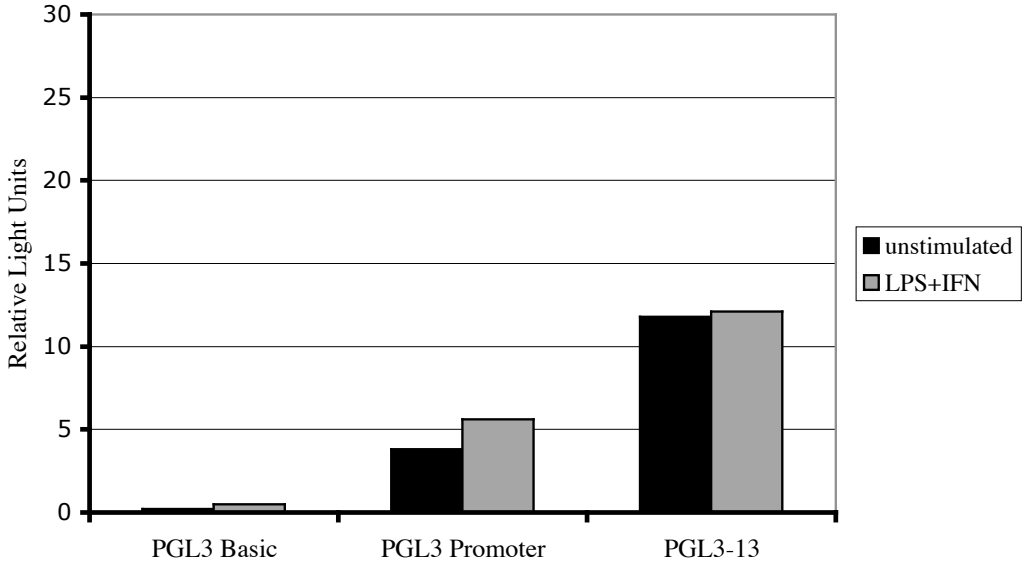

Fig. 10. Luciferase reporter gene assay. Relative light units (RLU) are calculated as the percent luciferase activity for each vector of a PGL3 control vector containing an endogenous SV40 promoter and downstream enhancer. PGL3 Basic vector lacks both a promoter and enhancer. PGL3 Promoter contains an SV40 promoter. PGL3-13 contains an SV40 promoter and the *SLC11A1* 3'UTR resistant allele.

did not show a significant increase in expression levels when compared to unstimulated cells.

Next, RAW264.7 cells were transfected with a eukaryotic expression vector expressing the R-associated 3'UTR sequence allele of *SLC11A* (GT13) and three S-associated 3'UTR sequence alleles (GT14, GT15, and GT16). A fifth transfectant carrying a 3'UTR GT18 dinucleotide repeat, equal to one full turn of the a-helix from the GT13 allele, was also prepared in order to determine if the 3'UTR sequence affects *SLC11A1* in a phase-dependant manner. The five transfectants were then tested for their capacity to retard intracellular replication of *B. abortus* in an *in vitro* killing assay (Fig 11). The macrophages expressing the GT13 resistant allele had a lower bacterial percent survival at 24hrs than did the macrophages expressing the three naturally occurring GT14, GT15, and GT16 susceptible alleles (Fig. 10), falling below the cutoff line of 80% Survival for the resistant phenotype. This result was consistent with the macrophage killing data previously published for the GT16-carrying SLC11A1 allele. When the results for the three susceptible alleles are compared to each other, no significant difference in macrophage killing ability was observed, with the percent bacterial survival for all three clones above 80%. The *B. abortus* percent survival rate for the GT18 expressing transfectant was significantly higher than that for the resistant allele and fell within the range for a susceptible phenotype. The GT18 failed to reconstitute the resistant phenotype in these cells, indicating that although variations in the (GT) repeat length affected the macrophage phenotype, they did not do so in a phase-dependant manner.

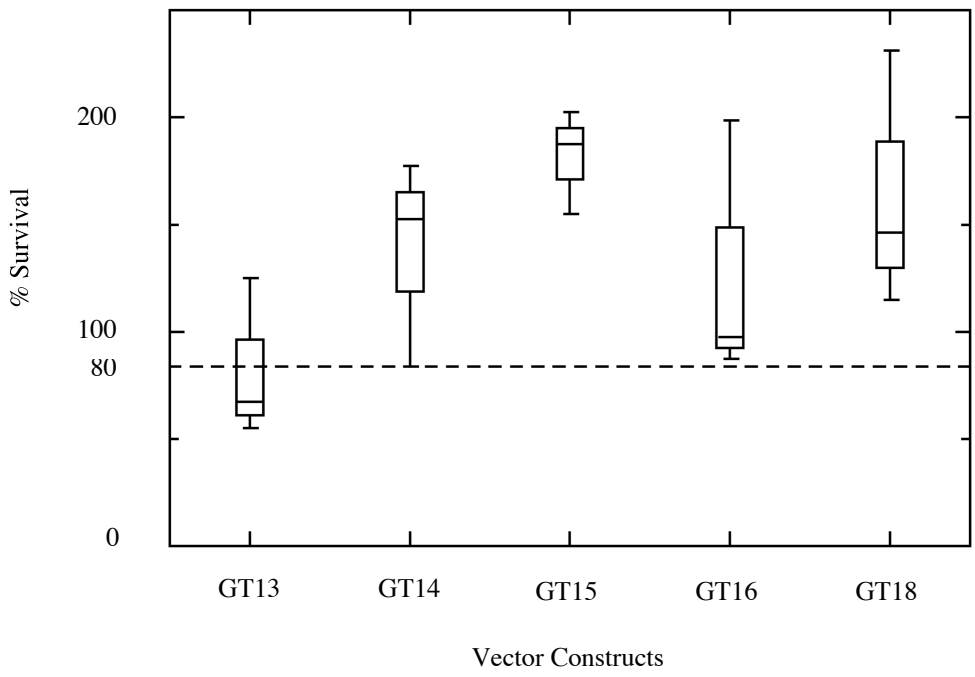

Fig. 11. Survival of *B. abortus* in *SLC11A1* transfected clones. Bacterial survival is expressed as percentage of survival in T=24 clones of T=0 clones. Results for each clone are from 3 independent experiments.

DISCUSSION

Previous studies on bovine *SLC11A1* have suggested that the 3' nucleotide sequence of this gene may act as a downstream enhancer of expression. When the nucleotide sequence of this region was scanned for transcriptional elements, numerous transcriptional factor-binding sites were found. However, four specific transcription factor binding elements stand out for being associated not only with the restricted tissue specific expression of this gene within cells of monocyte lineage, but also with expressed resistance to *B. abortus* infection in mice. The PU.1 site located at nucleotide position 10274 just upstream of the stop codon binds an ETS transcription factor implicated in cell growth and differentiation. The GAS (IFN-γ activation site) element at pst 10414 can mediate the response to IFN type II through Stat-1 binding. The ISRE site (pst 10801) can bind ICSBP, also known as Interferon Response Factor 8 (IRF8). ICSBP is a member of a family of proteins expressed in response to interferon and implicated in the transcriptional control of immune response genes. The multiple initiation response (Inr) sites found just downstream from the 3rd microsatellite repeat in the 3'UTR are normally found in the 5' promoter region of most genes. These transcriptional elements are part of the Miz-1/c-myc antagonist transcriptional factor regulatory structure. Miz-1 and c-myc have been shown to compete for binding on the Inr element. When c-myc is bound to Inr, associated gene expression is repressed. However when Miz-1 is bound to the Inr element, the associated gene expression is stimulated. Data from two published studies suggest a connection between these elements, the expression

of *SLC11A1*, and resistance to *B. abortus* infection in mice. In 2002, Ko *et al.* found that mice deficient in ICSBP had an increased susceptibility to brucellosis (50). A year later Alter-Koltunoff *et al.* documented that *SLC11A1* expression levels are controlled by ICSBP, PU.1, and Miz-1 transcription factors (6). In the murine *SLC11A1* promoter, it has been demonstrated that Miz-1 can bind Inr sites and act as a positive regulator of gene expression by antagonizing the inhibitory effect of c-Myc binding at this element (14). It has also been found that Pu.1 and ICSBP proteins interact to effect transcriptional regulation of genes other than *SLC11A1*. It is possible that these three nuclear binding factors, along with IFN-γ, can interact with each other as well as with their DNA binding elements to form a downstream enhancesome within the 3'UTR and upregulate transcription of the bovine *SLC11A1* gene in response to *B. abortus* infection.

In order to investigate the possibility of sequence specific enhancesome formation on the 3'UTR of bovine *SLC11A1*, an electrophoretic mobility shift assay was performed using a portion of the bovine 3' UTR as a probe. A preliminary study previously showed that a 93bp DNA fragment centered around the 1st microsatellite repeat of the bovine *SLC11A1* 3' UTR could form DNA-protein complexes when mixed with cell protein extract from bovine macrophages (32). However, this study failed to determine if the observed differences in the concentration of complexes formed were due to sequence differences in the probe or differences in the concentration of nuclear binding factor in the extracts isolated from resistant and susceptible macrophages.

In the study described here, the protein extracts from resting resistant and susceptible macrophages were assayed with both resistant and susceptible allele probes to negate any differences in DNA-protein complex formation due to differences in specific factor concentrations. Two specific gel-shifting bands were visualized from both the R and S macrophage protein extracts when combined with either the R or S allele probes. Addition of excess unlabeled probe resulted in a loss of both complexes indicating that the formation of this nucleo-protein complex is sequence specific and not driven by non-specific DNA-protein interactions. Surprisingly, the shift bands were seen in abundance in the reactions containing probe and cytoplasmic extract from these resting macrophages, while the reactions containing nuclear extract and probe resulted in little or no gel shift. The two gel-shift bands are observed in the nuclear fraction of macrophages challenged with *B. abortus* indicating that there may be a translocation of a *SLC11A1*-specific protein factor from the cytoplasm into the nucleus in response to *B. abortus* challenge in bovine macrophages. Cytoplasmic ICSBP has been shown to be translocated across the nuclear membrane in response IFN-γ treatment in some murine cell types (53). One possible scenario is that cytoplasmic ICSBP within *B. abortus*-challenged macrophages is translocated into the nucleus in response to an increase in IFN-γ production. Once in the nucleus, the ICSBP complexes with PU.1 and Miz-1 on the 3'UTR sequence of the *SLC11A1* gene. This nucleoprotein complex then functions as a downstream enhancesome to upregulate transcription of the gene above constitutive levels.

In previous studies, it has been shown that the length polymorphism found within the 3rd dinucleotide repeat of the bovine 3'UTR has a significant association with the resistance phenotype to *B. abortus* infection in cattle. It is possible that this microsatellite sequence variation may be in linkage disequilibrium with another functional variant associated with *B. abortus* resistance, and therefore is merely a marker polymorphism rather than a functional sequence variant. The haplotype reconstruction data from Chapter II. of this dissertation showed that polymorphisms in this dinucleotide repeat are in significant linkage disequilibrium with a non-synonymous point mutation in the 2nd exon. However, it has also been hypothesized that the length polymorphism may have a functional effect on the phenotype by altering the rate of transcription of bovine *SLC11A1*. The dinucleotide repeat may act as a linker between two or more flanking protein factor binding sites. The ligands for these sequence elements could interact with each other and bind cooperatively to the DNA to form a transcriptional enhancesome. If variations in the length of the linking microsatellite alter the spatial orientation of the DNA binding sites, the cooperative binding of the associated protein factors could be diminished and the efficiency of enhancesome formation reduced. If this were true, then changes in dinucleotide repeat length would affect the phenotype in a phase-dependent manner. A luciferase reporter gene construct with the 3'UTR of bovine *SLC11A1* inserted downstream of the coding region does slightly enhance luciferase production above constitutive levels, indicating that the 3'UTR of the bovine *SLC11A1* gene could function as a downstream enhancer region. However,

a mammalian expression vector construct employing the *SLC11A1* promoter sequence, along with the cDNA and 3'UTR of the gene, indicated that although variations in the microsatellite length repeat do effect cell phenotype in a *B. abortus* macrophage killing assay, it is not in a phase dependent manner. There was a difference in killing efficiencies between cells harboring the resistant GT13 allele and those containing the known susceptible alleles, GT14, GT15, and GT16. However, the absence of any discernable difference in killing efficiencies between susceptible alleles and an inability to reconstitute the resistant phenotype by inserting a full-turn incremental variant of 10 bases (GT18) into the microsatellite indicates that there is no phase dependent effect, and the microsatellite polymorphism must effect the cellular phenotype by an alternate mechanism than the one hypothesized here. It is possible that the increased microsatellite length leads to greater instability of *SLC11A1* specific message, and therefore a reduced production of the corresponding protein. A definitive assay that would allow the analysis of message concentration separate from the rate of transcription at the promoter, such as a nuclear run-on assay, would be required to make this statement with any certainty.

CHAPTER IV

CONCLUSION

The complete genomic sequence of bovine *SLC11A1* isolated from a bovine genomic BAC library consists of 10,816 bases from the start codon to the poly-A tail signal sequence. The presence of several important functional and signaling sequence motifs originally described in the cDNA sequence was confirmed in this study of the genomic sequence, and the 15 exons predicted in cDNA by the GT-AG donor-acceptor splice rule for intron-exon junctions were observed in a Pustell Matrix comparison between the cDNA and the genomic sequence for this gene. Comparison with the published complete genomic sequence of other species found that bovine *SLC11A1* has a greater than 50% conserved nucleotide identity with human (60.0%) and canine (50.5%) *SLC11A1*, while the conserved nucleotide identity with the mouse genomic *SLC11A1* is less (40.3%). Phylogenetic analysis of the genomic sequence places bovine and human *SLC11A1* within the same evolutionary clade. These data confirm that the *SLC11A1* is highly conserved across species, even at the genomic level. The highly conserved nature of the protein-coding regions of the gene was already known from comparisons of its cDNA from different species and implies that the structure and function of the gene and its product are extremely important in evolutionary terms. The fact that this conservation rate between species is high within the genomic sequence, including all the non-coding regions, suggests that the regulation of expression of this gene is very important. This conservation of sequence is observed for within species

comparisons as well. When the complete bovine *SLC11A1* genomic sequence for a *B. abortus* resistant animal and a *B. abortus* susceptible animal is compared, there is only 1% divergence across the entire 11kB sequence.

Previous searches for disease-associated sequence variations in the bovine *SLC11A1* gene relied on the expressed form of the gene. However, the concentration of the cDNA used in these SNP scans is dependent upon the level of expression of each allele and the resulting mRNA concentration. It is possible that certain alleles could be under-represented in the mRNA isolation and their associated nucleotide polymorphisms could be overlooked. It is known that susceptible alleles of bovine *SLC11A1* have lower mRNA expression levels than resistant alleles (10, 33). In addition, searching only within cDNA would fail to detect any functionally important nucleotide polymorphisms found within the introns that may be associated with regulation of the genes expression. SNP analysis of the exons within the genomic sequence of bovine *SLC11A1* revealed three previously undetected non-synonymous SNPs that associated with disease phenotypes in *B. abortus* infection. The adenine to guanine transition at position 695 within the second exon is highly significant and resulted in a replacement of histidine for arginine within the protein product. Arginine has a greater positive charge at neutral pH than histidine, and this amino acid change may result in a conformational change of the *SLC11A1* protein product. This conformational change could alter the function of the protein and the resulting ability to resist *B. abortus* challenge. When the individual sequences were concatenated and

haplotype reconstruction performed, this SNP located in the second exon was found to be in linkage disequilibrium with the previously described phenotype-associated 3'UTR microsatellite repeat, indicating a possible association of this non-synonymous SNP with natural disease resistance in cattle.

Another interesting finding was the differential rate of variability observed between the resistant and susceptible populations. The nucleotide diversity for the susceptible population was nearly 5x higher than in the resistant populations. An accumulation of mildly deleterious mutations within the whole gene may result in a less efficient protein product or reduced gene expression and could lead to a reduced ability to resist infection. In this case, the number of accumulated point mutations may be more significantly associated with susceptibility to intracellular pathogens than any one single SNP by itself. Complete sequence analysis of other genes in resistant and susceptible populations would help determine if this is a systemic phenomenon due to a less efficient repair mechanism or if the accumulation of nucleotide variants is specific for the *SLC11A1* gene.

Analysis of a possible functional association of the 3'UTR of bovine *SLC11A1* revealed that although the region does contain sequence specific DNA-protein binding elements, the association of these elements with their protein ligand does not appear to be affected by length variations in the microsatellite repeat in a phase-dependent manner. This would suggest that cooperative binding is not a key element in the effect of this part of the gene on the cellular phenotype. Although variations in the 3'UTR microsatellite do not appear to affect the gene on the level

of transcriptional regulation, there does appear to be a direct affect on gene function and phenotype, possibly at the level of mRNA stability. Xu *et al*. showed that AU-rich elements (ARE) homologous to mRNA destabilizing sequences found in many short-lived mRNAs are implicated in the regulation of human *SLC11A1*. Two similar ARE's are seen in bovine *SLC11A1* 3'UTR. However no variations in the sequence or number of these mRNA destabilizing elements have been identified in *B. abortus* resistant and susceptible cattle populations.

In conclusion, three new genomic sequence variants statistically associated with resistance and susceptibility to *B. abortus* infection in cattle were identified in *SLC11A1*, a differential rate of sequence variation occurred between resistant and susceptible populations, and variations in the previously identified disease associated 3'UTR microsatellite were associated with macrophage killing phenotype, albeit in a non-phase dependent manner. It remains to be determined if the newly discovered SNPs will be functionally associated with macrophage phenotype, and whether or not this is in conjunction with any functional association, such as mRNA stability, found for the 3'UTR microsatellite.

REFERENCES

1. **Aaltonen, L. A., P. Peltomaki, F. S. Leach, P. Sistonen, L. Pylkkanen, J. P. Mecklin, H. Jarvinen, S. M. Powell, J. Jen, S. R. Hamilton, et al.** 1993. Clues to the pathogenesis of familial colorectal cancer. Science **260:**812-816.

2. **Adams, L. G.** 2002. The pathology of brucellosis reflects the outcome of the battle between the host genome and the *Brucella* genome. Vet Microbiol **90:**553-561.

3. **Adams, L. G., and J. W. Templeton.** 1998. Genetic resistance to bacterial diseases of animals. Rev Sci Tech **17:**200-219.

4. **Agranoff, D., I. M. Monahan, J. A. Mangan, P. D. Butcher, and S. Krishna.** 1999. Mycobacterium tuberculosis expresses a novel pH-dependent divalent cation transporter belonging to the Nramp family. J Exp Med **190:**717-724.

5. **Alm, J. S., C. B. Sanjeevi, E. N. Miller, P. Dabadghao, G. Lilja, G. Pershagen, J. M. Blackwell, and A. Scheynius.** 2002. Atopy in children in relation to BCG vaccination and genetic polymorphisms at SLC11A1 (formerly NRAMP1) and D2S1471. Genes Immun **3:**71-77.

6. **Alter-Koltunoff, M., S. Ehrlich, N. Dror, A. Azriel, M. Eilers, H. Hauser, H. Bowen, C. H. Barton, T. Tamura, K. Ozato, and B. Z. Levi.** 2003. Nramp1-mediated innate resistance to intraphagosomal pathogens is regulated by IRF-8, PU.1, and Miz-1. J Biol Chem **278:**44025-44032.

7. **Altet, L., O. Francino, L. Solano-Gallego, C. Renier, and A. Sanchez.** 2002. Mapping and sequencing of the canine NRAMP1 gene and identification of mutations in leishmaniasis-susceptible dogs. Infect Immun **70:**2763-2771.

8. **Atkinson, P. G., J. M. Blackwell, and C. H. Barton.** 1997. Nramp1 locus encodes a 65 kDa interferon-gamma-inducible protein in murine macrophages. Biochem J **325:**779-786.

9. **Barnes, B., B. Lubyova, and P. M. Pitha.** 2002. On the role of IRF in host defense. J Interferon Cytokine Res **22:**59-71.

10. **Barthel, R., J. Feng, J. A. Piedrahita, D. N. McMurray, J. W. Templeton, and L. G. Adams.** 2001. Stable transfection of the bovine NRAMP1 gene into murine RAW264.7 cells: effect on *Brucella abortus* survival. Infect Immun **69:**3110-3119.

11. **Barton, C. H., T. E. Biggs, S. T. Baker, H. Bowen, and P. G. Atkinson.** 1999. Nramp1: a link between intracellular iron transport and innate resistance to intracellular pathogens. J Leukoc Biol **66:**757-762.

12. **Bellamy, R., C. Ruwende, T. Corrah, K. P. McAdam, H. C. Whittle, and A. V. Hill.** 1998. Variations in the NRAMP1 gene and susceptibility to tuberculosis in West Africans. N Engl J Med **338:**640-644.

13. **Blackwell, J. M., C. H. Barton, J. K. White, S. Searle, A. M. Baker, H. Williams, and M. A. Shaw.** 1995. Genomic organization and sequence of the human NRAMP gene: identification and mapping of a promoter region polymorphism. Mol Med **1:**194-205.

14. **Bowen, H., T. E. Biggs, E. Phillips, S. T. Baker, V. H. Perry, D. A. Mann, and C. H. Barton.** 2002. c-Myc represses and Miz-1 activates the murine natural resistance-associated protein 1 promoter. J Biol Chem **277:**34997-35006.

15. **Bruce, D.** 1888. The *Micrococcus* of Malta Fever. Practitioner **40:**241-249.

16. **Cai, L., J. F. Taylor, R. A. Wing, D. S. Gallagher, S. S. Woo, and S. K. Davis.** 1995. Construction and characterization of a bovine bacterial artificial chromosome library. Genomics **29:**413-425.

17. **Campbell, G. A., L. G. Adams, and B. A. Sowa.** 1994. Mechanisms of binding of *Brucella abortus* to mononuclear phagocytes from cows naturally resistant or susceptible to brucellosis. Vet Immunol Immunopathol **41:**295-306.

18. **Campuzano, V., L. Montermini, M. D. Molto, L. Pianese, M. Cossee, F. Cavalcanti, E. Monros, F. Rodius, F. Duclos, A. Monticelli, F. Zara, J. Canizares, H. Koutnikova, S. I. Bidichandani, C. Gellera, A. Brice, P. Trouillas, G. De Michele, A. Filla, R. De Frutos, F. Palau, P. I. Patel, S. Di Donato, J. L. Mandel, S. Cocozza, M. Koenig, and M. Pandolfo.** 1996. Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. Science **271:**1423-1427.

19. **Carey, M. F., and S. T. Smale.** 2000. Transcriptional regulation in eukaryotes: concepts, strategies, and techniques. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

20. **Celli, J., S. P. Salcedo, and J. P. Gorvel.** 2005. *Brucella* coopts the small GTPase Sar1 for intracellular replication. Proc Natl Acad Sci U S A **102:**1673-1678.

21. **Centers for Disease Control and Prevention.** 2000. Suspected brucellosis case prompts investigation of possible bioterrorism-related activity--New Hampshire and Massachusetts, 1999. Jama **284:**300-302.

22. **Chamaillard, M., S. E. Girardin, J. Viala, and D. J. Philpott.** 2003. Nods, Nalps and Naip: intracellular regulators of bacterial-induced inflammation. Cell Microbiol **5:**581-592.

23. **Corbel, M. J.** 1997. Brucellosis: an overview. Emerg Infect Dis **3:**213-221.

24. **Coussens, P. M., M. J. Coussens, B. C. Tooker, and W. Nobis.** 2004. Structure of the bovine natural resistance associated macrophage protein (NRAMP 1) gene and identification of a novel polymorphism. DNA Seq **15:**15-25.

25. **Coutinho, A., and A. Poltorack.** 2003. Innate immunity: from lymphocyte mitogens to Toll-like receptors and back. Curr Opin Immunol **15:**599-602.

26. **Crothers, D. M., M. R. Gartenberg, and T. E. Shrader.** 1991. DNA bending in protein-DNA complexes. Methods Enzymol **208:**118-146.

27. **Cutler, S. J., A. M. Whatmore, and N. J. Commander.** 2005. Brucellosis--new aspects of an old disease. J Appl Microbiol **98:**1270-1281.

28. **Dubaniewicz, A., S. E. Jamieson, M. Dubaniewicz-Wybieralska, M. Fakiola, E. Nancy Miller, and J. M. Blackwell.** 2005. Association between SLC11A1 (formerly NRAMP1) and the risk of sarcoidosis in Poland. Eur J Hum Genet **13:**829-834.

29. **Dyer, R. B., and N. K. Herzog.** 1995. Isolation of intact nuclei for nuclear extract preparation from a fragile B-lymphocyte cell line. Biotechniques **19:**192-195.

30. **Everett, C. M., and N. W. Wood.** 2004. Trinucleotide repeats and neurodegenerative disease. Brain **127:**2385-2405.

31. **Fearon, D. T., and R. M. Locksley.** 1996. The instructive role of innate immunity in the acquired immune response. Science **272:**50-53.

32. **Feng, J.** 1996. Isolation and Characterization of Bovine Nramp1 and Bison Nramp1-A Major Candidate Gene for Natural Resistance to Brucellosis. M.S. Thesis. Texas A&M University, College Station.

33. **Feng, J., Y. Li, M. Hashad, E. Schurr, P. Gros, L. G. Adams, and J. W. Templeton.** 1996. Bovine natural resistance associated macrophage protein 1 (Nramp1) gene. Genome Res **6:**956-964.

34. **Finlay, B. B., and S. Falkow.** 1989. Common themes in microbial pathogenicity. Microbiol Rev **53:**210-230.

35. **Ganz, T., and R. I. Lehrer.** 1998. Antimicrobial peptides of vertebrates. Curr Opin Immunol **10:**41-44.

36. **Greenfield, R. A., D. A. Drevets, L. J. Machado, G. W. Voskuhl, P. Cornea, and M. S. Bronze.** 2002. Bacterial pathogens as biological weapons and agents of bioterrorism. Am J Med Sci **323:**299-315.

37. **Gruenheid, S., E. Pinner, M. Desjardins, and P. Gros.** 1997. Natural resistance to infection with intracellular pathogens: the Nramp1 protein is recruited to the membrane of the phagosome. J Exp Med **185:**717-730.

38. **Hackam, D. J., O. D. Rotstein, W. Zhang, S. Gruenheid, P. Gros, and S. Grinstein.** 1998. Host resistance to intracellular infection: mutation of natural resistance-associated macrophage protein 1 (Nramp1) impairs phagosomal acidification. J Exp Med **188:**351-364.

39. **Harmon, B. G., L. G. Adams, J. W. Templeton, and R. Smith, 3rd.** 1989. Macrophage function in mammary glands of *Brucella abortus*-infected cows and cows that resisted infection after inoculation of *Brucella abortus*. Am J Vet Res **50:**459-465.

40. **Hebert, S., D. Mount, and G. Gamba.** 2004. Molecular physiology of cation-coupled Cl- cotransport: the SLC12 family. Pflugers Archiv European Journal of Physiology **447:**580-593.

41. **Hediger, M. A., M. F. Romero, J. B. Peng, A. Rolfs, H. Takanaga, and E. A. Bruford.** 2004. The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport. Pflugers Arch European Journal of Physiology **447:**465-468.

42. **Hochschild, A., and M. Ptashne.** 1986. Cooperative binding of lambda repressors to sites separated by integral turns of the DNA helix. Cell **44:**681-687.

43. **Horin, P., and J. Matiasovic.** 2000. Two polymorphic markers for the horse SLC11A1 (NRAMP1) gene. Anim Genet **31:**152.

44. **Hu, J., N. Bumstead, P. Barrow, G. Sebastiani, L. Olien, K. Morgan, and D. Malo.** 1997. Resistance to salmonellosis in the chicken is linked to NRAMP1 and TNC. Genome Res **7:**693-704.

45. **Hu, J., N. Bumstead, D. Burke, F. A. Ponce de Leon, E. Skamene, P. Gros, and D. Malo.** 1995. Genetic and physical mapping of the natural resistance-associated macrophage protein 1 (NRAMP1) in chicken. Mamm Genome **6:**809-815.

46. **Huntington's Disease Collaborative Research Group.** 1993. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell **72:**971-983.

47. **Hutt, F. B.** 1958. Genetic resistance to disease in domestic animals. Comstock Publishing Associates, Ithaca, NY.

48. **Kim, J. H., S. Y. Lee, S. H. Lee, C. Sin, J. J. Shim, K. H. In, S. H. Yoo, and K. H. Kang.** 2003. NRAMP1 genetic polymorphisms as a risk factor of tuberculous pleurisy. Int J Tuberc Lung Dis **7:**370-375.

49. **Kishi, F.** 1994. Isolation and characterization of human Nramp cDNA. Biochem Biophys Res Commun **204:**1074-1080.

50. **Ko, J., A. Gendron-Fitzpatrick, and G. A. Splitter.** 2002. Susceptibility of IFN regulatory factor-1 and IFN consensus sequence binding protein-deficient mice to brucellosis. J Immunol **168:**2433-2440.

51. **Ko, J., and G. A. Splitter.** 2003. Molecular host-pathogen interaction in brucellosis: current understanding and future approaches to vaccine development for mice and humans. Clin Microbiol Rev **16:**65-78.

52. **Latham, K. E., and M. E. Westhusin.** 2000. Nuclear transplantation and cloning in mammals. Methods Mol Biol **136:**405-425.

53. **Levi, B. Z., S. Hashmueli, M. Gleit-Kielmanowicz, A. Azriel, and D. Meraro.** 2002. ICSBP/IRF-8 transactivation: a tale of protein-protein interaction. J Interferon Cytokine Res **22:**153-160.

54. **Liu, X. F., and V. C. Culotta.** 1999. Post-translation control of Nramp metal transport in yeast. Role of metal ions and the BSD2 gene. J Biol Chem **274:**4863-4868.

55. **Mark, Z., A. Toren, N. Amariglio, G. Schiby, F. Brok-Simoni, and G. Rechavi.** 1998. Instability of dinucleotide repeats in Hodgkin's disease. Am J Hematol **57:**148-152.

56. **Marquet, S., F. O. Sanchez, M. Arias, J. Rodriguez, S. C. Paris, E. Skamene, E. Schurr, and L. F. Garcia.** 1999. Variants of the human NRAMP1 gene and altered human immunodeficiency virus infection susceptibility. J Infect Dis **180:**1521-1525.

57. **Matthews, G. D., and A. M. Crawford.** 1998. Cloning, sequencing and linkage mapping of the NRAMP1 gene of sheep and deer. Anim Genet **29:**1-6.

58. **McKie, A., and D. Barlow.** 2004. The SLC40 basolateral iron transporter family (IREG1/ferroportin/MTP1). Pflugers Archiv European Journal of Physiology **447:**801-806.

59. **Medzhitov, R., and C. A. Janeway, Jr.** 1997. Innate immunity: impact on the adaptive immune response. Curr Opin Immunol **9:**4-9.

60. **Meyer, K. B., and M. S. Neuberger.** 1989. The immunoglobulin kappa locus contains a second, stronger B-cell-specific enhancer which is located downstream of the constant region. EMBO J **8:**1959-1964.

61. **Mount, D., and M. Romero.** 2004. The SLC26 gene family of multifunctional anion exchangers. Pflugers Archiv European Journal of Physiology **447:**710.

62. **Ozer, J., K. Mitsouras, D. Zerby, M. Carey, and P. M. Lieberman.** 1998. Transcription factor IIA derepresses TATA-binding protein (TBP)-associated factor inhibition of TBP-DNA binding. J Biol Chem **273:**14293-14300.

63. **Paixao, T. A., C. Ferreira, A. M. Borges, D. A. Oliveira, A. P. Lage, and R. L. Santos.** 2006. Frequency of bovine Nramp1 (Slc11a1) alleles in Holstein and Zebu breeds. Vet Immunol Immunopathol **109:**37-42.

64. **Pappas, G., N. Akritidis, M. Bosilkovski, and E. Tsianos.** 2005. Brucellosis. N Engl J Med **352:**2325-2336.

65. **Qureshi, T., J. W. Templeton, and L. G. Adams.** 1996. Intracellular survival of *Brucella abortus*, *Mycobacterium bovis* BCG, *Salmonella dublin*, and *Salmonella typhimurium* in macrophages from cattle genetically resistant to *Brucella abortus*. Vet Immunol Immunopathol **50:**55-65.

66. **Rippe, K., P. H. von Hippel, and J. Langowski.** 1995. Action at a distance: DNA-looping and initiation of transcription. Trends Biochem Sci **20:**500-506.

67. **Rittig, M. G., M. T. Alvarez-Martinez, F. Porte, J. P. Liautard, and B. Rouot.** 2001. Intracellular survival of *Brucella* spp. in human monocytes involves conventional uptake but special phagosomes. Infect Immun **69:**3995-4006.

68. **Rodrigues, V., P. Y. Cheah, K. Ray, and W. Chia.** 1995. malvolio, the *Drosophila* homologue of mouse NRAMP-1 (Bcg), is expressed in macrophages and in the nervous system and is required for normal taste behaviour. EMBO J **14:**3007-3020.

69. **Rozas, J., J. C. Sanchez-DelBarrio, X. Messeguer, and R. Rozas.** 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics **19:**2496-2497.

70. **Rubinsztein, D. C., J. Leggo, R. Coles, E. Almqvist, V. Biancalana, J. J. Cassiman, K. Chotai, M. Connarty, D. Crauford, A. Curtis, D. Curtis, M. J. Davidson, A. M. Differ, C. Dode, A. Dodge, M. Frontali, N. G. Ranen, O. C. Stine, M. Sherr, M. H. Abbott, M. L. Franz, C. A. Graham, P. S. Harper, J. C. Hedreen, M. R. Hayden, et al.** 1996. Phenotypic characterization of individuals with 30-40 CAG repeats in the Huntington disease (HD) gene reveals HD cases with 36 repeats and apparently normal elderly individuals with 36-39 repeats. Am J Hum Genet **59:**16-22.

71. **Runstadler, J. A., H. Saila, A. Savolainen, M. Leirisalo-Repo, K. Aho, E. Tuomilehto-Wolf, J. Tuomilehto, and M. F. Seldin.** 2005. Association of SLC11A1 (NRAMP1) with persistent oligoarticular and polyarticular rheumatoid factor-negative juvenile idiopathic arthritis in Finnish patients: haplotype analysis in Finnish families. Arthritis Rheum **52:**247-256.

72. **Ryu, S., Y. K. Park, G. H. Bai, S. J. Kim, S. N. Park, and S. Kang.** 2000. 3'UTR polymorphisms in the NRAMP1 gene are associated with susceptibility to tuberculosis in Koreans. Int J Tuberc Lung Dis **4:**577-580.

73. **Schug, J., and G. C. Overton.** 1998. TESS: Transcription Element Search Software on the WWW. Technical Report CBIL-TR-1997-1001-v0.0. Computational Biology and Informatics Laboratory, School of Medicine University of Pennsylvania, Philadelphia.

74. **Searle, S., and J. M. Blackwell.** 1999. Evidence for a functional repeat polymorphism in the promoter of the human NRAMP1 gene that correlates with autoimmune versus infectious disease susceptibility. J Med Genet **36:**295-299.

75. **Stahl, P. D., and R. A. Ezekowitz.** 1998. The mannose receptor is a pattern recognition receptor involved in host defense. Curr Opin Immunol **10:**50-55.

76. **Stephens, M., and P. Donnelly.** 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet **73:**1162-1169.

77. **Stephens, M., N. J. Smith, and P. Donnelly.** 2001. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet **68:**978-989.

78. **Struhl, K.** 1999. Enzymatic manipulation of DNA and RNA. John Wiley and Sons, Inc., New York.

79. **Takeda, K., T. Kaisho, and S. Akira.** 2003. Toll-like receptors. Annu Rev Immunol **21:**335-376.

80. **Taleski, V., L. Zerva, T. Kantardjiev, Z. Cvetnic, M. Erski-Biljic, B. Nikolovski, J. Bosnjakovski, V. Katalinic-Jankovic, A. Panteliadou, S. Stojkoski, and T. Kirandziski.** 2002. An overview of the epidemiology and epizootology of brucellosis in selected countries of Central and Southeast Europe. Vet Microbiol **90:**147-155.

81. **Templeton, J. W., R. Smith, 3rd, and L. G. Adams.** 1988. Natural disease resistance in domestic animals. J Am Vet Med Assoc **192:**1306-1315.

82. **Thompson, C. B.** 1995. New insights into V(D)J recombination and its role in the evolution of the immune system. Immunity **3:**531-539.

83. **Tuggle, C. K., C. B. Schmitz, and D. Gingerich-Feil.** 1997. Rapid communication: cloning of a pig full-length natural resistance associated macrophage protein (NRAMP1) cDNA. J Anim Sci **75:**277.

84. **USA Patriot Act.** 2001. United and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism Act of 2001. 125-127 HR 3162 RDS:1st Session of the 107th Congress.

85. **Vidal, S., P. Gros, and E. Skamene.** 1995. Natural resistance to infection with intracellular parasites: molecular genetics identifies Nramp1 as the Bcg/Ity/Lsh locus. J Leukoc Biol **58:**382-390.

86. **Vidal, S., M. L. Tremblay, G. Govoni, S. Gauthier, G. Sebastiani, D. Malo, E. Skamene, M. Olivier, S. Jothy, and P. Gros.** 1995. The Ity/Lsh/Bcg locus: natural resistance to infection with intracellular parasites is abrogated by disruption of the Nramp1 gene. J Exp Med **182:**655-666.

87. **Vidal, S. M., E. Pinner, P. Lepage, S. Gauthier, and P. Gros.** 1996. Natural resistance to intracellular infections: Nramp1 encodes a membrane phosphoglycoprotein absent in macrophages from susceptible (Nramp1 D169) mouse strains. J Immunol **157:**3559-3568.

88. **Wright, S. D., R. A. Ramos, P. S. Tobias, R. J. Ulevitch, and J. C. Mathison.** 1990. CD14, a receptor for complexes of lipopolysaccharide (LPS) and LPS binding protein. Science **249:**1431-1433.

89. **Wyllie, S., P. Seu, and J. A. Goss.** 2002. The natural resistance-associated macrophage protein 1 Slc11a1 (formerly Nramp1) and iron metabolism in macrophages. Microbes Infect **4:**351-359.

VITA

Name:	Christopher John Schutta
Address:	4600 Winchester Drive Bryan, Texas 77802
Email Address:	cschutta@cvm.tamu.edu
Education:	B.S., Biology, Binghamton University, State University of New York, 1985
	M.S., Genetics, Texas A&M University, 1992
	Ph.D., Veterinary Microbiology, Texas A&M University, 2006