

**Identifying genome associations with unique mastitis phenotypes in response to
intramammary *Streptococcus uberis* challenge**

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Lydia Jean Siebert

May 2017

Dedication

To my best friend, my other half, my brain-twin, for always believing in me, convincing me I can do this, and making me a better version of myself.

Acknowledgements

I wish to express my deep gratitude to my major professor, Dr. Gina Pighetti, for supervising the continuation of my graduate education. She taught me many valuable skills both on farm and in the lab, as well as life lessons that will go with me for the rest of my life. In addition to Dr. Pighetti, I would like to thank the other members of my committee, Drs. Cheryl Kojima, Brynn Voy, Meg Staton, and Chunlei Su for their knowledge, input, and critiques.

The assistance and valuable time of Leszek Wojakiewicz and Kylie Stafford were essential to my graduate program, and for that I thank them greatly. I am similarly thankful for the time and knowledge given by the UT Milk Quality Lab and the UT Little River Dairy personnel.

I also wish to thank my parents, Rick and Sheila Siebert, my sister, Mary Lou Siebert, and my aunt, Mary Jo Siebert, immensely for their unfailing support and encouragement during my time as a graduate student. Lastly, I would like to thank all of my friends and fellow graduate students, in particular, Blake Lowe, Lindsey Tyner, and Rachel Dockery, for all their love, laughs, advice, and unwavering faith in me.

Abstract

Mastitis, the inflammation of the mammary gland, dramatically decreases dairy industry revenues and milk quality, making improved control and prevention methods a goal of the industry. Prior research has observed variation in response to experimental challenge with regards to inflammation indicators, *S. uberis* concentrations, and the need for antibiotics. To determine possible causes of the observed variation following experimental *S. uberis* intramammary challenge, we performed genome association analyses (N = 34 – 36 Holstein dairy cows), a linkage analysis, and *S. uberis* milk inoculation (N = 21 Holstein dairy cows) trials. Association analyses were done using Illumina's BovineSNP50 BeadChip and 10 novel phenotypes developed using *S. uberis* challenge data. A total of 52 SNPs were significantly associated ($p < 9.34 \times 10^{-5}$) with our novel phenotypes, with 17 SNPs in regions with prior evidence for mastitis or mastitis associated traits. Linkage analyses were performed using Haploview on BTA 2 to identify loci in linkage disequilibrium with the IL-8 receptor, CXCR1, which has demonstrated importance for neutrophil function critical for mastitis clearance. We determined that CXCR1 is part of two separate haplotype blocks, one that is 13 kb and the other is 440 kb. Between association and linkage analyses, 44 potential candidate genes were identified, where the majority have functions related to inflammation/immunity or gene expression regulation. *S. uberis* milk inoculation trials were conducted in aseptic milk collected within 7 days post-calving from clinically healthy cows with differing CXCR1 haplotype combinations. Milk samples were inoculated with three strains of *S. uberis* (UT366, UT387, and UT888) in vitro and growth was monitored for 10 hours. Both *S. uberis* strain and CXCR1 haplotype combination significantly influenced the rate of *S. uberis* growth in milk ($p < 0.0001$) suggesting an inhibitory or permissive factor towards *S. uberis* growth may exist with concentration varying dependent on CXCR haplotype concentration. Further investigation of the identified candidate genes and proteomics analysis of milk factors could identify more efficient or better genetic selection methods or novel preventative/treatment compounds to combat *S. uberis* mastitis.

Table of Contents

Chapter I: Introduction.....	1
References.....	5
Chapter II: Literature Review.....	9
Mastitis.....	10
Genetic Selection.....	11
Heritability	13
Epistasis.....	16
Genome Wide Association Studies (GWAS).....	17
Loci Associated with Mastitis Phenotypes	18
Candidate Genes Identified by Loci Associated with Mastitis Phenotypes	22
References.....	26
Chapter III: Genome wide association study identifies loci associated with somatic cell count phenotypes following experimental challenge with <i>Streptococcus uberis</i>	33
Abstract.....	34
Introduction	34
Materials and Methods.....	36
Animal Selection	36
Challenge	37
Sampling and Phenotype Description	37
SNP Chip Analysis.....	38
Results	38
Discussion.....	41
Acknowledgements	50
References.....	51
Chapter IV: Genome wide association study identifies loci associated with novel phenotypes following experimental challenge with <i>Streptococcus uberis</i>	55
Abstract.....	56
Introduction	57
Materials and Methods.....	58

Animal Selection	58
Challenge	58
Sampling and Phenotype Description	59
Association Analysis	60
Results	61
Discussion.....	70
Acknowledgements.....	79
References.....	80
Chapter V: Investigation of <i>Bos taurus</i> CXCR1 gene identifies loci with potential mastitis implications using GWAS and linkage analyses.....	85
Abstract.....	86
Introduction	86
Materials and Methods.....	88
Animal Selection	88
Haplotyping.....	89
Challenge	89
Sampling and Phenotype Description	90
Association and Linkage Analysis.....	90
Results	91
Discussion.....	96
Candidate Loci in LD with CXCR1:	96
Candidate Loci Associated with Antibiotic Therapy Groupings:	103
References.....	105
Chapter VI: Determining growth potential of <i>Streptococcus uberis</i> in aseptic milk samples obtained from cows with differing CXCR1 genetic backgrounds.....	108
Abstract.....	109
Introduction	110
Materials and Methods.....	111
Haplotyping.....	111
Milk Sampling and Collection.....	112

Evaluation of Leukocytes in Milk.....	112
<i>S. uberis</i> Cultures	113
Inoculating Milk with <i>S. uberis</i>	113
Statistical Analysis	114
Results	114
Discussion.....	118
Acknowledgements.....	124
References.....	125
Chapter VII: Conclusion	129
References.....	133
Appendix	135
Vita	140

List of Tables

Table 1: Identified loci associated with developed novel SCC phenotypes.	39
Table 2: Identified loci associated with developed novel phenotypes.	62
Table 3: SNPs identified to be in linkage disequilibrium with CXCR1 SNPs using Haploview based on BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) positions.	92
Table 4: Genes and RNAs contained within the haplotype blocks surrounding CXCR1 SNPs based on the <i>Bos taurus</i> UMD_3.1.1 (GCF_000003055.6) genome build.	95
Table 5: Frequency of haplotype combinations observed in our sample population and their corresponding antibiotic therapy groupings.	97
Table 6: Identified loci associated with the CXCR1 antibiotic therapy phenotype.....	98
Table 7: Frequency of haplotype combinations observed in cows whose milk was inoculated with the 3 strains of <i>S. uberis</i>	115
Table A- 1: Comprehensive list of SNPs identified using novel phenotypes with distance to neighboring SNPs and genotype frequencies indicated. Final column indicates status of the SNP following SNP pruning based on minor allele frequencies.....	136

List of Figures

- Figure 1:** Number of cows per SCC cure phenotype category. Ranges of days represent the time range individual cow's SCC returned to below 200,000 cells/mL, which is industry standard for healthy..... 42
- Figure 2:** Manhattan plot of GWAS for SCC cure phenotype using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) in Holsteins (N = 34). The x-axis shows chromosome and SNP order within a chromosome. The y-axis show the $-\log(p\text{-value})$ for each SNP. P-value thresholds are indicated by dashed lines: top black line $p = 1.95 \times 10^{-5}$; bottom gray line $p = 9.34 \times 10^{-5}$ 43
- Figure 3:** Minimum, median, and maximum SCC AUC values for both 7 and 28 d phenotypes displayed in millions of (cells/mL)*d. 44
- Figure 4:** Manhattan plot of GWAS for SCC 7 day AUC phenotype using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) in Holsteins (N = 36). The x-axis shows chromosome and SNP order within a chromosome. The y-axis show the $-\log(p\text{-value})$ for each SNP. P-value thresholds are indicated by dashed lines: top black line $p = 1.95 \times 10^{-5}$; bottom gray line $p = 9.34 \times 10^{-5}$ 45
- Figure 5:** Manhattan plot of GWAS for SCC 28 day AUC phenotype using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) in Holsteins (N = 35). The x-axis shows chromosome and SNP order within a chromosome. The y-axis show the $-\log(p\text{-value})$ for each SNP. P-value thresholds are indicated by dashed lines: top black line $p = 1.95 \times 10^{-5}$; bottom gray line $p = 9.34 \times 10^{-5}$ 46
- Figure 6:** Minimum, median, and maximum for *S. uberis* 1st clearance, severity score, and chronic score phenotypes. *S. uberis* 1st clearance is represented by the number of days until *S. uberis* CFU reached zero based on milk culturing. Severity and chronic scores are represented by the sum of milk and mammary scores over the initial 7 d post-challenge or the entire 28 d challenge period, respectively. Minimums for both severity and chronic score were zero. 64
- Figure 7:** Manhattan plot of GWAS for *S. uberis* 1st clearance phenotype using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) in Holsteins (N = 34). The x-axis shows chromosome and SNP order within a chromosome. The y-axis show the $-\log_{10}(P\text{-value})$ for each SNP. P-value thresholds are indicated by dashed lines: top black line $p = 1.95 \times 10^{-5}$; bottom gray line $p = 9.34 \times 10^{-5}$ 65
- Figure 8:** Number of cows per *S. uberis* shedding cycle category, where the x-axis indicates the number of times *S. uberis* CFU reached zero within the 28 day sampling period. 66
- Figure 9:** Manhattan plot of GWAS for *S. uberis* shedding cycles phenotype using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) in Holsteins (N = 34). The x-axis shows chromosome and SNP order within a chromosome. The y-axis

show the $-\log_{10}(\text{P-value})$ for each SNP. P-value thresholds are indicated by dashed lines: top black line $p = 1.95 \times 10^{-5}$; bottom gray line $p = 9.34 \times 10^{-5}$	67
Figure 10: Minimum, median, and maximum <i>S. uberis</i> CFU AUC values for both 7 and 28 d phenotypes displayed in millions of CFU*d.....	68
Figure 11: Manhattan plot of GWAS for 7 d <i>S. uberis</i> CFU AUC phenotype using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) in Holsteins (N = 36). The x-axis shows chromosome and SNP order within a chromosome. The y-axis show the $-\log_{10}(\text{P-value})$ for each SNP. P-value thresholds are indicated by dashed lines: top black line $p = 1.95 \times 10^{-5}$; bottom gray line $p = 9.34 \times 10^{-5}$	69
Figure 12: Manhattan plot of GWAS for 28 d <i>S. uberis</i> CFU AUC phenotype using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) in Holsteins (N = 35). The x-axis shows chromosome and SNP order within a chromosome. The y-axis show the $-\log_{10}(\text{P-value})$ for each SNP. P-value thresholds are indicated by dashed lines: top black line $p = 1.95 \times 10^{-5}$; bottom gray line $p = 9.34 \times 10^{-5}$	71
Figure 13: Manhattan plot of GWAS for severity score phenotype using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) in Holsteins (N = 36). The x-axis shows chromosome and SNP order within a chromosome. The y-axis show the $-\log_{10}(\text{P-value})$ for each SNP. P-value thresholds are indicated by dashed lines: top black line $p = 1.95 \times 10^{-5}$; bottom gray line $p = 9.34 \times 10^{-5}$	72
Figure 14: Manhattan plot of GWAS for chronic score phenotype using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) in Holsteins (N = 34). The x-axis shows chromosome and SNP order within a chromosome. The y-axis show the $-\log_{10}(\text{P-value})$ for each SNP. P-value thresholds are indicated by dashed lines: top black line $p = 1.95 \times 10^{-5}$; bottom gray line $p = 9.34 \times 10^{-5}$	73
Figure 15: Linkage analysis identifying SNPs on BTA 2 in linkage disequilibrium with CXCR1 SNPs using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA). Blocks are outlined in thick black lines. Colors indicate likelihood of linkage disequilibrium (LD) as calculated by logarithm of odds (LOD; $\log_{10}(\text{likelihood of LD/likelihood of equilibrium where higher values indicate more likely LD})$ and D' (disequilibrium coefficient normalized for the observed allele frequencies; values closer to 1 or 100% indicate higher LD; numbers in squares). Red squares = $\text{LOD} \geq 2$, $D' = 1$. Shades of pink = $\text{LOD} \geq 2$, $D' < 1$. Blue = $\text{LOD} < 2$, $D' = 1$. White = $\text{LOD} < 2$, $D' < 1$; strong evidence of recombination.....	93
Figure 16: Spatial diagram SNPs identified to be in linkage disequilibrium with CXCR1 SNPs using Haploview based on BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) positions.	94
Figure 17: Number of cows per antibiotic therapy phenotype grouping.....	99
Figure 18: Manhattan plot of GWAS for antibiotic therapy phenotype using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) in Holsteins (N = 35). The	

x-axis shows chromosome and SNP order within a chromosome. The y-axis show the $-\log_{10}(\text{P-value})$ for each SNP. P-value thresholds are indicated by dashed lines: top black line $p = 1.95 \times 10^{-5}$; bottom gray line $p = 9.34 \times 10^{-5}$	100
Figure 19: Growth of three strains of <i>S. uberis</i> , represented as back transformed least squares means from $\log_{10}$, in aseptic whole milk obtained from Holstein cows within 7 days post-calving. Letters above the bars indicated means that differ ($p < 0.0001$).	116
Figure 20: The effect of <i>S. uberis</i> strain on growth over time in aseptic whole milk obtained from Holstein cows within 7 days post-calving. All slopes differed from each other significantly over time ($p < 0.0001$).	117
Figure 21: The effect of haplotype combination on <i>S. uberis</i> growth over time in aseptic whole milk obtained from Holstein cows within 7 days post-calving. All slopes differed from each other significantly over time ($p < 0.0001$).	119
Figure 22: Total leukocyte counts observed for each haplotype combination based on QScout analysis performed on aseptic whole milk obtained from Holstein cows within 7 days post-calving. Haplotype combination did not influence total leukocyte counts $p = 6424$	120
Figure 23: Percentages of leukocyte sub-populations observed for each haplotype combination based on QScout analysis performed on aseptic whole milk obtained from Holstein cows within 7 days post-calving. Haplotype combination did not influence the percentages of neutrophils ($p = 0.1509$), lymphocytes ($p = 0.7863$), or macrophages ($p = 0.4903$) observed.	121

Abbreviations and Symbols

AUC	Area under the curve
BTA	<i>Bos taurus</i> autosome
CFU	Colony forming units
CM	Clinical mastitis
DIM	Days in milk
GWAS	Genome wide association study (ies)
IL	Interleukin
LD	Linkage disequilibrium
MY	Milk yield
PL	Productive life
PTA	Predicted transmitting ability
QTL	Quantitative trait loci
SCC	Somatic cell count
SCS	Somatic cell score
SNP	Single nucleotide polymorphism
STA	Standard transmitting ability

Chapter I: Introduction

Mastitis, the inflammation of the mammary gland, is one of the most devastating diseases in the dairy industry. Data shows mastitis can affect 20-50% of any herd causing an increase in cow morbidity and mortality (Wilson et al. 1997). Additionally, mastitis costs the dairy industry upwards of \$2 billion annually (NMC 2005) in the form of lost profits, treatment costs, decreased milk yield, and additional labor costs (Akers and Nickerson 2011; Fetrow 2000). Furthermore, mastitis greatly impacts milk quality, which is observed by a decreased ability to make cheese, decreased shelf life of dairy products, and decreased palatability of dairy products by consumers (Barbano et al. 1991; Klei et al. 1998; Ma et al. 2000).

Mastitis frequently results from pathogenic bacteria invading the mammary gland. Once bacteria invade the mammary gland, epithelial cells and resident macrophages release cytokines which recruit polymorphonuclear neutrophils (PMN) and other leukocytes to the site of infection (Paape et al. 2003). The recruitment of cytokines results in a large influx of leukocytes into the milk, where leukocytes are commonly referred to as somatic cells. Somatic cell counts (SCC), represented as cells/mL, are routinely monitored by the majority of dairy producers and are used as an indicator for the level of intramammary infection or mastitis. There are some resident leukocyte populations in the mammary gland and/or milk (i.e.: macrophages) meaning even when pathogenic bacteria are not present in the gland, SCC is greater than zero. Industry standard considers a SCC below 200,000 cells/mL a healthy gland and SCC above 200,000 cells/mL to indicate either subclinical or clinical mastitis (Dohoo et al. 2011; Schepers et al. 1997; Schukken et al. 2003). Subclinical mastitis is indicated by only an increase in SCC. Clinical mastitis presents with the hallmarks of inflammation including redness, swelling, heat, and pain to the touch, accompanied by a reduction in milk production (NMC 1999). Clinical mastitis also alters milk resulting in flakes, clots, and/or strings composed of bacteria, immune cells, and immune related proteins to be present. In severe clinical cases, milk can become yellow, watery, or even bloody.

Mastitis pathogens can be contagious or environmental in nature. Common control practices utilized on most US dairy farms, have been successful in controlling most contagious pathogens, however, environmental pathogens are still an issue especially in regions with high temperatures/humidity. One common environmental pathogen is *Streptococcus uberis*. *S. uberis* is responsible for a high proportion of both clinical and subclinical mastitis in addition to being commonly isolated during the dry (non-lactating period) (Bradley et al. 2007; Petrovski et al. 2011; Riekerink et al. 2008; Verbeke et al. 2014). *S. uberis* has some unique pathology characteristics that can make clearing the infection from the mammary gland more problematic. Once *S. uberis* enters the mammary gland, it binds lactoferrin, which is a common milk protein (Fang

and Oliver 1999). Lactoferrin can then act as a bridging molecule and allow binding of *S. uberis* to the mammary epithelial cells (Fang et al. 2000). After binding, *S. uberis* then invades the host cells via caveolae dependent endocytosis (Almeida and Oliver 2006; Matthews et al. 1994). Once internalized, *S. uberis* can exploit host cell transduction and persist in the cytoplasm for hours (Almeida et al. 2000; Tamilselvam et al. 2006) effectively evading the host immune system. Additionally, *S. uberis* is especially problematic because, while it typically originates as an environmental pathogen, it can spread throughout the herd in a contagious manner once introduced (Zadoks 2007). Due to the prevalence and impact of *S. uberis* mastitis, increased preventative and control methods are a long standing goal of the industry.

One way the industry has attempted to control mastitis incidence is through genetic selection. Traditional genetic selection strategies have focused on selecting sires whose daughters have decreased SCC or more desirable udder traits that help prevent bacteria entering into the mammary gland (i.e.: tight teat end sphincters) (Rupp and Boichard 2003; Seykora and McDaniel 1986). However, these have had limited success. More recent selection strategies utilizing these same phenotypes have focused on increasing sample sizes and marker density in order to improve identification of loci of interest (Meredith et al. 2013; Meredith et al. 2012; Schnabel et al. 2005). Additionally, genome wide association studies (GWAS) have branched into using clinical mastitis (CM) incidence as a phenotype for selection (Lund et al. 2008; Sahana et al. 2014; Tiezzi et al. 2015). Unfortunately, success has still been limited primarily by low heritabilities and further complicated by the 100+ different pathogens that can cause mastitis, each with its own unique pathology. Thus, the use of pathogen specific phenotypes could be advantageous for the identification of loci that could serve to select against particular causative pathogens, but few such studies have been conducted (Sørensen et al. 2008).

A recent experimental intramammary *S. uberis* challenge demonstrated that cow responses with respect to inflammation indicators, *S. uberis* CFU, and necessity of antibiotics differed (Siebert 2013). Such varying responses could be used as novel phenotypes specific to *S. uberis* mastitis. *We hypothesize that using pathogen specific phenotypes will identify regions of the bovine genome and candidate genes associated with S. uberis mastitis.* Furthermore, the CXCR1 genetic background of the cow significantly influenced these observed responses (Siebert 2013). CXCR1 is the membrane bound receptor for the immune mediator IL-8. Prior studies have demonstrated that CXCR1 is critical for proper neutrophil functions which are critical for mastitis clearance (Rambeaud et al. 2006; Rambeaud and Pighetti 2007; Rambeaud and Pighetti 2005). However, other genes could be in linkage disequilibrium (LD) with

CXCR1 and be attributing the varying responses observed. To address our hypothesis, we propose the following three objectives:

Objective 1: Identify loci and candidate genes associated with *S. uberis* mastitis using novel phenotypes developed from *S. uberis* challenge data.

We hypothesize loci linked to the immune response and S. uberis pathology (i.e.: inflammatory cytokines, lactoferrin) will be associated with the novel S. uberis phenotypes.

Objective 2: Determine the haplotype block surrounding CXCR1.

We hypothesize other genes within the CXCR1 haplotype block may be contributing to the variation in response to S. uberis challenge observed because they are in LD with CXCR1.

Objective 3: Determine if CXCR1 genetic background influences the ability of *S. uberis* to grow in aseptic whole milk in vitro.

We hypothesize CXCR1 genetic background will influence the concentration of factors (i.e.: cellular components, antimicrobial factors, plasminogen, etc.) in milk that could be inhibitory or permissive to S. uberis growth.

References

- Akers, R. M., and Nickerson, S. C. (2011). "Mastitis and its impact on structure and function in the ruminant mammary gland." *Journal of Mammary Gland Biology and Neoplasia*, 16(4), 275-289.
- Almeida, R. A., Calvinho, L. F., and Oliver, S. P. (2000). "Influence of protein kinase inhibitors on *Streptococcus uberis* internalization into bovine mammary epithelial cells." *Microbial Pathogenesis*, 28(1), 9-16.
- Almeida, R. A., and Oliver, S. P. (2006). "Trafficking of *Streptococcus uberis* in bovine mammary epithelial cells." *Microbial Pathogenesis*, 41(2-3), 80-89.
- Barbano, D. M., Rasmussen, R. R., and Lynch, J. M. (1991). "Influence of milk somatic cell count and milk age on cheese yield." *Journal of Dairy Science*, 74(2), 369-388.
- Bradley, A. J., Leach, K. A., Breen, J. E., Green, L. E., and Green, M. J. (2007). "Survey of the incidence and aetiology of mastitis on dairy farms in England and Wales." *The Veterinary Record*, 160, 253-258.
- Dohoo, I. R., Smith, J., Andersen, S., Kelton, D. F., and Godden, S. (2011). "Diagnosing intramammary infections: Evaluation of definitions based on a single milk sample." *Journal of Dairy Science*, 94(1), 250-261.
- Fang, W., Almeida, R. A., and Oliver, S. P. (2000). "Effects of lactoferrin and milk on adherence of *Streptococcus uberis* to bovine mammary epithelial cells." *American Journal of Veterinary Research*, 61(3), 275-279.
- Fang, W., and Oliver, S. P. (1999). "Identification of lactoferrin-binding proteins in bovine mastitis-causing *Streptococcus uberis*." *Federation of European Biochemical Societies Microbiology Letters*, 176(1), 91-96.
- Fetrow, J. "Mastitis: An economic consideration." *Presented at Council, N.M. (ed.), 39th Annual Meeting of the National Mastitis Council*.
- Klei, L., Yun, J., Sapru, A., Lynch, J., Barbano, D., Sears, P., and Galton, D. (1998). "Effects of milk somatic cell count on cottage cheese yield and quality." *Journal of Dairy Science*, 81(5), 1205-1213.
- Lund, M. S., Guldbrandtsen, B., Buitenhuis, A. J., Thomsen, B., and Bendixen, C. (2008). "Detection of quantitative trait loci in Danish Holstein cattle affecting clinical mastitis, somatic cell score, udder conformation traits, and assessment of associated effects on milk yield." *Journal of Dairy Science*, 91(10), 4028-4036.

- Ma, Y., Ryan, C., Barbano, D. M., Galton, D. M., Rudan, M. A., and Boor, K. J. (2000). "Effects of somatic cell count on quality and shelf-life of pasteurized fluid milk." *Journal of Dairy Science*, 83(2), 264-274.
- Matthews, K. R., Almeida, R. A., and Oliver, S. P. (1994). "Bovine mammary epithelial cell invasion by *Streptococcus uberis*." *Infection and Immunity*, 62(12), 5641-6.
- Meredith, B. K., Berry, D. P., Kearney, F., Finlay, E. K., Fahey, A. G., Bradley, D. G., and Lynn, D. J. (2013). "A genome-wide association study for somatic cell score using the Illumina high-density bovine beadchip identifies several novel QTL potentially related to mastitis susceptibility." *Frontiers in Genetics*, 4(229).
- Meredith, B. K., Kearney, F. J., Finlay, E. K., Bradley, D. G., Fahey, A. G., Berry, D. P., and Lynn, D. J. (2012). "Genome-wide associations for milk production and somatic cell score in Holstein-Friesian cattle in Ireland." *BMC Genetics*, 13(1), 1-11.
- NMC. (1999). *Laboratory handbook on bovine mastitis*: NMC.
- NMC. (2005). "The cost of mastitis: Dairy insight research 2005/2006: Final report." *National Mastitis Advisory Committee*.
- Paape, M. J., Bannerman, D. D., Zhao, X., and Lee, J.-W. (2003). "The bovine neutrophil: Structure and function in blood and milk." *Veterinary Research*, 34(5), 597-627.
- Petrovski, K. R., Williamson, N. B., Lopez-Villalobos, N., Parkinson, T. J., and Tucker, I. G. (2011). "Culture results from milk samples submitted to veterinary diagnostic laboratories from August 2003 to December 2006 in New Zealand." *New Zealand Veterinary Journal*, 59(6), 317-322.
- Rambeaud, M., Clift, R., and Pighetti, G. M. (2006). "Association of a bovine CXCR2 gene polymorphism with neutrophil survival and killing ability." *Veterinary Immunology and Immunopathology*, 111(3-4), 231-238.
- Rambeaud, M., and Pighetti, G. (2007). "Differential calcium signaling in dairy cows with specific CXCR1 genotypes potentially related to interleukin-8 receptor functionality." *Immunogenetics*, 59(1), 53-58.
- Rambeaud, M., and Pighetti, G. M. (2005). "Impaired neutrophil migration associated with specific bovine CXCR2 genotypes." *Infection and Immunity*, 73(8), 4955-4959.
- Riekerink, R. G. M. O., Sampimon, O. C., Eerland, V. J., Swarts, M. J., and Lam, T. J. G. M. "Comparing bacterial counts on bare hands with gloved hands during milking." *Presented at Mastitis control: From science to practice. Proceedings of*

International Conference, The Hague, Netherlands, 30 September-2 October 2008.

- Rupp, R., and Boichard, D. (2003). "Genetics of resistance to mastitis in dairy cattle." *Veterinary Research*, 34(5), 671-688.
- Sahana, G., Guldbrandtsen, B., Thomsen, B., Holm, L. E., Panitz, F., Brøndum, R. F., Bendixen, C., and Lund, M. S. (2014). "Genome-wide association study using high-density single nucleotide polymorphism arrays and whole-genome sequences for clinical mastitis traits in dairy cattle." *Journal of Dairy Science*, 97(11), 7258-7275.
- Schepers, A. J., Lam, T. J., Schukken, Y. H., Wilmink, J. B., and Hanekamp, W. J. (1997). "Estimation of variance components for somatic cell counts to determine thresholds for uninfected quarters." *Journal of Dairy Science*, 80(8), 1833-40.
- Schnabel, R. D., Sonstegard, T. S., Taylor, J. F., and Ashwell, M. S. (2005). "Whole-genome scan to detect QTL for milk production, conformation, fertility and functional traits in two US Holstein families." *Animal Genetics*, 36(5), 408-416.
- Schukken, Y. H., Wilson, D. J., Welcome, F., Garrison-Tikofsky, L., and Gonzalez, R. N. (2003). "Monitoring udder health and milk quality using somatic cell counts." *Veterinary Research*, 34(5), 579-596.
- Seykora, A. J., and McDaniel, B. T. (1986). "Genetics statistics and relationships of teat and udder traits, somatic cell counts, and milk production." *Journal of Dairy Science*, 69(9), 2395-2407.
- Siebert, L. J. (2013). "Bovine mammary interleukin-8 receptor expression and genetic association with *Streptococcus uberis* based mastitis."
- Sørensen, L. P., Guldbrandtsen, B., Thomasen, J. R., and Lund, M. S. (2008). "Pathogen-specific effects of quantitative trait loci affecting clinical mastitis and somatic cell count in Danish Holstein cattle." *Journal of Dairy Science*, 91(6), 2493-2500.
- Tamilselvam, B., Almeida, R. A., Dunlap, J. R., and Oliver, S. P. (2006). "*Streptococcus uberis* internalizes and persists in bovine mammary epithelial cells." *Microbial Pathogenesis*, 40(6), 279-285.
- Tiezzi, F., Parker-Gaddis, K. L., Cole, J. B., Clay, J. S., and Maltecca, C. (2015). "A genome-wide association study for clinical mastitis in first parity US Holstein cows using single-step approach and genomic matrix re-weighting procedure." *PLOS ONE*, 10(2), e0114919.

- Verbeke, J., Piepers, S., Supré, K., and De Vliegher, S. (2014). "Pathogen-specific incidence rate of clinical mastitis in Flemish dairy herds, severity, and association with herd hygiene." *Journal of Dairy Science*, 97(11), 6926-6934.
- Wilson, D. J., Gonzalez, R. N., and Das, H. H. (1997). "Bovine mastitis pathogens in New York and Pennsylvania: Prevalence and effects on somatic cell count and milk production." *Journal of Dairy Science*, 80(10), 2592-2598.
- Zadoks, R. N. (2007). "Sources and epidemiology of *Streptococcus uberis*, with special emphasis on mastitis in dairy cattle." *Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*, 2(30), 15.

Chapter II: Literature Review

Mastitis

Mastitis, or mammary gland inflammation, is an ongoing issue for the dairy industry which results in billions in lost revenues annually (NMC 2005). Losses include discarded milk from infected cows, decreased milk production from infected cows, antibiotic treatments, increased labor for treatment of mastitis cows, and replacement of culled cows (Akers and Nickerson 2011; Fetrow 2000; Jones et al. 1994). Additionally, mastitis decreases milk quality, where increased mastitis instance decreases the palatability of dairy products, decreases the shelf life of products, and decreases the ability to make cheese (Barbano et al. 1991; Klei et al. 1998; Ma et al. 2000). Mastitis is defined as the inflammation of the mammary gland, which is often caused by bacteria invading the gland through the teat end. Once inside the gland, bacteria encounter host epithelial cells and macrophages which release cytokines that recruit neutrophils and other leukocytes to the site of infection (Paape et al. 2002a; Paape et al. 2003; Paape et al. 2002b). This recruitment of leukocytes results in a large influx of immune cells into the milk, where they are commonly referred to as somatic cells. Somatic cell counts (SCC) or somatic cell scores (SCS; the $\log_{10}$ of SCC) are routinely monitored by dairy producers and processors, and are used as an indicator of mastitis. There are some resident somatic cells in the mammary gland and/or milk (i.e.: macrophages) meaning even when pathogenic bacteria are not present in the gland, SCC is greater than zero. Industry standard considers a SCC below 200,000 cells/mL a healthy gland and SCC above 200,000 cells/mL to indicate mastitis (Dohoo et al. 2011; Schepers et al. 1997; Schukken et al. 2003). Mastitis can be clinical or subclinical, where both are hallmarked by an increase in SCC above 200,000 cells/mL, but clinical mastitis is also indicated by inflammation and altered milk production, as well as reduced milk production (NMC 1999).

Several factors can influence the probability of a cow getting mastitis including age, parity, stage of lactation, genetics, breed, herd management, climate, and causative pathogen. Mastitis incidence increases with parity, or the more offspring a cow has the more likely it is she will contract mastitis (Ruegg 2003). Stage of lactation also influences mastitis risk, with dry (non-lactating) cows that are not being regularly handled/milked and cows in early lactation are more susceptible to developing mastitis than cows in late lactation (Oliver and Mitchell 1983). Climate can also influence mastitis incidence, specifically warm and moist climates favor the growth of most pathogens, which increases their overall numbers in the cow's environment and increases the probability that the pathogen comes into contact with the cow and causes mastitis (Hogan and Smith 1987). The influence of these various factors on mastitis incidence lead to a high degree of variation observed in mastitis incidence across herds

and breeds of dairy cattle, where most herds demonstrate a mastitis incidence of 20-50% (Wilson et al. 1997).

Mastitis causative pathogens can be environmental or contagious in nature. Contagious pathogens are transferred from cow to cow. Spread of contagious pathogens is often facilitated by the milking process where bacteria present on one cow can be passed on to the next cow via milker's hands or the milking claw (Hogan and Smith 1987). Environmental pathogens are those present in the general environment of the cow (i.e.: soils) and gain entrance to gland via proximity (Hogan and Smith 1987; Hogan et al. 1989; Smith and Hogan 1993). Common mastitis control and prevention strategies (i.e.: pre-dips, post-dips, etc.) used on most US dairies have been successful in controlling most contagious pathogens, but environmental pathogens remain an issue, particularly in areas with warm and moist climates. *Streptococcus uberis* is a common mastitis pathogen that is responsible for a high proportion of both clinical and subclinical mastitis and is commonly isolated from the dry (non-lactating) period (Bradley et al. 2007; Jayarao et al. 1999; Oliver 1988; Petrovski et al. 2011; Phuektes et al. 2001; Riekerink et al. 2008; Verbeke et al. 2014). Furthermore, while *S. uberis* originates as an environmental pathogen, it can spread through a herd in a contagious manner (Zadoks 2007). Thus, control and prevention of mastitis, particularly, *S. uberis* mastitis, is a long standing goal of the dairy industry.

Genetic Selection

One method employed to help control and prevent mastitis is genetic selection. The idea of using genetics in animal breeding to improve the overall health of any given population by reducing diseases and improving animal performance (i.e.: increasing milk yield) has been around for decades (Soller and Beckmann 1983). Genetic selection is the process that makes certain desirable traits more prevalent than other less-desirable traits within a given population by carefully breeding individuals with the desired traits to make the next generation. Genetic selection is facilitated by the use of PTAs (predicted transmitting abilities) or STAs (standard transmitting abilities), are estimates of the genetic superiority or inferiority that a cow or bull will transmit to their offspring for given traits (The Holstein Foundation 2015).

PTAs are available for a variety of traits including: milk yield, fat percentage, type traits (i.e.: traits specific to the breed), udder characteristics, somatic cell score (SCS), and others, and are often used to rank bulls by their genetic merit. PTAs are calculated using best linear unbiased predictions (BLUP), which take into account not only the trait measurements, but also information on pedigree, environment, and relationships between traits to determine the best prediction of the trait that will be passed on to

offspring (Campbell and Marshall 1975). PTAs are represented as differences from the genetic base, or average for that trait specific to breed, and are expressed in units contingent on the properties of the trait being measured (Wattiaux 1995). Thus, displaying PTAs for different traits for one individual bull on the same graph would help to determine if a given bull is an extreme for one trait over others, but is difficult because of varying units. Standard transmitting abilities (STAs) are determined from PTAs of linear type traits so that traits are represented on a standardized scale and able to be depicted on the same graph (The Holstein Foundation 2015). STAs are typically depicted as within 3 standardized units of zero, positive or negative, where values closer to zero indicate the bull is more likely to pass on the breed average phenotype for that trait to its offspring (Holstein Association USA 2013).

Also important when looking at PTAs, is reliability, often expressed as %R. Reliability refers to the estimated accuracy of the calculated PTA and is a factor of the amount of information used to calculate the PTA (Wattiaux 1995). Reliability is influenced by several factors including: reliability of PTA values for parents/other relatives of the bull, number of records available for that bull's offspring, and the number of herds (varying environments) those offspring are located in (Wattiaux 1995). As the amount of information known about a given sire increases, the reliability of their calculated PTAs increases. When a calf is born, their PTAs are based solely on the PTAs of their parents and relatives since the calf does not have offspring yet, which results in relatively low reliability. The reliability of calf PTAs can be improved using genomic testing. Genomic testing involves genotyping the calf at many single nucleotide polymorphisms (SNPs). The resulting genotypes are then compared to a subset of individuals (i.e.: training population) that have reliable PTAs based on their offspring and have also been genotyped, in order to determine what components of PTAs can be attributed to genetics (Wiggans et al. 2011). Based on the comparison, a selection index is developed that combines the parent/relative PTA information and genetics to improve the reliability of the calf PTA (The Holstein Foundation 2015; VanRaden et al. 2009). As the calf ages and has offspring whose phenotypes can be measured, this information is also included in the PTA calculation, further improving the reliability since more information is used. Bulls that have 100 daughters in 100 different herds typically have about 88% reliability without genetic information included (Wattiaux 1995). However, even when information about a bull's parent/relative PTAs, hundreds or thousands of daughters, and genetics are included in PTA calculation, the highest reliability reachable is 99% since the set of genes the daughters receive from the bull is always random (The Holstein Foundation 2015).

PTAs (i.e.: genetic selection) have been used to decrease mastitis incidence. Such selection began with selecting against udder traits that would make bacteria entrance into the gland easier. Selection against deep udders, long teats, loose udder attachment, and unbalanced udders were correlated with less mastitis incidence likely due to reduce proximity of the udder to organic matter on the ground that is heavily laden with mastitis causing bacteria (Lund et al. 1994; Rupp and Boichard 1999; Sørensen et al. 2000). Selection efforts against mastitis progressed to selection against cows with high somatic cell scores (SCS, $\log_{10}$ of SCC). Since somatic cell counts are used as an indicator of mastitis, where industry standard considers concentrations above 200,000 cells/mL to indicate mastitis (Dohoo et al. 2011; Schepers et al. 1997; Schukken et al. 2003), cows with higher average SCS over a lactation would correlate to cows with frequent mastitis bouts or chronically mastitic cows. More recently, there has been some selection based on clinical mastitis, but this has been limited to Scandinavian countries where treatment of clinical mastitis is performed only veterinarians, which facilitates the record keep process (Heringstad et al. 2003; Philipsson et al. 1995). Success using these mastitis selection traits has been limited, likely due to the complexity of the disease and the variation in causative pathogens with each having unique pathology. Furthermore, complex traits, like SCS, are easily influenced by other factors including heritability and epistasis.

Heritability

Heritability is defined as the proportion of phenotypic variance attributed to differences in genotypes (Lush 1949). When attempting to identify causes of phenotypic differences, the differences are best represented by their variance (i.e.: the differences squared), and can be broken into two subset definitions (Lush 1949). The broad sense of heritability is a function of genotype variance divided by phenotype variance and is used in contrast to the environmental effects. However, this broad definition (represented as H^2) relies on the genotype being transmitted as one unit, which is unrealistic. Instead, genes that combine to produce a phenotype can segregate and recombine to generate new combinations. Such recombinations could cause unique gene interactions resulting in genotype averages that are substantially different than the average genotype within the general population. Thus, the narrow sense of heritability (represented as h^2) includes only the additive effects or variance of the genes (the variance attributed to parent-offspring resemblance) divided by the phenotype variance. The result of narrow heritability is the expected fraction of phenotypic variation between parents that can be expected to be passed on to the next generation (Lush 1949).

The representation of heritability as h^2 became popular because of Sewall Wright's use of h to represent heredity in describing his path coefficient model. The path coefficient model describes the correlation squared (i.e.: h^2) as the proportion of phenotype attributed to 'the path from genotype to phenotype' (Wright 1920). In addition, Ronald Fisher described the resemblance among relatives using regression and correlation coefficients while estimating the variance in phenotype that is attributed to genotypes as a percentage (Fisher 1919). However, the term heritability is credited to Jay L. Lush, who used it to describe proportion of variance due to hereditary factors (Lush 1940). Over the years, statistical methodology in partitioning variation and estimating heritability has been thoroughly developed and has largely converged across disciplines and species (Lynch and Walsh 1998). The percentages of variance, first described by Fisher, have come to correspond to the current definitions of broad- and narrow-sense heritability (Visscher et al. 2008).

Presently, the equation used to calculate heritability is a ratio of variances. The numerator is the variation from additive genetic values of a given population or breeding values, defined as the sum of the parents' genes average effect that results in a phenotypic mean value for all the parents' progeny (Falconer 1975). The denominator is the total observed variation that has been adjusted to exclude variation attributed to certain fixed effects (i.e.: sex, age) (Visscher et al. 2008). The result of the heritability equation is a proportion, represented as a value between 0 and 1, called the heritability index (HI). The higher the heritability index, the higher the probability of the measured phenotype being passed on to the next generation through parent genotypes within the population it was measured.

Heritability estimates are based on observed differences between actual and expected resemblance between relatives. When selection pressure is applied, the ratio of the amount of change in the mean phenotype between generations to the difference in mean phenotype between parents is used to estimate heritability (Visscher et al. 2008). Determined over several generations, the heritability is the regression or ratio of the cumulative selection response to the cumulative selection differentials (Falconer 1975). Traditionally, heritability was measured using simple balanced designs such as those involving regressions between offspring and parent phenotypes, correlation of siblings' phenotypes, and correlations between twin pairs of monozygotic and dizygotic origins (Falconer 1975). More recent linear mixed models have allowed for heritability calculation from more complex designs including those where phenotype information is available for a variety of relationships, unbalanced designs, and those that include environmental variances. The 'animal model' is one linear mixed model, which was originally designed for use in animal breeding, where an additive genetic relationship is

fitted for each individual in the pedigree and weighted appropriately according to their sampling variance (Lynch and Walsh 1998; Quaas and Pollak 1980). Estimations of genetic variance components from complex pedigrees are accomplished using residual maximum likelihood methodology or Markov chain Monte Carlo methods, with the latter deriving more conservative Bayesian estimates (Patterson and Thompson 1971; Sorensen and Gianola 2007). To obtain accurate estimates of heritability, hundreds of observations are necessary to reduce the sampling error which is contingent on low sampling variance and controlled by large observation numbers (Falconer 1975).

Considering heritabilities and the variances used to calculate them are parameters of a given population it is possible only to estimate heritability opposed to determine a defined value of heritability across populations. This is due to the fact that heritability is not merely a factor of genetics and can be influenced by several factors. The main factor that influences heritability is the environment. In fact, Jacquard (1983) states that heritability is actually a result of two factors: the genes transmitted to offspring at the time of conception, and the environmental contribution, in all its forms, during the entirety of development. Additionally, heritability can be confounded by factors such as sex, age, and cohort. Thus when determining heritability, these factors are adjusted for when calculating the variance in phenotype, which is the denominator of the heritability equation (Visscher et al. 2008). If identifiers for a confounding factor are not present, thus not allowing for correction of the confounding factor, phenotypic variance calculated will be higher and result in a lower heritability calculation (Visscher et al. 2008). Another factor that can influence heritability is time. If over time a selection pressure is applied to a particular phenotype, then the variance of that phenotype decreases. Decreased phenotype variance results in a higher calculated heritability such as observed by the increase in first lactation milk yield between the 1970's and the present day (Berry et al. 2003). The fact that heritability can be influenced by a variety of factors is supported by the fact that heritabilities calculated for the same trait in different populations often vary due to the different combined effects of the environmental factors each population is subjected to (Visscher et al. 2008). However, heritabilities calculated for the same trait can be very similar to each other regardless of the population they are measured in.

With respect to artificial insemination programs such as those used within the dairy industry, heritability is used to speculate the effectiveness of using phenotypic information to determine breeding value when selecting mating pairs to improve herd genetics. High heritabilities for traits that are easy to measure (i.e.: milk yield) become the basis for producers to develop easy and effective breeding schemes that improve the overall herd over time. However, traits with low heritability (i.e.: somatic cell count)

are more difficult to develop breeding schemes around since they require analysis of many related individuals to predict somewhat accurately and can be strongly influenced by environmental factors like causative pathogen.

Epistasis

Epistasis refers to the interaction of genes that are not alleles and it is thought to influence many complex traits/diseases such as SCS/mastitis. Selection of traits/diseases with epistatic influences is more difficult because you need to select for a collection of genes working together to cause a desired phenotype instead of one gene causing the phenotype. The term 'epistatic' was first used in the early 1900's to describe one loci preventing the manifestation of another loci effect, which was seen as an extension of allele dominance (Bateson and Mendel 1909). Since its first description, there have been many refinements to the concept of epistasis, which has resulted in variations in the definition. The original definition is most similar to the definition used today by biologists where epistasis refers to the effect of one locus on another, often resulting in the 'masking' of one locus in the presence of the other (Siemiatycki and Thomas 1981). The original definition implies directional epistatic effects where loci A affects loci B, but not vice versa, whereas, the more modern definition used by biologists does not imply directionality, making this definition of epistasis more generalized than the original (Bateson and Mendel 1909; Neuman et al. 1992). In models where epistasis is detected, it is often assumed this has biological significance about the mechanisms of disease, which may not be the case depending on how the model itself was composed. There is not always agreement between statistical and biological models of epistasis and significant statistical interactions do not imply biological or mechanistic interactions (Witte 2005).

There are many different molecular mechanisms that are theorized to cause epistasis. First are metabolic models, which describe epistasis between mutations that affect one target enzyme or a family of enzymes with similar structure and/or function (de Visser et al. 2011). Because enzymes typically function within a metabolic network/pathway, the interactions that characterize the network/pathway often dictate epistatic interactions (Szathmary 1993). In fact, models of yeast pathways have demonstrated mutations in successive steps of a pathway typically display redundant epistatic effects, whereas, mutations in unrelated pathways demonstrated varied effects congruent to the relatedness of the mutated pathways (Segre et al. 2005). Second is pleiotrophy, where mutations can be affecting multiple phenotypes (de Visser et al. 2011). Pleiotrophy is frequently demonstrated within proteins where one mutation may increase the activity rate of an enzyme but a subsequent mutation that stabilizes the same enzyme causes enzyme neutralization or loss of function (DePristo et al. 2005;

Wang et al. 2002). Third is the environment, where varying environmental conditions can allow for different and/or novel interactions amongst genetic loci (de Visser et al. 2011). In fact, the degree of environmental complexity is thought to influence the evolution of epistasis, and can be demonstrated by bacteria showing greater divergence when subjected to complex environments opposed to simple environments (Cooper and Lenski 2010; Korona et al. 1994; Rozen et al. 2008). Environmental effects on epistasis are most easily demonstrated where a mutation results in an inactive/truncated protein wherein alternative proteins/pathways compensate for the lost function (Crow and Kimura 1979). Fourth is robustness or the insensitivity of organisms to the effects of a mutation due to the diversity/complexity of the genome (Visser et al. 2003; Wagner 2013). If the genome consists of overlapping structures and functions, then any one mutation will have a lower impact since other sequences likely have the same or overlapping functions (Sanjuán and Elena 2006; Sanjuán and Nebot 2008). Lastly is evolvability, defined as the ability to adapt quickly to a varying environment, likely due to mutation/recombination (de Visser et al. 2011). High mutation rates and recombination can inadvertently select for genetic robustness facilitating the epistatic influences discussed with robustness (Azevedo et al. 2006; Visser et al. 2003; Wilke et al. 2001). However, recombination can cause a reduction in epistasis because recombination selects for 'generalist' mutations which benefit an array of genetic backgrounds opposed to selecting co-adapted complexes (Azevedo et al. 2006; Malmberg 1977). As previously discussed, variation in mastitis phenotype is heavily influenced by the environment indicating which could also be influencing the epistatic relationships within the genome. Thus, allowing for the inclusion of multiple traits in disease modeling will hopefully allow the identification of genetic variants that may otherwise remain undetected for complex disease traits such as mastitis.

Genome Wide Association Studies (GWAS)

Genetic selection has had limited success in reducing mastitis by selecting for more desirable udder characteristics, reduced SCS, and reduced clinical mastitis, genome likely due to low heritability and potential epistatic effects throughout the genome. As a result, other methods, including genome wide association studies (GWAS), have emerged to help improve overall mastitis phenotypes in dairy cows through genetics. GWAS attempt to identify associations with specific phenotypes and loci throughout the genome. Such associations could identify causal mutations related to the phenotype of interest and/or loci in linkage disequilibrium (LD) with causal mutations. Additionally, GWAS can identify candidate genes, genes that the associated SNPs are in or near, whose functions can be investigated further to identify their functional relationship to the associated phenotype.

Loci Associated with Mastitis Phenotypes

Many studies have utilized the GWAS approach to identify loci associated with mastitis phenotypes. Pighetti and Elliott (2011) reviewed much of the research completed before 2011, identifying 10 quantitative trait loci (QTLs) for SCC/SCS located in similar regions of the genome in more than one study. Since then, several more studies have been completed which have helped to confirm previously identified QTLs and identified novel loci associated with mastitis phenotypes. A QTL website (<http://www.animalgenome.org/cgi-bin/QTLdb/BT/index>) helps to collect the most recent information and is updated regularly (Hu et al. 2010). Since 2011, GWAS have discovered loci associated with mastitis phenotypes on every autosome of the bovine genome in a variety of breeds and populations. At least half of the loci discussed were previously identified themselves or overlap with previously identified regions in the same breed/population and/or new ones. Newer studies take advantage of advances in statistical methodology and new software releases to help improve prior results or narrow wide ranges with known associations.

One more recent GWAS utilized two separate populations of Holstein-Friesians, one of sires and one of cows, to identify associations between daughter yield deviations (DYD) for SCS, among other economic dairy traits, and the over 50,000 SNPs included in the medium density BeadChip from Illumina (Meredith et al. 2012). Both single SNP regression and Bayesian methods were utilized in analysis, where SNPs were included as continuous variables. Additionally, pedigree information for at least the prior four generations was included for each animal and accounted for by calculating average additive relationships between animals to reduce the likelihood of identifying false positives due to close animal pedigrees (Gilmour et al. 2009). SNP regression analyses included individual animal effects as random and multiple testing was accounted for by using false discovery rate reduction (Storey and Tibshirani 2003). The BayesB method for Bayesian testing was utilized but modified to account for the reliability of DYD phenotype information (Meuwissen et al. 2001). They identified 9 SNPs associated with SCS ($p < 0.05$) using a single SNP regression analysis approach including loci on BTAs 6, 10, 15, and 20. Three of the identified SNPs are located in previously identified QTL regions on BTAs 6 and 10 (Bennewitz et al. 2004; Daetwyler et al. 2008; Rupp and Boichard 2003). The remaining 6 loci were located in novel regions for SCS association with the strongest association on BTA 20. However, when a Bayesian analysis was used in place of a single SNP regression analysis to reduce false discoveries, only one SNP on BTA 4 was associated with SCS which had been previously identified by (Tal-Stein et al. 2010). The marked reduction in SNP associations with the Bayesian approach is speculated to be from the decreased reliability of the SCS proofs used to

develop the SCS DYD phenotypes. Decreased reliability of SCS phenotypes likely originates from less certainty in the true breeding value of the animal because environment plays such a heavy role in mastitis incidence for cows.

A second study by Meredith et al. (2013), again utilized Holstein-Friesians to test associations with SCS and clinical mastitis DYD, that were based on a larger number of progeny, to test associations with SNPs arrayed on Illumina's high density BeadChip with over 777,000 SNPs. P-values were corrected for multiple testing using the Benjamini-Hochberg procedure (Benjamini and Hochberg 1995) and QTL boundaries were determined based on pairwise linkage disequilibrium (LD), or the non-random association of loci due to close proximity, using Plink (Purcell et al. 2007). If the boundaries of two QTL overlapped, they are combined into one using the R/Bioconductor package "GenomicRanges" (Aboyoun et al. 2010; Gentleman et al. 2004). Post analysis, a total of 138 significant SNPs were identified across BTA 1, 3, 4, 5, 6, 9, 10, 13, 17, 20, 21, 22, 23, 24, 25, and 26. Eleven of the identified SNPs fell within known regions of SCS association, leaving 17 novel significant SNPs. In comparing the results between their two studies, Meredith et al. (2013) noted the SNPs identified using the high density SNP array (777,000 SNPs) were more strongly associated and had more defined QTL boundaries than those they identified previously with the medium density array (54,000 SNPs) (Meredith et al. 2012). In attempts to gain more information from the prior study, Meredith et al. (2013) performed imputation on the medium density genotypes with BEAGLE software (Browning and Browning 2009) using the high density array genotypes as a reference, but results proved unsuccessful with no consistency of imputation score across data sets and no significant associations being detected with the combined two data sets from both studies. Speculation was made that, either the reference array was of insufficient sample size (n=724) or the variation in the surrounding LD structure of significant SNPs caused the inconsistency of imputation (de Bakker et al. 2008; Spencer et al. 2009).

Another 2013 study by Sahana et al. (2013), performed a genome wide association analysis for mastitis traits on Nordic Holsteins using the medium density array (54,000 SNPs) from Illumina in attempts to confirm and refine the mapped position of QTLs identified previously using 356 microsatellite markers by Lund et al. (2008). Before running associations, data mining was completed to remove SNPs with call scores below 0.65 and minor allele frequencies below 0.05. Additionally, genotypes missing on the array were imputed using Fast-Phase software (Scheet and Stephens 2006). To determine associations, a mixed model was used where individual SNPs were added step-wise to the model as fixed effects and the polygenic genetic effect was a random effect (Yu et al. 2006). A Bonferroni correction was performed to reduce

family-wise error rate, but was found to be too conservative in correcting p-values (Han et al. 2009), so instead corrections to reduce false discovery rate as described by Benjamini & Hochberg (1995) were applied to help reduce the instance of false positives but not be overly conservative. A total of 143 significant SNP associations with clinical mastitis phenotypes were identified with the majority on BTAs 6, 14, 16, and 20. Overall, 6 of the identified SNPs (BTA 5, 9, and 24) were within the regions previously identified as associated with mastitis QTLs, despite the definitions of mastitis used differing slightly across studies (Lund et al. 2008; Lund et al. 2007; Nilsen et al. 2009; Sahana et al. 2008; Sodeland et al. 2011).

As a follow up to their prior work, Sahana et al. (2014) performed a targeted association study attempting to fine-map QTL regions previously identified in Nordic Holsteins for clinical mastitis and SCC (Lund et al. 2008; Lund et al. 2007; Sahana et al. 2013) using sequence level variants and confirm if the same QTL regions existed in Danish Jerseys or if the two breed populations segregated. Estimated breeding values from artificial insemination sires for clinical mastitis and SCC traits were associated with SNPs on a medium density SNP array (54,000 SNPs) from Illumina. Extensive data mining was performed again pre-association analysis to remove low quality base calls, keep minor allele frequencies below 5%, and prevent deviation from Hardy-Weinberg equilibrium, to increase power and reduce false discovery rate. Like other studies, imputation was also used, but here it was targeted to 9 mastitis and SCC traits in Nordic Holsteins located on BTAs 5, 6, 13, 16, 19, and 20 using high density array (777,000 SNPs) information. Only SNPs with greater than 0.9 imputation certainty were kept to perform associations. Following association analysis with Bonferroni multiple testing corrections, a total of 6 wide-peaked regions of association located on the 6 BTA that were targeted. In conjunction with regular association tests, association tests with the most important SNP as a cofactor in the model were also run. These cofactor analyses resulted in a collapse of the wide peaks of association on all chromosomes, likely indicating the single genetic variant within the targeted region, except for on BTA 5. The collapsed peak for each chromosome association likely indicates the causative single genetic variant for each region, but extensive LD does not allow for eliminating the possibility of multiple causal variants for the regions. Additionally, loci on BTA 16 and 20 previously identified in Nordic Holsteins were confirmed present in Danish Jerseys.

Tiezzi et al. (2015) also attempted to identify loci associated with mastitis using GWAS. They utilized producer reported instance of clinical mastitis on N=103,585 Holsteins associated with their sire's genotypes using a BovineSNP50 BeadChip from Illumina. However, they employed a single-step genomic BLUP method that

incorporated both pedigree and genomics of each animal into a matrix that weighted each marker's contribution by the variance of that marker to better account for relatedness of the animals and the genetic variance demonstrated by the population measured. Additionally, the binary nature of the clinical mastitis phenotype was accommodated using a threshold-liability model. This analysis identified 10 genomic regions that explained significant portions of the variation observed in clinical mastitis incidence which were located on BTAs 2, 8, 11, 14, 16, 19, 20, and 29, with the most significance being attributed to the regions on BTA 2, 14, and 20. While other studies have identified loci associated with clinical mastitis on these chromosomes (Klungland et al. 2001; Sodeland et al. 2011; Youngerman et al. 2004), only the region identified on BTA 20 (32.17-61.61) was also identified using a clinical mastitis phenotype by another study (Sodeland et al. 2011).

Lastly, Biffani et al. (2016) took a novel approach by using a case-control design to identify loci associated with mastitis. Their approach looked at Italian Holstein cows that were repeatedly screened for mastitis over time (N=550 with mastitis cases, N=571 without – controls) and compared them to the results of a prior GWAS on SCC breeding values of Italian Holstein sires. The comparison focused on 9 previously identified SNPs that are located on BTA 6 and 14 (Minozzi et al. 2011) and was completed using a single-SNP regression model in the 'qtscore' function in the GenABEL package (Aulchenko et al. 2007) in R. The results agreed with the prior studies finding 2 of the SNPs on BTA 6 and 4 of the SNPs on BTA 14 significantly associated with mastitis incidence which validated using historical case-control data for GWAS.

These recent studies focused on confirming and narrowing the loci previously identified in mastitis trait association studies. Overall, results seemed to favor loci located on BTA 6, 14, 16, and 20, but loci were identified on every chromosome of the genome. Bayesian based approaches seem to provide great benefit in reducing false discoveries, although studies correcting for family wise error rate through pedigrees met with less success than those correcting for false discovery rates, since the latter is less abruptly conservative and does not reduce power as heavily as the former. Studies utilizing imputation showed varying success rates, likely the result of varying sample sizes and inconsistent LD structure across the genome. Lastly, higher density SNP arrays provided more defined QTL boundaries than those obtained with medium density arrays because closer SNPs allow better definition of the local LD structure. The general trend for GWAS approach seems to focus on increasing analysis performance through statistics, increased sample size, and increased SNP array density, with little focus on the phenotype. Thus, exploring the phenotypes used and the conditions they are generated under could provide important information in the context of GWAS.

Candidate Genes Identified by Loci Associated with Mastitis Phenotypes

While numerous genome loci have been associated with mastitis phenotypes using GWAS, the potential candidate genes surrounding these loci are even more numerous. Functions of potential candidate genes can be implicated in many aspects of mastitis progression and are summarized below. Of note, the previous review by Pighetti and Elliott (2011) covered a range of candidate genes including those for pathogen recognition, immune cell migration, bacterial elimination, and resolution of infection.

The innate immune system is composed of non-specific mechanisms that are initiated within hours or immediately upon antigen recognition, including things like anti-microbial peptides. Several of the loci associated with mastitis phenotypes have provided candidate genes that fall within the functions of innate immunity. Within a QTL located on BTA 6, is the GC gene that codes for Gc-globulin. Gc-globulin leads to cathelicidin antimicrobial peptide synthesis by delivering vitamin D needed for synthesis, to target monocytes and also helps to enhance macrophage phagocytosis (Hewison 2011; Yamamoto and Homma 1991). Regions on BTA 13 associated with SCS include genes for β -defensins such as: DEFB117, DEFB119, DEFB122, DEFB122A, DEFB129 and 2 precursor genes (Cormican et al. 2008). Another QTL region on BTA 17 is in close proximity to SDF2L1 (stromal cell-derived factor 2-like protein 1) gene which plays a role in processing many β -defensins (Zheng et al. 2012). β -defensins are a family of anti-microbial peptides that are responsible for initiating the immune response and preventing microbial colonization of epithelial surfaces which are known to be expressed in mammary tissue (Ganz 2003; Goldammer et al. 2004; Swanson et al. 2004). The gene for C9 is located near a QTL region identified on BTA 20 and codes for the precursor for the pore-forming unit of the complement membrane attack complex. The complement complex aids in the clearance of microorganisms by introducing pores into microbe membranes and mutations in the C9 gene have been linked to increases susceptibility to some infections (Horiuchi et al. 1998; Kira et al. 1998; Ricklin et al. 2010; Zoppi et al. 1990). Also nearby to the identified BTA 20 region is the LIFR gene which produces leukemia inhibitory factor receptor, subunit α . Once the receptor binds leukemia inhibitory factor its effects include modulation of the inflammatory response and stimulation of hepatocytes to produce acute-phase proteins (Baumann and Wong 1989; Linker et al. 2008). Each of these functions aid in the clearance of invading bacteria and could be implicated in mastitis incidence.

Once bacteria are recognized by the immune system, the recruitment of immune cells to the site of infection to aid in infection resolution begins. Identified regions on BTA 1, 13, 17, and 20 are close to genes with immune cell recruitment implications. An

identified QTL on BTA 1 is located within the gene for CD166 antigen or ALCAM. CD166 antigen is heavily implicated in leukocyte adhesion, migration and the activation of T-cells in human cancers and shows ties to mastitis resistance in challenged somatic cells from sheep (Bonnetfont et al. 2011; Cayrol et al. 2008; Hassan et al. 2004; Zimmerman et al. 2006). The gene for EDN3 (endothelin 3) is located near a QTL region on BTA 13. EDN3, expressed by several cell types including macrophages, has vasoconstrictive effects and potentially activates neutrophils needed to phagocytose microorganisms (Elferink and De Koster 1998; Giaid et al. 1991; Selsted and Ouellette 2005). Also near the QTL on BTA 13 is Phactr 3 (Phosphatase and actin regulator 3) which interacts with cell cytoskeleton to facilitate cell migration and spreading essential for mobilizing immune cells (Luster et al. 2005; Sagara et al. 2009). The IL31R α gene, which acts in the IL-31 receptor complex, is located near a BTA 20 QTL. Upon binding cytokine, the IL-31 receptor complex activates JAK-STAT and MAPK pathways which in turn initiate a several other immune processes (Cornelissen et al. 2012; Zhang et al. 2008). Additionally, the gene for MAPK1, which is a member of the MAPK family and is essential for leukocyte recruitment following bacteria invasion, is located near a QTL on BTA 17 (Herlaar and Brown 1999; Kaminska 2005). Deficiencies in the recruitment of immune cells to the site of infection could allow for increased incidence of mastitis or chronic infections develop.

In addition to response genes above, other genes related to T-cells, B-cells, and antibodies are potential candidate genes. A region identified on BTA 6, has its most significantly associated SNP within the DCK (deoxycytidine kinase) gene. The DCK gene catalyzes the rate limiting step deoxyribonucleoside salvage and results in low production of T- and B-cells when knocked out in mice linking its function to lymphopoiesis (Toy et al. 2010). Also in this region on BTA 6, is the gene for IGJ (immunoglobulin J polypeptide), which aids assembly of IgM pentamers and IgA dimers. Both IgM and IgA play critical roles in defending hosts from microorganisms, with IgM being the first antibody produced after microbial recognition and IgA adhering to microbes on mucosal surfaces to help prevent adherence to epithelial cells (Lamm 1997; Racine and Winslow 2009). Also on BTA 6, associated with another region, is the candidate gene DAPP1 or Bam32, which has suspected roles in B-cell receptor signaling (Pierce 2002). B-cell receptors, upon binding antigen, help designate B-cell fates like proliferation, differentiation, receptor internalization, and T-cell adhesion and subsequent activation (Al-Alwan et al. 2010; Han et al. 2003; Niirio et al. 2004). Additionally, the gene for NPFFR2, a member of the NPFF RFamide neuropeptides, is located in this region. These neuropeptides decrease macrophage nitric oxide production producing anti-inflammatory activity in mouse models and can influence T-cell proliferation (Jhamandas and Goncharuk 2014; Lecron et al. 1992; Minault et al.

1995; Sun et al. 2013). Lastly, on BTA 19, an identified QTL region is close to the mTORC1 (mTOR complex 1) gene which influences T-cell functions (Zeng et al. 2013). Functions above involved in specific antigen responses have the potential to alter response to mastitis pathogens.

A few candidate genes were identified based on changes in expression level following challenge with mastitis pathogens or their components. Both the IGLL-1 (lambda-like polypeptide 1 precursor) and the SNAP29 (synaptosomal-associated protein 29) genes are located near QTL identified on BTA 17 and have demonstrated up-regulated expression following bacterial challenge in different models (Swanson et al. 2009; Wesolowski et al. 2012). Additionally, a micro-RNA, mir301b, has demonstrated down regulation following lipopolysaccharide (LPS) challenge. It is located near the same QTL region on BTA 17 (Zheng et al. 2012). While it is unclear the actual function these genes may play in mastitis, their changes in expression level following challenge suggest they are influenced by pathogenic bacteria presence. Perhaps such genes could be used in PCR based diagnostic techniques if their expression changes are consistent across pathogens.

Lastly, some loci on BTA 5, 6, 13, 20, associated with mastitis phenotypes are also associated with milk phenotypes, particularly with decreased milk yield (Grisart et al. 2002; Meredith et al. 2012; Sahana et al. 2014; Sahana et al. 2013). The concurrent association with mastitis and milk yield traits suggests the effects of the detected genes are pleiotropic and support the widely accepted correlation between SCS and decreased milk production (Carlén et al. 2004). Within these identified regions for milk and mastitis traits there are genes that could be implicated in pathways for milk synthesis. The region located on BTA 5 is within the intron of the MGST1 (microsomal glutathione S-transferase 1) gene. MGST1 is localized to sites of lipid synthesis within the cell, thought to protect against oxidative stress, and can catalyze lipid hydroperoxides making it an excellent candidate for a milk fat QTL (Johansson et al. 2010; Mosialou et al. 1995). Close to the region identified on BTA 6 are several Casein related proteins. Caseins are the primary protein component of bovine milk, and a QTL for milk protein exists in this region while not actually mapping to within one of the casein related genes, thus indicating the potential for multiple causative polymorphisms (Sahana et al. 2014).

The range of candidate genes discussed above, while far from all-inclusive, demonstrates QTL associated with mastitis phenotypes are numerous and include a variety of functions. Trends in identified genes are: antimicrobial peptides, immune cell recruitment and antigen specific immunity. General functions follow trends observed from the prior review of candidate genes with some being identified in both (Pighetti and

Elliott 2011). Despite the wide range of genes identified utilizing older and newer GWAS and QTL analyses, there has been little translation to on farm selection practices.

In summary, many regions of the bovine genome have been identified as associated with mastitis phenotypes using a variety of original and revised GWAS analyses. Many studies have identified similar and/or overlapping regions of association. Newer technologies strive to improve the results of GWAS analyses by reducing the probability of false discoveries and narrowing the regions of association to causative variants. These identified regions have provided any number of candidate genes with potential to influence a cow's response to mastitis. Generalized functions of candidate genes include bacterial recognition, recruitment of immune cells, and antigen specific immune responses. However, little conversion of these associated loci is observed to commonly used selection parameters on farm. Frequently, the easiest and least expensive way to acquire access to sample sizes adequate for GWAS analyses in the dairy industry is to rely on public data bases or government records for phenotype information. This is largely because acquiring the resources and financial support to house, milk, and collect data on a large population of dairy cows is not feasible for one researcher or even a team of researchers. As a result, phenotypes utilized in dairy GWAS are frequently derived from daughter yield deviations (DYD) or estimated breeding values (EBV). Both EBVs and DYDs are estimates or predictions of a sire's progeny based on averages of all the bull's existing, mature offspring, and do not represent that individual cow's performance directly. Thus, there is little ability to account for the effect of environment or the proportion of genetics received from the dam, both of which may cause fluctuation in cows' mastitis phenotype. Furthermore, Pszczola and Calus (2015) determined that databases used to determine DYD or EBV need to be updated with a minimum of 600 individuals per generation to remain reliable in their ability to predict, which can become costly depending on the phenotype in question. These discrepancies, could account for a portion of the reason why a great deal of research is being completed in identifying regions, but little evidence of translation to on farm use of selection for or against discovered loci is observed. Perhaps developing novel phenotypes from data collected on smaller sample sizes in response to direct challenge with common mastitis pathogens could narrow the field of identified associations. Additionally, identified loci would be associated to the pathology of the challenge organism, leading to better candidate genes increasing the likelihood of developing novel treatments, preventatives or vaccinations specific to pathogens.

References

- Akers, R. M., and Nickerson, S. C. (2011). "Mastitis and its impact on structure and function in the ruminant mammary gland." *Journal of Mammary Gland Biology and Neoplasia*, 16(4), 275-289.
- Aulchenko, Y. S., Ripke, S., Isaacs, A., and Van Duijn, C. M. (2007). "GenABEL: an R library for genome-wide association analysis." *Bioinformatics*, 23(10), 1294-1296.
- Azevedo, R. B. R., Lohaus, R., Srinivasan, S., Dang, K. K., and Burch, C. L. (2006). "Sexual reproduction selects for robustness and negative epistasis in artificial gene networks." *Nature*, 440(7080), 87-90.
- Barbano, D. M., Rasmussen, R. R., and Lynch, J. M. (1991). "Influence of milk somatic cell count and milk age on cheese yield." *Journal of Dairy Science*, 74(2), 369-388.
- Bateson, W., and Mendel, G. (1909). *Mendel's principles of heredity*: Courier Corporation.
- Berry, D. P., Buckley, F., Dillon, P., Evans, R. D., Rath, M., and Veerkamp, R. F. (2003). "Genetic parameters for body condition score, body weight, milk yield, and fertility estimated using random regression models." *Journal of Dairy Science*, 86(11), 3704-3717.
- Biffani, S., Del Corvo, M., Capoferri, R., Pedretti, A., Luini, M., Williams, J. L., Pagnacco, G., Minvielle, F., and Minozzi, G. (2016). "An alternative experimental case-control design for genetic association studies on bovine mastitis." *Animal*, 1-6.
- Bradley, A. J., Leach, K. A., Breen, J. E., Green, L. E., and Green, M. J. (2007). "Survey of the incidence and aetiology of mastitis on dairy farms in England and Wales." *The Veterinary Record*, 160, 253-258.
- Campbell, J. R., and Marshall, R. T. (1975). *Dairy production and processing: The science of milk and milk products*: McGraw Hill Book Company.
- Cooper, T. F., and Lenski, R. E. (2010). "Experimental evolution with *E. coli* in diverse resource environments. I. Fluctuating environments promote divergence of replicate populations." *BMC Evolutionary Biology*, 10(1), 1.
- Crow, J. F., and Kimura, M. (1979). "Efficiency of truncation selection." *Proceedings of the National Academy of Sciences*, 76(1), 396-399.

- de Visser, J. A. G. M., Cooper, T. F., and Elena, S. F. (2011). "The causes of epistasis." *Proceedings of the Royal Society of London B: Biological Sciences*, 278(1725), 3617-3624.
- DePristo, M. A., Weinreich, D. M., and Hartl, D. L. (2005). "Missense meanderings in sequence space: a biophysical view of protein evolution." *Nature Reviews Genetics*, 6(9), 678-687.
- Dohoo, I. R., Smith, J., Andersen, S., Kelton, D. F., and Godden, S. (2011). "Diagnosing intramammary infections: Evaluation of definitions based on a single milk sample." *Journal of Dairy Science*, 94(1), 250-261.
- Falconer, D. S. (1975). *Introduction to quantitative genetics*: Pearson Education India.
- Fetrow, J. "Mastitis: An economic consideration." *Presented at Council, N.M. (ed.), 39th Annual Meeting of the National Mastitis Council*.
- Fisher, R. A. (1919). "The Correlation between Relatives on the Supposition of Mendelian Inheritance." *Earth and Environmental Science Transactions of the Royal Society of Edinburgh*, 52(02), 399-433.
- Heringstad, B., Rekaya, R., Gianola, D., Klemetsdal, G., and Weigel, K. A. (2003). "Genetic change for clinical mastitis in Norwegian cattle: A threshold model analysis." *Journal of Dairy Science*, 86(1), 369-375.
- Hogan, J. S., and Smith, K. L. (1987). "A practical look at environmental mastitis." *Compendium on Continuing Education for the Practicing Veterinarian*, 9(10).
- Hogan, J. S., Smith, K. L., Hoblet, K. H., Todhunter, D. A., Schoenberger, P. S., Hueston, W. D., Pritchard, D. E., Bowman, G. L., Heider, L. E., Brockett, B. L., and et al. (1989). "Bacterial counts in bedding materials used on nine commercial dairies." *Journal of Dairy Science*, 72(1), 250-258.
- Holstein Association USA. (2013). "How to read Holstein sire information". City: Holstein Association USA, Inc.
- Jacquard, A. (1983). "Heritability: One word, three concepts." *Biometrics*, 39(2), 465-477.
- Jayarao, B. M., Gillespie, B. E., Lewis, M. J., Dowlen, H. H., and Oliver, S. P. (1999). "Epidemiology of *Streptococcus uberis* intramammary infections in a dairy herd." *Journal of Veterinary Medicine, Series B*, 46(7), 433-442.
- Jones, W. P., Hansen, L. B., and Chester-Jones, H. (1994). "Response of health care to selection for milk yield of dairy cattle." *Journal of Dairy Science*, 77(10), 3137-3152.

- Klei, L., Yun, J., Sapru, A., Lynch, J., Barbano, D., Sears, P., and Galton, D. (1998). "Effects of milk somatic cell count on cottage cheese yield and quality." *Journal of Dairy Science*, 81(5), 1205-1213.
- Klungland, H., Sabry, A., Heringstad, B., Olsen, H. G., Gomez-Raya, L., Våge, D. I., Olsaker, I., Ødegård, J., Klemetsdal, G., Schulman, N., Vilkkki, J., Ruane, J., Aasland, M., Rønningen, K., and Lien, S. (2001). "Quantitative trait loci affecting clinical mastitis and somatic cell count in dairy cattle." *Mammalian Genome*, 12(11), 837-842.
- Korona, R., Nakatsu, C. H., Forney, L. J., and Lenski, R. E. (1994). "Evidence for multiple adaptive peaks from populations of bacteria evolving in a structured habitat." *Proceedings of the National Academy of Sciences*, 91(19), 9037-9041.
- Lund, T., Miglior, F., Dekkers, J. C. M., and Burnside, E. B. (1994). "Genetic relationships between clinical mastitis, somatic cell count, and udder conformation in Danish Holsteins." *Livestock Production Science*, 39(3), 243-251.
- Lush, J. L. (1940). "Intra-sire correlations or regressions of offspring on dam as a method of estimating heritability of characteristics." *Proceedings of the American Society of Animal Nutrition*, 1940(1), 293-301.
- Lush, J. L. (1949). "Heritability of quantitative characters in farm animals." *Hereditas*, 35(S1), 356-375.
- Lynch, M., and Walsh, B. (1998). *Genetics and analysis of quantitative traits*: Sinauer Sunderland, MA.
- Ma, Y., Ryan, C., Barbano, D. M., Galton, D. M., Rudan, M. A., and Boor, K. J. (2000). "Effects of somatic cell count on quality and shelf-life of pasteurized fluid milk." *Journal of Dairy Science*, 83(2), 264-274.
- Malmberg, R. L. (1977). "The evolution of epistasis and the advantage of recombination in populations of bacteriophage T4." *Genetics*, 86(3), 607-621.
- Minozzi, G., Nicolazzi, E. L., Strozzi, F., Stella, A., Negrini, R., Ajmone-Marsan, P., and Williams, J. L. "Genome wide scan for somatic cell counts in Holstein bulls." *Presented at BMC Proceedings*.
- Neuman, R. J., Rice, J. P., and Chakravarti, A. (1992). "Two-Locus models of disease." *Genetic Epidemiology*, 9(5), 347-365.
- NMC. (1999). *Laboratory handbook on bovine mastitis*: NMC.
- NMC. (2005). "The cost of mastitis: Dairy insight research 2005/2006: Final report." *National Mastitis Advisory Committee*.

- Oliver, S. P. (1988). "Frequency of isolation of environmental mastitis-causing pathogens and incidence of new intramammary infection during the nonlactating period." *American Journal of Veterinary Research*, 49(11), 1789-93.
- Oliver, S. P., and Mitchell, B. A. (1983). "Susceptibility of bovine mammary gland to infections during the dry period." *Journal of Dairy Science*, 66(5), 1162-1166.
- Paape, M., Mehrzad, J., Zhao, X., Detilleux, J., and Burvenich, C. (2002a). "Defense of the bovine mammary gland by polymorphonuclear neutrophil leukocytes." *Journal of Mammary Gland Biology and Neoplasia*, 7(2), 109-21.
- Paape, M. J., Bannerman, D. D., Zhao, X., and Lee, J.-W. (2003). "The bovine neutrophil: Structure and function in blood and milk." *Veterinary Research*, 34(5), 597-627.
- Paape, M. J., Shafer-Weaver, K., Capuco, A. V., Oostveldt, K., and Burvenich, C. (2002b). "Immune surveillance of mammary tissue by phagocytic cells", in J. A. Mol and R. A. Clegg, (eds.), *Biology of the Mammary Gland*. Springer US, pp. 259-277.
- Patterson, H. D., and Thompson, R. (1971). "Recovery of inter-block information when block sizes are unequal." *Biometrika*, 58(3), 545-554.
- Petrovski, K. R., Williamson, N. B., Lopez-Villalobos, N., Parkinson, T. J., and Tucker, I. G. (2011). "Culture results from milk samples submitted to veterinary diagnostic laboratories from August 2003 to December 2006 in New Zealand." *New Zealand Veterinary Journal*, 59(6), 317-322.
- Philipsson, J., Ral, G., and Berglund, B. (1995). "Somatic cell count as a selection criterion for mastitis resistance in dairy cattle." *Livestock Production Science*, 41(3), 195-200.
- Phuektes, P., Mansell, P. D., Dyson, R. S., Hooper, N. D., Dick, J. S., and Browning, G. F. (2001). "Molecular epidemiology of *Streptococcus uberis* isolates from dairy cows with mastitis." *Journal of Clinical Microbiology*, 39(4), 1460-1466.
- Quaas, R. L., and Pollak, E. J. (1980). "Mixed model methodology for farm and ranch beef cattle testing programs." *Journal of Animal Science*, 51(6), 1277-1287.
- Riekerink, R. G. M. O., Sampimon, O. C., Eerland, V. J., Swarts, M. J., and Lam, T. J. G. M. "Comparing bacterial counts on bare hands with gloved hands during milking." *Presented at Mastitis control: From science to practice. Proceedings of International Conference, The Hague, Netherlands, 30 September-2 October 2008*.

- Rozen, D. E., Habets, M. G. J. L., Handel, A., de Visser, J., and Arjan, G. M. (2008). "Heterogeneous adaptive trajectories of small populations on complex fitness landscapes." *PLoS One*, 3(3), e1715.
- Ruegg, P. L. (2003). "Investigation of mastitis problems on farms." *Veterinary Clinics of North America: Food Animal Practice*, 19(1), 47-73.
- Rupp, R., and Boichard, D. (1999). "Genetic parameters for clinical mastitis, somatic cell score, production, udder type traits, and milking ease in first lactation Holsteins." *Journal of Dairy Science*, 82(10), 2198-2204.
- Sanjuán, R., and Elena, S. F. (2006). "Epistasis correlates to genomic complexity." *Proceedings of the National Academy of Sciences*, 103(39), 14402-14405.
- Sanjuán, R., and Nebot, M. R. (2008). "A network model for the correlation between epistasis and genomic complexity." *PLoS One*, 3(7), e2663.
- Schepers, A. J., Lam, T. J., Schukken, Y. H., Wilmink, J. B., and Hanekamp, W. J. (1997). "Estimation of variance components for somatic cell counts to determine thresholds for uninfected quarters." *Journal of Dairy Science*, 80(8), 1833-40.
- Schukken, Y. H., Wilson, D. J., Welcome, F., Garrison-Tikofsky, L., and Gonzalez, R. N. (2003). "Monitoring udder health and milk quality using somatic cell counts." *Veterinary Research*, 34(5), 579-596.
- Segre, D., DeLuna, A., Church, G. M., and Kishony, R. (2005). "Modular epistasis in yeast metabolism." *Nature Genetics*, 37(1), 77-83.
- Siemiatycki, J., and Thomas, D. C. (1981). "Biological models and statistical interactions: An example from multistage carcinogenesis." *International Journal of Epidemiology*, 10(4), 383-387.
- Smith, K. L., and Hogan, J. S. (1993). "Environmental mastitis." *Veterinary Clinics of North America: Food Animal Practice*, 9(3), 489-98.
- Sodeland, M., Kent, M. P., Olsen, H. G., Opsal, M. A., Svendsen, M., Sehested, E., Hayes, B. J., and Lien, S. (2011). "Quantitative trait loci for clinical mastitis on chromosomes 2, 6, 14 and 20 in Norwegian Red cattle." *Animal Genetics*, 42(5), 457-465.
- Soller, M., and Beckmann, J. S. (1983). "Genetic polymorphism in varietal identification and genetic improvement." *Theoretical and Applied Genetics*, 67(1), 25-33.
- Sorensen, D., and Gianola, D. (2007). *Likelihood, Bayesian, and MCMC methods in quantitative genetics*: Springer Science & Business Media.

- Sørensen, M. K., Jensen, J., and Christensen, L. G. (2000). "Udder conformation and mastitis resistance in danish first-lactation cows: Heritabilities, genetic and environmental corelations." *Acta Agric Scand (A)*, 50(2), 72-82.
- Szathmary, E. (1993). "Do deleterious mutations act synergistically? Metabolic control theory provides a partial answer." *Genetics*, 133(1), 127-132.
- The Holstein Foundation. (2015). "Understanding genetics and the sire summaries". City: The Holstein Foundation.
- Tiezzi, F., Parker-Gaddis, K. L., Cole, J. B., Clay, J. S., and Maltecca, C. (2015). "A genome-wide association study for clinical mastitis in first parity US Holstein cows using single-step approach and genomic matrix re-weighting procedure." *PLOS ONE*, 10(2), e0114919.
- VanRaden, P. M., Van Tassell, C. P., Wiggans, G. R., Sonstegard, T. S., Schnabel, R. D., Taylor, J. F., and Schenkel, F. S. (2009). "Reliability of genomic predictions for North American Holstein bulls." *Journal of Dairy Science*, 92(1), 16-24.
- Verbeke, J., Piepers, S., Supré, K., and De Vliegher, S. (2014). "Pathogen-specific incidence rate of clinical mastitis in Flemish dairy herds, severity, and association with herd hygiene." *Journal of Dairy Science*, 97(11), 6926-6934.
- Visscher, P. M., Hill, W. G., and Wray, N. R. (2008). "Heritability in the genomics era- Concepts and misconceptions." *Nature Reviews Genetics*, 9(4), 255-266.
- Visser, J. A. G. M., Hermisson, J., Wagner, G. P., Meyers, L. A., Bagheri-Chaichian, H., Blanchard, J. L., Chao, L., Cheverud, J. M., Elena, S. F., and Fontana, W. (2003). "Perspective: Evolution and detection of genetic robustness." *Evolution*, 57(9), 1959-1972.
- Wagner, A. (2013). *Robustness and evolvability in living systems*: Princeton University Press.
- Wang, X., Minasov, G., and Shoichet, B. K. (2002). "Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs." *Journal of Molecular Biology*, 320(1), 85-95.
- Wattiaux, M. A. (1995). "Predicted transmitting ability and reliability." *Dairy Essentials*, 16, 61-64.
- Wiggans, G. R., VanRaden, P. M., and Cooper, T. A. (2011). "The genomic evaluation system in the United States: Past, present, future." *Journal of Dairy Science*, 94(6), 3202-3211.

- Wilke, C. O., Wang, J. L., Ofria, C., Lenski, R. E., and Adami, C. (2001). "Evolution of digital organisms at high mutation rates leads to survival of the flattest." *Nature*, 412(6844), 331-333.
- Wilson, D. J., Gonzalez, R. N., and Das, H. H. (1997). "Bovine mastitis pathogens in New York and Pennsylvania: Prevalence and effects on somatic cell count and milk production." *Journal of Dairy Science*, 80(10), 2592-2598.
- Witte, J. S. (2005). "Gene-environment interaction", *Encyclopedia of Biostatistics*. John Wiley & Sons, Ltd.
- Wright, S. (1920). "The relative importance of heredity and environment in determining the piebald pattern of guinea-pigs." *Proceedings of the National Academy of Sciences*, 6(6), 320-332.
- Youngerman, S. M., Saxton, A. M., Oliver, S. P., and Pighetti, G. M. (2004). "Association of CXCR2 polymorphisms with subclinical and clinical mastitis in dairy cattle." *Journal of Dairy Science*, 87(8), 2442-2448.
- Zadoks, R. N. (2007). "Sources and epidemiology of *Streptococcus uberis*, with special emphasis on mastitis in dairy cattle." *Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*, 2(30), 15.

**Chapter III: Genome wide association study identifies loci associated with
somatic cell count phenotypes following experimental challenge with
*Streptococcus uberis***

Genome wide association study identifies loci associated with somatic cell count phenotypes following experimental challenge with *Streptococcus uberis*

L. Siebert¹, M.E. Staton², S.P. Oliver^{1, 3}, and G.M. Pighetti¹

¹The University of Tennessee, Knoxville, TN, Department of Animal Science

²The University of Tennessee, Knoxville, TN, Department of Entomology and Plant Pathology

³The University of Tennessee, Knoxville, TN, AgResearch

Abstract

Mastitis is a detrimental disease in the dairy industry that costs upwards of \$2 billion annually and decreases milk quality. Often, mastitis results from bacteria entering the gland through the teat end. A common mastitis causing pathogen is *S. uberis*, which is responsible for a high percentage of subclinical and clinical mastitis cases. Following an intramammary experimental challenge with *S. uberis* on Holstein cows (n=40), milk samples were collected and somatic cell counts (SCC) were determined by the Dairy Herd Improvement Association Laboratory. Traditional GWAS have utilized test day SCC, SCC over an entire lactation, or sire EPD/DYD SCC to identify loci of interest for mastitis. Our approach utilizes SCC collected following direct *S. uberis* experimental challenge to generate three novel phenotypes: area under the curve (AUC) of SCC for 0-7 d and 0-28 d post-challenge; and individual cows were placed in one of three categories (<21 d, 21-28 d, or >28 d) based on when their SCC returned to below 200,000 cells/ml post-challenge. To identify loci of interest a 50K SNP chip analysis was performed using the BovineSNP50 v2 DNA BeadChip from Illumina. Associations were tested using Plink and identified 16 significant ($p < 9.34 \times 10^{-5}$) SNPs across the 3 phenotypes. Five of the identified SNPs are located within previously identified quantitative trait loci (QTLs) for traits with links to our generated phenotypes. Additionally, six identified SNPs are located within known genes whose functions can be linked to SCC. The identified loci should be further investigated to potentially identify causation behind the observed phenotypes. Such investigations could lead to novel treatment and prevention compounds/protocols for *S. uberis* mastitis or genetic selection methods for cows with greater potential to resist *S. uberis* infection.

Introduction

Mastitis, the inflammation of the mammary gland, decreases milk quality and farm revenues generated. Data shows mastitis can affect 20-50% of a herd causing an increase in cow morbidity and mortality (Bradley et al. 2007; Østerås et al. 2006; Piepers et al. 2007; Wilson et al. 1997). Due to the prevalence and impact of mastitis,

increased preventative and control methods are a long standing goal of the dairy industry. Mastitis frequently results from pathogenic bacteria invading the mammary gland. Common mastitis causing pathogens include *Escherichia coli*, *Streptococcus uberis*, and *Staphylococcus aureus*, where each pathogen has its own unique pathology (Schukken et al. 2011). *S. uberis* is responsible for a high percentage of both clinical and subclinical mastitis and is commonly isolated during the dry (non-lactating) period (Bradley et al. 2007; Petrovski et al. 2011; Riekerink et al. 2008; Verbeke et al. 2014). Due to *S. uberis*'s prevalence and its classification as both an environmental and contagious pathogen, it is often used as a model organism in challenge studies (Zadoks 2007). Once bacteria invade the mammary gland, epithelial cells and resident macrophages release cytokines which recruit polymorphonuclear neutrophils (PMN) and other leukocytes to the site of infection (Paape et al. 2003). This cytokine recruitment results in a large influx of leukocytes into the milk, where leukocytes are commonly referred to as somatic cells. Somatic cell counts (SCC), represented as cells/mL, are routinely monitored by the majority of dairy producers and are used as an indicator for the level of intramammary infection or mastitis. There are some resident leukocyte populations in the mammary gland and/or milk (i.e.: macrophages) meaning, even when pathogenic bacteria are not present in the gland, SCC is greater than zero. Industry standard considers a SCC below 200,000 cells/mL a healthy gland and SCC above 200,000 cells/mL to indicate either subclinical or clinical mastitis (Dohoo et al. 2011; Schepers et al. 1997; Schukken et al. 2003). Subclinical mastitis is indicated by an increase in SCC, whereas clinical mastitis is represented by an increase in SCC and the presence of clinical signs (i.e.: inflammation, altered milk production). Additionally, SCC are often converted to somatic cell scores (SCS), which are the log transformation of SCC.

Traditional selection for cows with higher milk production has inadvertently selected for cows that are also more susceptible to diseases including mastitis (Rupp and Boichard 2003). Existing mastitis/SCC heritability calculations, ranging from 0.05 to 0.27, are based on test day measurements or measurements over an entire lactation, which do not account for variation in SCC based on causative pathogen (Haile-Mariam et al. 2003; Kennedy et al. 1982; Weller et al. 1992). Additionally, the low SCC heritability means the selection response is proportionally small and gene interaction(s) and/or environmental factor(s) likely influence SCC heritability. The advent of high throughput genotyping arrays has allowed researchers to perform genome wide association studies (GWAS) to identify specific loci or single nucleotide polymorphisms (SNPs) within the genome that are associated with SCC or related phenotypes instead of relying solely on traditional bull proofs, which are based on observed phenotype selection. The results obtained from GWAS are primarily influenced by two factors: the

coverage of the array (i.e.: number of SNPs genotyped) and the number of individuals genotyped, amongst other smaller factors (Spencer et al. 2009). The GWAS approach relies heavily on the SNPs included in the array being in linkage disequilibrium (LD), the non-random association of alleles at different loci, with the causal variant. The greater the LD between the SNP and the causal variant, the greater the power to detect the association (Spencer et al. 2009). Genome arrays are available in a variety of coverage density, which refers to the number of SNPs genotyped, from low density (7,000 SNPs) to high density (780,000 SNPs). Genotyping a larger number of SNPs by using a higher density array can increase the power of GWAS results, however, due to the high degree of LD present in the Holstein breed, a medium density array (54,000 SNPs) should provide sufficient detection power (de Roos et al. 2008).

Recent mastitis GWAS have utilized medium density arrays to detect associations successfully (Cole et al. 2009; Meredith et al. 2012; Pryce et al. 2010). These studies took advantage of national genetic breeding evaluations (EBV or DYD) for SCC/SCS kept by departments of agriculture or breed associations to develop phenotypes for association testing. Both estimated breeding values (EBVs) and daughter yield deviations (DYDs) are estimates or predictions of a bull's progeny based on averages of a sampling of the bull's offspring, and do not represent an individual cow's performance directly nor do they account for variation in SCC based on stage of lactation or causative pathogen. To circumvent this confounding variation and directly utilize an individual cow's performance against a particular pathogen as the phenotype of interest, we propose utilizing SCC data collected from direct experimental challenge with *S. uberis* to develop phenotypes and test associations. We hypothesize utilizing our novel SCC phenotypes derived from *S. uberis* experimental challenge will detect significant associations with SNPs and identify potential target genes directly related to *S. uberis* mastitis pathogenesis post-partum. The objective of this study was to perform a GWAS utilizing 40 Holstein dairy cows that were genotyped using the Illumina BovineSNP50 v2 BeadChip (54,609 SNPs) to test associations with novel SCC phenotypes derived SCC collected following direct experimental *S. uberis* challenge.

Materials and Methods

Animal Selection

Holstein dairy cows (N = 40) that were entering their second or third lactation were enrolled in a vaccine trial to test the effectiveness of a novel vaccine candidate against *S. uberis*, i.e.: *Streptococcus uberis* adhesion molecule (SUAM) that involved a direct *S. uberis* challenge (USDA-NIFA AFRI 2011-67015-30168). Cows were in overall

good health, pregnant, at the end of their first or second lactation, had quarter SCC <250,000 cells/mL, and were free of Johne's, brucellosis, tuberculosis, and bovine leukemia virus (BLV). Both *S. uberis* and SUAM baseline titers were assessed and cows with the lowest titers were enrolled. All cows were housed at the East Tennessee Education and Research Center (ETREC) Little River Dairy (Walland, TN). Upon arrival at ETREC, cows were acclimated to the facility for two weeks prior to the initiation of research protocols. Cows were housed separately from the core herd and cared for in accordance with ETREC Little River Dairy standard protocols. All animal based work was approved by and done in accordance with Institutional Care and Use Committee (IACUC) guidelines.

Challenge

Approximately one week prior to calving and at first milking post-calving, secretion samples were collected from each quarter to screen for bacterial presence to select a challenge quarter free of pathogens. A frozen stock of *Streptococcus uberis* (UT888) was thawed in a 37°C water bath and an aliquot used to inoculate a blood agar plate. After 23-24 h of growth on the agar, 3 representative colonies were used to inoculate Todd Hewitt Broth for an additional 7 h of growth. The 1 mL aliquot of the resulting liquid culture was serially diluted in PBS to approximately 2,000 CFU for infusion. A total of 5 mL of the diluted culture was used for challenge, resulting in a total of 10,000 CFU delivered. Challenge was performed within 3 days post-calving with 5 mL of the diluted *S. uberis* culture immediately following milking. Prior to intramammary infusion, teat ends were cleaned thoroughly with sterile alcohol pads and the inoculum was infused using sterile disposable syringes with teat cannulas. The challenge inoculum was massaged upward into the teat and gland cisterns and plated (100µL) to determine the actual CFU/mL of *S. uberis* delivered which averaged 2122 ± 371 CFU/mL and ranged from 1185 to 3202 CFU/mL.

Sampling and Phenotype Description

Milk samples were collected from the challenge quarter in Capitol Vials with B-14 milk preservative pills (Neta Scientific, Inc., Hainesport, NJ) at the following time points: immediately before challenge (0 d) and then 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 11, 14, 18, 21, 25, and 28 d post-challenge. Collected samples were used to determine somatic cell counts (SCC), represented as cells/mL, by the Tennessee Dairy Herd Information Association (DHIA) lab (Knoxville, TN). From the SCC data, three phenotypes were developed. Industry standard considers a SCC below 200,000 cells/mL healthy, individual cows were placed in one of three categories (<21 d, 21-28 d, or >28 d) based on when their SCC returned to below 200,000 cells/mL post-

challenge, which will be referred to as **SCC cure**. Additionally, area under the curve (AUC) for SCC cells/mL for 0-7 days and 0-28 days post-challenge were calculated from linear representations of SCC over time for each cow using SAS (Cary, NC), and will be referred to as **7d AUC** and **28d AUC** respectively. AUCs were calculated in order to have a single value for each cow to test for genetic associations, since looking at effects over time is not possible with a GWAS approach. Select cows were sacrificed throughout the post-challenge sampling as part of the SUAM vaccine trial which resulted in incomplete data sets to generate phenotypes for some cows. As a result, the number of cows per phenotype differed as follows: SCC cure had N=34, 7d AUC had N=36, and 28d AUC had N=35.

SNP Chip Analysis

Blood samples were collected from each cow via jugular puncture and DNA was isolated using UltraClean BloodSpin DNA Isolation Kit (Mo Bio, Carlsbad, CA). Resulting DNA was used to perform a 50K SNP chip analysis using the BovineSNP50 DNA Analysis BeadChip (Illumina, San Diego, CA) (Gunderson et al. 2005; Steemers et al. 2006). The SNP chip has 54,609 SNPs that span all 29 autosomes of the bovine genome with an average SNP spacing of 49.4kb and a median spacing of 36.9kb. Associations were tested between developed phenotypes and SNPs using Plink (Purcell et al. 2007). Effects of sex were excluded from the analysis because all individuals were female and SNPs did not represent sex chromosomes. The p-value cutoffs of $p < 1.95 \times 10^{-5}$ for highly significant SNPs and $p = 1.95 \times 10^{-5}$ to 9.34×10^{-5} for suggestively significant SNPs were set based on guidelines for reducing false positives and increasing reproducibility (Manly 2005). For identified SNPs to be reported, the minor allele had to be represented by more than one cow, unless the SNP aligned with prior evidence for mastitis or mastitis related traits. A complete index of identified SNPs, including those removed based on the minor allele criteria, is included as an appendix (Table A-1). To allow depiction of the numerous SNPs, their genome locations, and their variable association p-values, generated Plink results were visualized using Manhattan plots created with Integrated Genome Viewer (IGV) (Robinson et al. 2011). All reported gene and SNP locations were reported based on the *Bos taurus* UMD_3.1.1 (GCF_000003055.6) genome build.

Results

A total of 16 SNPs associated with unique SCC phenotypes generated from *S. uberis* experimental challenge data were identified. The identified SNPs spanned 11 chromosomes, with 3 SNPs on BTA 9 and 20, 2 SNPs on BTA 3, and the remaining 8 SNPs being unique to individual chromosomes (Table 1). The first phenotype, SCC

Table 1: Identified loci associated with developed novel SCC phenotypes.

Trait	rs	Locus (chr: bp ¹)	SNP type	Allele freq ²	Nearest gene(s)	Prior evidence	P-value
Loci with prior evidence							
7d AUC	rs109961240	20: 63,407,185	5' upstream	A: 0.32 G: 0.68	TAS2R1	SCS QTL (Meredith et al., 2013)	4.14x10 ⁻⁶
28d AUC	rs41655875	5: 26,621,180	intron variant	C: 0.07 T: 0.93	ATF7	MY QTLs (Bennewitz et al., 2004, Awad et al., 2010)	2.80x10 ⁻⁵
28d AUC	rs109513885	3: 34,263,481	intron variant	C: 0.41 T: 0.59	CELSR2	MY QTLs (Viitala et al., 2003, Ashwell et al., 2004)	5.67x10 ⁻⁵
SCC cure	rs42488893	8: 19,967,642	5' upstream	C: 0.62 T: 0.38	U6 (snRNA)	SCC QTL, CM QTL (Klungland et al., 2001)	6.89x10 ⁻⁵
SCC cure	rs29015249	18: 24,862,274	intergenic	C: 0.36 T: 0.64	BREH1, NUP93	PL QTL, SCS QTL, MY QTL (Olsen et al., 2002, Muncie et al., 2006)	8.88x10 ⁻⁵
Newly identified loci							
SCC cure	rs109484182	9: 93,183,596	intron variant	A: 0.51 G: 0.49	TIAM2	--	8.54x10 ⁻⁷
28d AUC	rs43224998	1: 34,903,570	intergenic	A: 0.72 G: 0.28	CHMP2B, VGLL3	--	5.57x10 ⁻⁶
SCC cure	rs110216702	9: 100,163,987	intron variant	A: 0.76 G: 0.24	LOC100336821	--	8.00x10 ⁻⁶
SCC cure	rs29019936	17: 562,366	intron variant	C: 0.45 T: 0.55	CPE	--	1.82x10 ⁻⁵
7d AUC	rs41951389	20: 59,627,852	5' upstream	C: 0.43 T: 0.57	DNAH5	--	2.82x10 ⁻⁵
28d AUC	rs110556992	7: 65,047,126	intron variant	C: 0.56 T: 0.44	GLRA1	--	4.86x10 ⁻⁵
SCC cure	rs41594141	9: 73,368,179	intron variant	A: 0.60 G: 0.40	SGK1	--	5.43x10 ⁻⁵

Table 1: Continued

Trait	rs	Locus (chr: bp ¹)	SNP type	Allele freq ²	Nearest gene(s)	Prior evidence	P-value
7d AUC	rs41612729	28: 5,052,476	5' upstream	A: 0.56 G: 0.44	DISC1	--	6.48x10 ⁻⁵
SCC cure	rs41565622	13: 24,688,809	intron variant	A: 0.53 G: 0.47	LOC530173	--	8.32x10 ⁻⁵
28d AUC	rs43499732	3: 6,480,313	intergenic	G: 0.20 T: 0.80	RGS5, CCDC190	--	8.48x10 ⁻⁵
7d AUC	rs109243453	20: 58,928,800	intergenic	A: 0.40 G: 0.60	TRIO, TRNAC- ACA	--	8.59x10 ⁻⁵

¹Reported positions are based on the *Bos taurus* UMD_3.1.1 (GCF_000003055.6) genome build.

²Allele frequencies are representative of the current study population.

cure, placed cows in one of 3 categories based on when during the post-challenge sampling each cow's SCC returned to below 200,000 cells/mL (Figure 1). The majority of cows (44%) were in the >28 days category, indicating their SCC had not returned to below 200,000 cells/mL before the sampling period ended at 28 days post-challenge. The next largest category was <21 days, which included 32% of cows that reached 200,000 cells/mL before 21 days post-challenge. The last category of 21 – 28 days was the smallest category and represented only 24% of cows. Association analyses identified 3 highly significant ($p < 1.95 \times 10^{-5}$) and 4 suggestively significant ($p = 1.95 \times 10^{-5}$ to 9.34×10^{-5}) SNPs associated with SCC cure (Figure 2). One suggestively significant SNP is located on each BTA 8 (rs42488893), 13 (rs41565622), and 18 (rs29015249) in regions not associated with any known functions. A significant SNP (rs29019936) is located on BTA 17 within the CPE (carboxypeptidase E) gene and the remaining three SNPs are part of a cluster of SNPs located on BTA 9, with SNPs in both the TIAM2 (t-cell lymphoma invasion and metastasis 2) (rs109484182), SGK1 (serum/glucocorticoid regulated kinase 1) (rs41594141), and LOC100336821 (rs41565622) genes.

The second and third phenotypes were a quantitative representation of SCC AUC over time, either for the first 7 days post-challenge (7d AUC) or for the entire 28 day sampling period (28d AUC). For 7d AUC, areas ranged from 1.7×10^6 (cells/mL)*d to 91.3×10^6 (cells/mL)*d, and had a median of 30.6×10^6 (cells/mL)*d (Figure 3). Association analyses identified 1 highly significant ($p < 1.95 \times 10^{-5}$) and 3 suggestively significant ($p = 1.95 \times 10^{-5}$ to 9.34×10^{-5}) SNPs associated with 7d AUC (Figure 4). One SNP (rs41612729) is located on BTA 28 and the remaining three SNPs are part of a cluster of SNPs on BTA 20 (rs109961240, rs41951389, and rs109243453). None of the SNPs identified with 7d AUC are located in regions with known functions. For 28d AUC, areas ranged from 2.5×10^6 (cells/mL)*d to 246.1×10^6 (cells/mL)*d, with a median of 69.4×10^6 (cells/mL)*d (Figure 3). Association analyses identified 1 highly significant ($p < 1.95 \times 10^{-5}$) and 4 suggestively significant ($p = 1.95 \times 10^{-5}$ to 9.34×10^{-5}) SNPs associated with 28 day AUC (Figure 5). One SNP was identified on each BTA 1, 5, and 7, where the SNP on BTA 5 (rs41655875) is in the ATF7 (activating transcription factor 7) gene and the SNP on BTA 7 (rs110556992) is in the GLRA1 (glycine receptor, alpha 1) gene. The remaining 2 SNPs were both on BTA 3, with one located in the CELSR2 (cadherin, EGF LAG seven-pass G-type receptor 2) gene (rs109513885).

Discussion

This study successfully identified associations between SNPs in the bovine genome and novel SCC phenotypes developed from SCC following direct *S. uberis* experimental challenge. To our knowledge, this study is the first to utilize data generated from an experimental challenge to develop GWAS mastitis phenotypes.

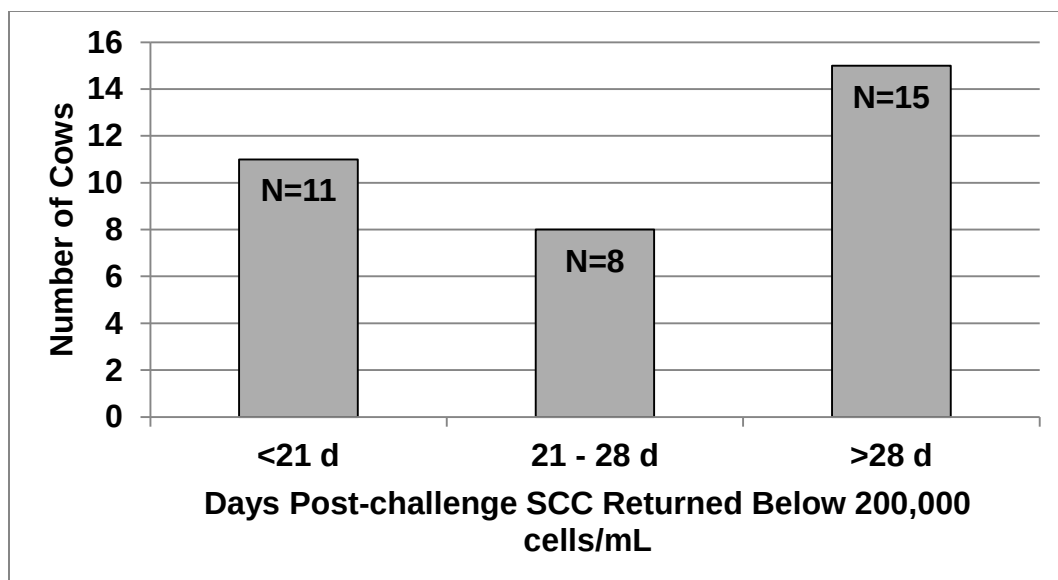


Figure 1: Number of cows per SCC cure phenotype category. Ranges of days represent the time range individual cow's SCC returned to below 200,000 cells/mL, which is industry standard for healthy.

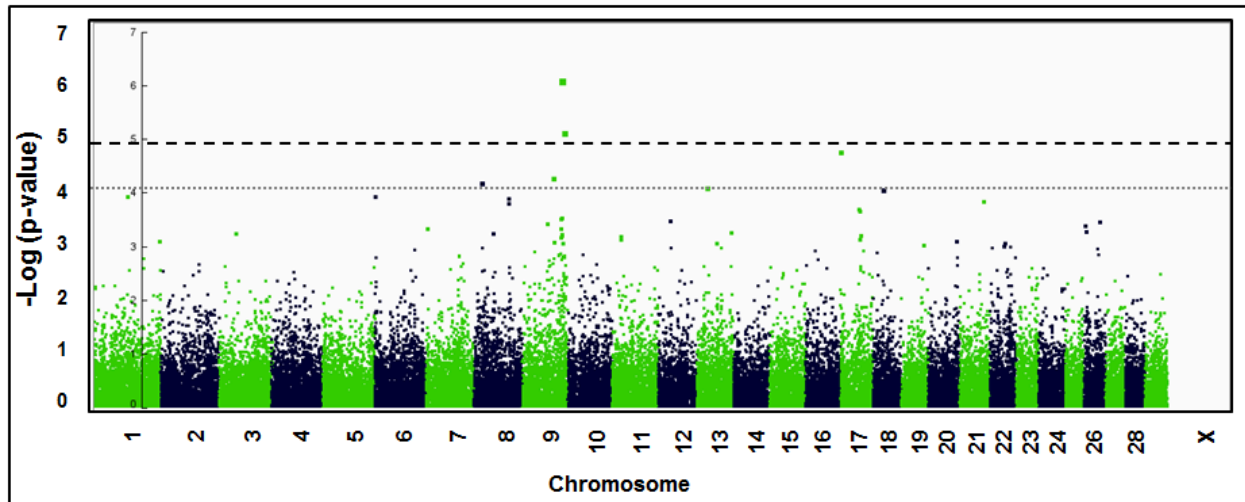


Figure 2: Manhattan plot of GWAS for SCC cure phenotype using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) in Holsteins (N = 34). The x-axis shows chromosome and SNP order within a chromosome. The y-axis show the $-\log(p\text{-value})$ for each SNP. P-value thresholds are indicated by dashed lines: top black line $p = 1.95 \times 10^{-5}$; bottom gray line $p = 9.34 \times 10^{-5}$.

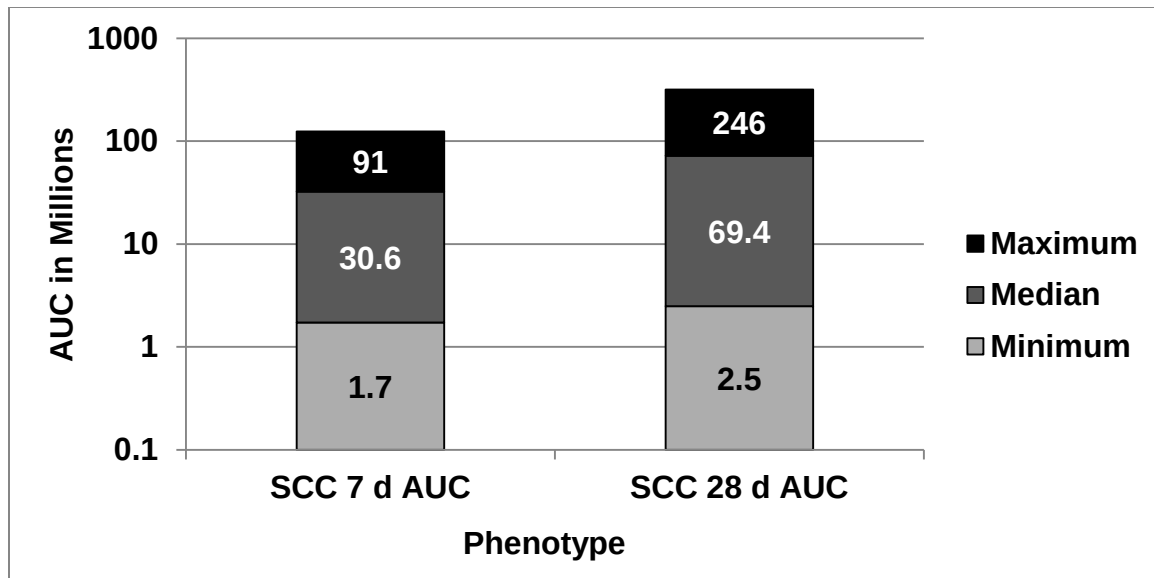


Figure 3: Minimum, median, and maximum SCC AUC values for both 7 and 28 d phenotypes displayed in millions of (cells/mL)*d.

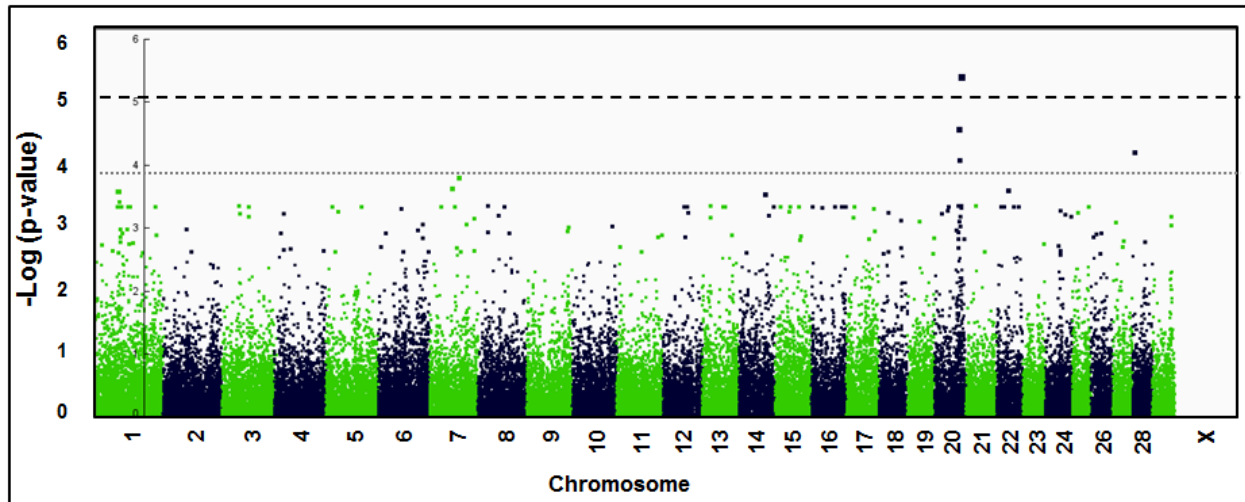


Figure 4: Manhattan plot of GWAS for SCC 7 day AUC phenotype using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) in Holsteins (N = 36). The x-axis shows chromosome and SNP order within a chromosome. The y-axis show the $-\log(p\text{-value})$ for each SNP. P-value thresholds are indicated by dashed lines: top black line $p = 1.95 \times 10^{-5}$; bottom gray line $p = 9.34 \times 10^{-5}$.

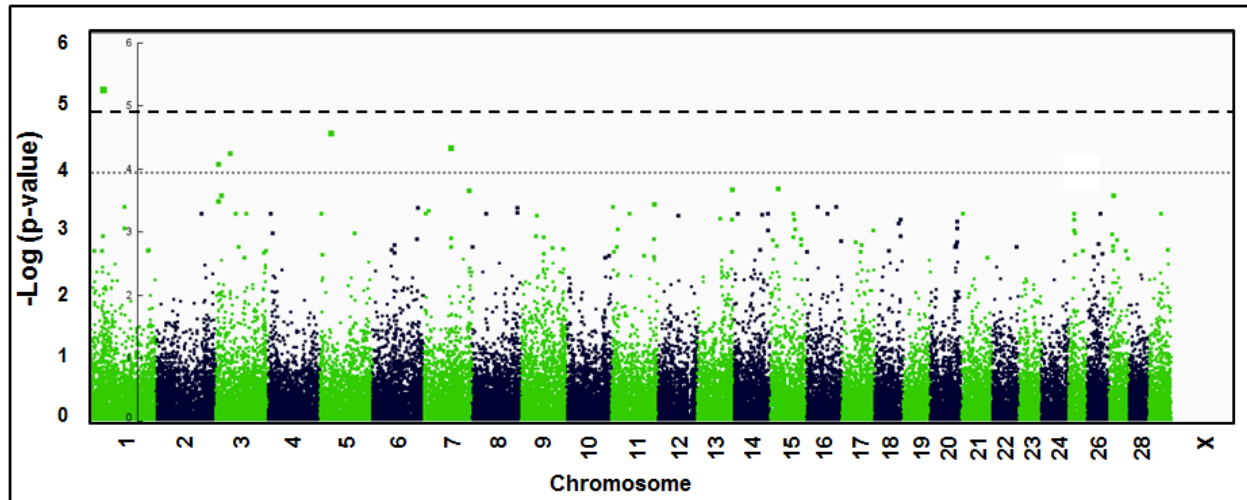


Figure 5: Manhattan plot of GWAS for SCC 28 day AUC phenotype using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) in Holsteins (N = 35). The x-axis shows chromosome and SNP order within a chromosome. The y-axis show the $-\log(p\text{-value})$ for each SNP. P-value thresholds are indicated by dashed lines: top black line $p = 1.95 \times 10^{-5}$; bottom gray line $p = 9.34 \times 10^{-5}$.

Additionally, this study is unique because of its power, with N=34-36 Holstein cows compared to hundreds of individuals in other GWAS studies focused on mastitis (Meredith et al. 2013; Sahana et al. 2014; Tal-Stein et al. 2010). However, 5 of the identified SNPs in this study (rs109961240, rs41655875, rs109513885, rs42488893, and rs29015249) are located in genomic regions with prior evidence associated with mastitis or related phenotypes (Table 1). Of these, the SCC cure phenotype was associated with SNP rs42488893 which was also within a clinical mastitis (CM) QTL (Klungland et al. 2001). This CM QTL was a binary trait where cows either had or did not have CM from 15 d pre-calving to 120 d post-calving. Our SCC cure phenotype focused on the first 28 d post-calving, during which we observed 62.5% of cows develop clinical mastitis, indicated by inflammation, altered milk production, and an elevated (>200,000 c/mL) SCC (data not shown). The highly inflammatory response observed with clinical mastitis suggests a strongly stimulated immune response, which could lead to a more efficient clearance of invading bacteria and result in a shorter period of time with SCC >200,000 c/mL, correlating to a lower SCC cure phenotype. Three SNPs significantly associated with the phenotypes 7d AUC (rs109961240) and SCC cure (rs42488893, rs29015249) were located within QTLs associated with SCC and SCS (Klungland et al. 2001; Meredith et al. 2013; Muncie et al. 2006). These QTLs used SCC lactation means or SCC averaged across a sire's daughters as phenotypes meaning loci identified are likely associated with cows with higher propensity for infections or cows with longer recovery times post-infection. The 7d AUC phenotype accumulated SCC for the first 7 days post-challenge, where cows with higher 7d AUC would have higher propensity for infection and those with lower 7d AUC would have lower propensity, which links our 7d AUC phenotype directly to the previously identified QTLs. The SCC cure phenotype put cows in categories based on when their SCC returned to below 200,000 c/mL, indicating a healthy gland, which would correlate to cows with length of time it takes for cows to recover post-infection linking SCC cure directly to the previously identified SCC/SCS QTLs. Of interest, SNPs associated with 28d AUC (rs41655875, rs109513885) and SCC cure (rs29015249) phenotypes were located in regions previously identified as having QTLs for productive life (PL) and/or milk yield (MY) (Ashwell et al. 2004; Awad et al. 2010; Bennewitz et al. 2004; Muncie et al. 2006; Olsen et al. 2002; Viitala et al. 2003). Productive life refers to the length of time a cow remains in the herd producing milk, which can be affected by many factors including health, fertility, and milk yield (Roxström and Strandberg 2002). A longer SCC cure (i.e.: >28d) and/or stronger inflammatory responses for 28d (AUC) indicates cows are not as effective at resolving infection. This would increase their risk of being culled from the herd because of mastitis directly or other negative impacts associated with mastitis such as reduced fertility and milk yield. For decades, an association between mastitis incidence and reduced milk production has been recognized (Meredith et al.

2012; Shaw and Beam 1935; Wilson et al. 1997) since fighting infections sequesters energy that would otherwise be partitioned to milk production. Since SCC is used as an indicator of mastitis, is the foundation for both our SCC cure and 28d AUC phenotypes, and is associated with decreased milk production, the identification of SNPs in regions with MY QTLs is logical. The identification of SNPs within known QTL regions with prior evidence helps establish the validity of using these novel SCC based phenotypes in response to direct, experimental *S. uberis* challenge for GWAS. Furthermore, while the associations identified in this study are associated with our novel *S. uberis* SCC phenotypes, since some of the identified SNPs are located in regions with prior evidence linked to varying mastitis and/or MY phenotypes, it is possible these regions could be associated with other causative pathogens besides *S. uberis*.

While the identified SNPs themselves may not be causal but rather in linkage disequilibrium is the causal variant, they do provide several interesting candidate genes based on the genes they are located within. Of the previously identified loci, the two identified using 28d AUC are also located within genes with known functions. First, rs41655875 is located within the ATF7 (activating transcription factor 7) gene. ATF7 binds the cAMP response element of many cellular promoters including many immune genes and helps control the epigenetic regulation of the same genes resulting in the speculation that ATF7 is involved in the controlling the 'memory' of the innate immune system through these epigenetic changes that persist for long periods of time (Yoshida et al. 2015). Control of the innate immune system through gene regulation would allow for faster responses to infection states and allow for rapid clearance of invading bacteria and be associated with overall lower 28d AUC. The other SNP, rs109513885, is located within the CELSR2 (cadherin EGF LAG seven-pass G-type receptor 2) gene that is primarily involved in central nervous system dendrite morphogenesis (Shima et al. 2007). The functions of CELSR2 are linked to calcium signaling, cell adhesion, and cell survival, which are functions shared with many g-protein receptors (Beall et al. 2005; Shima et al. 2007) frequently involved in migration of immune cells to the site of infection (Raman et al. 2010; Rambeaud et al. 2006; Rambeaud and Pighetti 2007; Richardson et al. 2003), thus impacts in CELSR2 might impact observed SCC and be linked to a higher or lower 28d AUC.

In addition to the SNPs identified in regions with prior evidence, this study identified 12 novel SNPs. The SCC cure phenotype identified 5 novel SNPs, of these, 3 are within genes with known functions. First, rs109484182, located on BTA 9 at 93.2Mb, is within the TIAM2 (t-cell invasion and metastasis-2) gene, which encodes a guanine nucleotide exchange factor that plays a role in the function of small GTPases. Roles of TIAM2 based on human research have included increasing cell invasions in

cancer development, cellular migration, adhesion, and apoptosis (Chen et al. 2012; Rooney et al. 2010). While the majority of existing research has focused on TIAM2's role in cancer development and metastasis, it has documented roles in general immune functions like migration and apoptosis that could affect a cow's SCC influx in response to invading pathogenic bacteria and thus its SCC cure phenotype. Another SNP, rs41594141, also located on BTA 9, but at 73.4Mb, is within the SGK1 (serum/glucocorticoid regulated kinase 1) gene, which is a serine/threonine kinase that is expressed in virtually all tissues where it regulates ion channels (Lang et al. 2009; Lang and Shumilina 2013). SGK1 also helps regulate the expression of NFkB, which is a transcription factor for many genes involved in effector functions of innate immune system (Baeuerle and Henkel 1994; Lang et al. 2010). NFkB regulated effector functions include the up regulation of the pro-inflammatory cytokine IL-8 and the adhesion molecules ICAM-1 and VCAM-1, which are vital for the recruitment of neutrophils to the site of infection (Collins et al. 1995; Matsusaka et al. 1993). Alterations in the ability to recruit neutrophils to the mammary gland during mastitis would likely permit the infection to persist and cause a prolonged period of mildly elevated SCC that is not sufficient enough to clear the infection and correlating to a high SCC cure phenotype observed. The third SNP located within a gene identified with the SCC cure phenotype, rs29019936, is located on BTA 17 at 0.6Mb, and is within the CPE (carboxypeptidase E) gene that primarily functions within the endocrine and nervous systems for the biosynthesis of neuropeptides and peptide hormones (Fricker 1988). Neuropeptides and peptide hormones have diverse roles in the regulation of the immune system including cytokine secretion, adhesion molecule expression, and feedback regulation as reviewed by Sternberg (2006). Additionally, CPE directly has been demonstrated to play a role in growth and migration of cancer cells in humans (Lee et al. 2011). The effects of these roles combined could dictate the ability of immune cells to migrate to the mammary gland resulting in overall lower SCC observed which would correlate to lower SCC cure phenotypes.

The 7d AUC phenotype identified 3 novel SNPs. While none of these SNPs were directly located in genes, there were a cluster of SNPs identified on BTA 20 including SNPs at positions 58.9Mb (rs109243453), 59.6Mb (rs41951389), and 63.4Mb (rs109961240, previously identified). Several other GWAS have identified associations with loci on BTA20 and SCC based phenotypes (Cole et al. 2011; Meredith et al. 2013; Meredith et al. 2012; Sahana et al. 2014; Sodeland et al. 2011; Tal-Stein et al. 2010). Within this BTA 20 cluster region there are 11 known genes, 4 tRNAs, 1 miRNA, and 7 ncRNAs, providing a number of possible candidate genes and functional RNAs. One gene of interest is death-associated protein (DAP)-1 which is a positive mediator of interferon (IFN)- γ induced programmed cell death (Deiss et al. 1995). Higher rates of

programmed cell death could lead to lower SCC and would correlate to lower 7d AUC phenotypes observed. Lastly, the 28d AUC phenotype identified the remaining 3 novel SNPs, where one SNP is located within a gene with known function. The SNP on BTA 7 at 65Mb is located within the GLRA1 (glycine receptor alpha 1) gene. GLRA1 is a subunit of an ion-gated, chlorine channel glycine receptor. Glycine and its receptor primarily function within the central nervous system to mediate synapse impulses (Van den Eynden et al. 2009). Glycine receptors also influence calcium ion fluxes in neutrophils and monocytes leading to reduced pro-inflammatory cytokine production, increased anti-inflammatory production, and reduced superoxide production resulting in overall less stimulation of the immune response to invading pathogens (Spittler et al. 1999; Wheeler et al. 2000). Less stimulation of the immune response would result in fewer immune cells recruited to the site of infection and lead to overall lower observed SCC and lower 28d AUC phenotypes.

In summary, our study identified SNPs associated with novel SCC based phenotypes derived from SCC collected in response to direct, experimental *S. uberis* challenge. While our study was limited by a low number of individual cows being genotyped, many of the identified SNPs (31%) are within genomic regions previously reported in considerably larger studies as associated with mastitis related traits, providing validity to the effectiveness of our novel phenotypes for identifying regions of interest. The identified SNPs provide several loci to investigate as potential target genes that could be linked directly to *S. uberis* pathogenesis. Furthermore, investigations of these identified loci or those identified in future analyses using challenge data could lead to novel treatment or preventative strategies for *S. uberis* mastitis and potentially identify genetic selection markers for cows with greater resistance to *S. uberis* mastitis.

Acknowledgements

The authors would like to thank Dr. Arnold Saxton of the department of Animal Science at the University of Tennessee for his assistance in converting the data format. The challenge portion of this study was supported by USDA-NIFA (2011-67015-30168). Additional support was obtained from the UT AgResearch AgInnovation Fund for SNP chip analyses.

References

- Ashwell, M. S., Heyen, D. W., Sonstegard, T. S., Van Tassell, C. P., Da, Y., VanRaden, P. M., Ron, M., Weller, J. I., and Lewin, H. A. (2004). "Detection of quantitative trait loci affecting milk production, health, and reproductive traits in Holstein cattle." *Journal of Dairy Science*, 87(2), 468-475.
- Awad, A., Russ, I., Emmerling, R., Forster, M., and Medugorac, I. (2010). "Confirmation and refinement of a QTL on BTA5 affecting milk production traits in the Fleckvieh dual purpose cattle breed." *Animal Genetics*, 41(1), 1-11.
- Baeuerle, P. A., and Henkel, T. (1994). "Function and activation of NF-kappaB in the immune system." *Annual review of immunology*, 12(1), 141-179.
- Beall, S. A., Boekelheide, K., and Johnson, K. J. (2005). "Hybrid GPCR/cadherin (Celsr) proteins in rat testis are expressed with cell type specificity and exhibit differential sertoli cell—germ cell adhesion activity." *Journal of Andrology*, 26(4), 529-538.
- Bennewitz, J., Reinsch, N., Guiard, V., Fritz, S., Thomsen, H., Looft, C., Kuhn, C., Schwerin, M., Weimann, C., Erhardt, G., Reinhardt, F., Reents, R., Boichard, D., and Kalm, E. (2004). "Multiple quantitative trait loci mapping with cofactors and application of alternative variants of the false discovery rate in an enlarged granddaughter design." *Genetics*, 168(2), 1019-27.
- Chen, J. S., Su, I. J., Leu, Y. W., Young, K. C., and Sun, H. S. (2012). "Expression of T-cell lymphoma invasion and metastasis 2 (TIAM2) promotes proliferation and invasion of liver cancer." *International Journal of Cancer*, 130(6), 1302-1313.
- Cole, J. B., Wiggans, G. R., Ma, L., Sonstegard, T. S., Lawlor, T. J., Jr., Crooker, B. A., Van Tassell, C. P., Yang, J., Wang, S., Matukumalli, L. K., and Da, Y. (2011). "Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary U.S. Holstein cows." *BMC Genomics*, 12, 408.
- Collins, T., Read, M. A., Neish, A. S., Whitley, M. Z., Thanos, D., and Maniatis, T. (1995). "Transcriptional regulation of endothelial cell adhesion molecules: NF-kappa B and cytokine-inducible enhancers." *The FASEB Journal*, 9(10), 899-909.
- Deiss, L. P., Feinstein, E., Berissi, H., Cohen, O., and Kimchi, A. (1995). "Identification of a novel serine/threonine kinase and a novel 15-kD protein as potential mediators of the gamma interferon-induced cell death." *Genes & development*, 9(1), 15-30.
- Fricker, L. D. (1988). "Carboxypeptidase E." *Annual Review of Physiology*, 50(1), 309-321.

- Klungland, H., Sabry, A., Heringstad, B., Olsen, H. G., Gomez-Raya, L., Våge, D. I., Olsaker, I., Ødegård, J., Klemetsdal, G., Schulman, N., Vilkkki, J., Ruane, J., Aasland, M., Rønningen, K., and Lien, S. (2001). "Quantitative trait loci affecting clinical mastitis and somatic cell count in dairy cattle." *Mammalian Genome*, 12(11), 837-842.
- Lang, F., Artunc, F., and Vallon, V. (2009). "The physiological impact of the serum- and glucocorticoid-inducible kinase SGK1." *Current Opinion in Nephrology and Hypertension*, 18(5), 439-448.
- Lang, F., and Shumilina, E. (2013). "Regulation of ion channels by the serum- and glucocorticoid-inducible kinase SGK1." *The FASEB Journal*, 27(1), 3-12.
- Lang, F., Strutz-Seeböhm, N., Seeböhm, G., and Lang, U. E. (2010). "Significance of SGK1 in the regulation of neuronal function." *The Journal of Physiology*, 588(18), 3349-3354.
- Lee, T. K., Murthy, S. R. K., Cawley, N. X., Dhanvantari, S., Hewitt, S. M., Lou, H., Lau, T., Ma, S., Huynh, T., and Wesley, R. A. (2011). "An N-terminal truncated carboxypeptidase E splice isoform induces tumor growth and is a biomarker for predicting future metastasis in human cancers." *The Journal of Clinical Investigation*, 121(3), 880-892.
- Matsusaka, T., Fujikawa, K., Nishio, Y., Mukaida, N., Matsushima, K., Kishimoto, T., and Akira, S. (1993). "Transcription factors NF-IL6 and NF-kappa B synergistically activate transcription of the inflammatory cytokines, interleukin 6 and interleukin 8." *Proceedings of the National Academy of Sciences*, 90(21), 10193-10197.
- Meredith, B. K., Berry, D. P., Kearney, F., Finlay, E. K., Fahey, A. G., Bradley, D. G., and Lynn, D. J. (2013). "A genome-wide association study for somatic cell score using the Illumina high-density bovine beadchip identifies several novel QTL potentially related to mastitis susceptibility." *Frontiers in Genetics*, 4(229).
- Meredith, B. K., Kearney, F. J., Finlay, E. K., Bradley, D. G., Fahey, A. G., Berry, D. P., and Lynn, D. J. (2012). "Genome-wide associations for milk production and somatic cell score in Holstein-Friesian cattle in Ireland." *BMC Genetics*, 13(1), 1-11.
- Muncie, S. A., Cassady, J. P., and Ashwell, M. S. (2006). "Refinement of quantitative trait loci on bovine chromosome 18 affecting health and reproduction in US Holsteins." *Animal Genetics*, 37(3), 273-275.
- Olsen, H. G., Gomez-Raya, L., Vage, D. I., Olsaker, I., Klungland, H., Svendsen, M., Adnøy, T., Sabry, A., Klemetsdal, G., Schulman, N., Kramer, W., Thaller, G., Rønningen, K., and Lien, S. (2002). "A genome scan for quantitative trait loci

- affecting milk production in Norwegian dairy cattle." *Journal of Dairy Science*, 85(11), 3124-30.
- Raman, D., Sai, J., Neel, N. F., Chew, C. S., and Richmond, A. (2010). "LIM and SH3 protein-1 modulates CXCR2-mediated cell migration." *Public Library of Science One*, 5(4), e10050.
- Rambeaud, M., Clift, R., and Pighetti, G. M. (2006). "Association of a bovine CXCR2 gene polymorphism with neutrophil survival and killing ability." *Veterinary Immunology and Immunopathology*, 111(3–4), 231-238.
- Rambeaud, M., and Pighetti, G. (2007). "Differential calcium signaling in dairy cows with specific CXCR1 genotypes potentially related to interleukin-8 receptor functionality." *Immunogenetics*, 59(1), 53-58.
- Richardson, R. M., Marjoram, R. J., Barak, L. S., and Snyderman, R. (2003). "Role of the cytoplasmic tails of CXCR1 and CXCR2 in mediating leukocyte migration, activation, and regulation." *The Journal of Immunology*, 170(6), 2904-11.
- Rooney, C., White, G., Nazgiewicz, A., Woodcock, S. A., Anderson, K. I., Ballestrin, C., and Malliri, A. (2010). "The Rac activator STEF (Tiam2) regulates cell migration by microtubule-mediated focal adhesion disassembly." *EMBO Reports*, 11(4), 292-298.
- Roxström, A., and Strandberg, E. (2002). "Genetic analysis of functional, fertility-, mastitis-, and production-determined length of productive life in Swedish dairy cattle." *Livestock Production Science*, 74(2), 125-135.
- Sahana, G., Guldbrandtsen, B., Thomsen, B., Holm, L. E., Panitz, F., Brøndum, R. F., Bendixen, C., and Lund, M. S. (2014). "Genome-wide association study using high-density single nucleotide polymorphism arrays and whole-genome sequences for clinical mastitis traits in dairy cattle." *Journal of Dairy Science*, 97(11), 7258-7275.
- Shaw, A. O., and Beam, A. L. (1935). "The effect of mastitis upon milk production." *Journal of Dairy Science*, 18(6), 353-357.
- Shima, Y., Kawaguchi, S.-y., Kosaka, K., Nakayama, M., Hoshino, M., Nabeshima, Y., Hirano, T., and Uemura, T. (2007). "Opposing roles in neurite growth control by two seven-pass transmembrane cadherins." *Nature Neuroscience*, 10(8), 963-969.
- Sodeland, M., Kent, M. P., Olsen, H. G., Opsal, M. A., Svendsen, M., Sehested, E., Hayes, B. J., and Lien, S. (2011). "Quantitative trait loci for clinical mastitis on chromosomes 2, 6, 14 and 20 in Norwegian Red cattle." *Animal Genetics*, 42(5), 457-465.

- Spittler, A., Reissner, C. M., Oehler, R., Gornikiewicz, A., Gruenberger, T., Manhart, N., Brodowicz, T., Mittlboeck, M., Boltz-Nitulescu, G., and Roth, E. (1999). "Immunomodulatory effects of glycine on LPS-treated monocytes: reduced TNF- α production and accelerated IL-10 expression." *The FASEB Journal*, 13(3), 563-571.
- Sternberg, E. M. (2006). "Neural regulation of innate immunity: a coordinated nonspecific host response to pathogens." *Nature Reviews Immunology*, 6(4), 318-328.
- Tal-Stein, R., Fontanesi, L., Dolezal, M., Scotti, E., Bagnato, A., Russo, V., Canavesi, F., Friedmann, A., Soller, M., and Lipkin, E. (2010). "A genome scan for quantitative trait loci affecting milk somatic cell score in Israeli and Italian Holstein cows by means of selective DNA pooling with single- and multiple-marker mapping." *Journal of Dairy Science*, 93(10), 4913-4927.
- Van den Eynden, J., Horwood, N., Carmans, S., Brône, B., Hellings, N., Steels, P., Harvey, R. J., and Rigo, J.-M. (2009). "Glycine and glycine receptor signalling in non-neuronal cells."
- Viitala, S. M., Schulman, N. F., de Koning, D. J., Elo, K., Kinoshita, R., Virta, A., Virta, J., Maki-Tanila, A., and Vilkkii, J. H. (2003). "Quantitative trait loci affecting milk production traits in Finnish Ayrshire dairy cattle." *Journal of Dairy Science*, 86(5), 1828-36.
- Wheeler, M., Stachlewitz, R. F., Yamashina, S., Ikejima, K., Morrow, A. L., and Thurman, R. G. (2000). "Glycine-gated chloride channels in neutrophils attenuate calcium influx and superoxide production." *The FASEB Journal*, 14(3), 476-484.
- Wilson, D. J., Gonzalez, R. N., and Das, H. H. (1997). "Bovine mastitis pathogens in New York and Pennsylvania: Prevalence and effects on somatic cell count and milk production." *Journal of Dairy Science*, 80(10), 2592-2598.
- Yoshida, K., Maekawa, T., Zhu, Y., Renard-Guillet, C., Chatton, B., Inoue, K., Uchiyama, T., Ishibashi, K.-i., Yamada, T., Ohno, N., Shirahige, K., Okada-Hatakeyama, M., and Ishii, S. (2015). "The transcription factor ATF7 mediates lipopolysaccharide-induced epigenetic changes in macrophages involved in innate immunological memory." *Nature Immunology*, 16(10), 1034-1043.

Chapter IV: Genome wide association study identifies loci associated with novel phenotypes following experimental challenge with *Streptococcus uberis*

Genome wide association study identifies loci associated with novel phenotypes following experimental challenge with *Streptococcus uberis*

L. Siebert¹, M.E. Staton², S.P. Oliver^{1, 3}, and G.M. Pighetti¹

¹The University of Tennessee, Knoxville, TN, Department of Animal Science

²The University of Tennessee, Knoxville, TN, Department of Entomology and Plant Pathology

³The University of Tennessee, Knoxville, TN, AgResearch

Abstract

Mastitis is a detrimental disease in the dairy industry that costs upwards of \$2 billion annually and decreases milk quality. Often, mastitis results from bacteria entering the gland through the teat end. A common mastitis causing pathogen is *S. uberis*, which is responsible for a high percentage of subclinical and clinical mastitis cases. Following an intramammary experimental challenge with *S. uberis* on Holstein cows (n=40), mammary and milk scores were observed to evaluate the level of inflammation present and aseptic milk samples were collected to determine CFU of *S. uberis* present in milk. Traditional GWAS have utilized somatic cell count (SCC) based phenotypes or clinical mastitis (CM) observations to identify loci of interest for mastitis. We proposed using novel phenotypes developed from data collected following direct *S. uberis* experimental challenge to generate 6 phenotypes. Four phenotypes were generated using *S. uberis* CFU: *S. uberis* 1st clearance which enumerated the days until CFU reached zero, *S. uberis* cycles which enumerated the number of times CFU reached zero within the 28d sampling period, area under the curve (AUC) of *S. uberis* CFU for 0-7 d and 0-28 d post-challenge. The last two phenotypes were developed using mammary and milk scores: severity score summed all scores for the highly inflammatory 0-7 d post-challenge and chronic score summed all scores for the entire 0-28 d sampling period post-challenge. To identify loci of interest a 50K SNP chip analysis was performed using the BovineSNP50 v2 DNA BeadChip from Illumina. Associations were tested using Plink and identified 29 significant ($p < 9.34 \times 10^{-5}$) SNPs across the 6 phenotypes. Nine of the identified SNPs are located within previously identified quantitative trait loci (QTLs) for traits with links to our generated phenotypes. Additionally, 14 identified SNPs are located within known genes, with most having functions related to the immune response or gene/protein regulation. The identified loci should be further investigated to potentially identify causation behind the observed phenotypes. Such investigations could lead to novel treatment and prevention compounds/protocols for *S. uberis* mastitis or genetic selection methods for cows with greater potential to resist *S. uberis* infection.

Introduction

Mastitis affects 20-50% of cows within any given herd resulting in dramatic decreases in milk quality and farm revenues (Akers and Nickerson 2011; Barbano et al. 1991; Bradley et al. 2007; Klei et al. 1998; Ma et al. 2000; NMC 2005; Østerås et al. 2006; Piepers et al. 2007; Wilson et al. 1997). Because of the impact and high incidence of mastitis, improved control and prevention is a constant goal of the dairy industry. Mastitis frequently results from pathogenic bacteria invading the mammary gland. Common mastitis causing pathogens include *Escherichia coli*, *Streptococcus uberis*, and *Staphylococcus aureus*, where each pathogen has its own unique pathology (Schukken et al. 2011). *S. uberis* is responsible for a high percentage of both clinical and subclinical mastitis and is commonly isolated during the dry (non-lactating) period (Bradley et al. 2007; Petrovski et al. 2011; Riekerink et al. 2008; Verbeke et al. 2014). *S. uberis* is especially problematic because, while it typically originates as an environmental pathogen, it can spread throughout the herd in a contagious manner once introduced (Zadoks 2007). Intramammary infections are characterized by increases in somatic cell counts (SCC), which enumerate the number of immune cells recruited to the infection site to combat invading bacteria. Some resident immune cells (i.e.: macrophages) reside in the mammary gland to provide protection at all time, thus, industry standard considers a SCC below 200,000 cells/mL a healthy gland and SCC above 200,000 cells/mL to indicate either subclinical or clinical mastitis (Dohoo et al. 2011; Schepers et al. 1997; Schukken et al. 2003). Subclinical *S. uberis* infections are indicated by an increase in SCC, whereas clinical *S. uberis* infections are represented by an increase in SCC and the presence of clinical signs (i.e.: inflammation, altered milk production) (Rambeaud et al. 2003). Additionally, *S. uberis* has some unique pathology characteristics that can make clearing the infection from the mammary gland more problematic. Once *S. uberis* enters the mammary gland, it binds lactoferrin, which is a common milk protein (Fang and Oliver 1999). Lactoferrin can then act as a bridging molecule and allow binding of *S. uberis* to the mammary epithelial cells (Fang et al. 2000). After binding, *S. uberis* then invades the host cells via caveolae dependent endocytosis (Almeida and Oliver 2006; Matthews et al. 1994). Once internalized, *S. uberis* can exploit host cell transduction and persist in the cytoplasm for hours (Almeida et al. 2000; Tamilselvam et al. 2006) effectively evading the host immune system. Since *S. uberis* mastitis is a common problem dairy producers combat and *S. uberis* has mechanisms to evade the immune response, identifying specific host genes involved in *S. uberis* pathogenesis could prove beneficial for the control and prevention of *S. uberis* mastitis on dairy farms.

Many studies attempting to identify genes important for mastitis have been conducted (Cole et al. 2009; Meredith et al. 2013; Meredith et al. 2012; Pryce et al. 2010). These genome wide association studies (GWAS) have utilized SCC based phenotypes or observations of clinical mastitis to test for associations. Our prior study, (Siebert et al., 2017, unpublished) was the first to utilize SCC collected in response to direct experimental intramammary challenge to identify associations with genomic regions. For the present study, we propose utilizing other subsets of data (i.e.: CFU of *S. uberis* in milk, mammary score, milk scores) collected in response to direct experimental intramammary *S. uberis* challenge to develop novel phenotypes and test associations with the bovine genome. We hypothesize utilizing our novel phenotypes derived from experimental challenge will increase specificity of associations and identify potential target genes directly related to *S. uberis* mastitis pathogenesis. The objective of this study was to perform a GWAS utilizing 40 Holstein dairy cows that were genotyped using the Illumina BovineSNP50 v2 BeadChip (54,609 SNPs) to test associations with novel phenotypes derived data collected following direct experimental *S. uberis* challenge.

Materials and Methods

Animal Selection

Holstein dairy cows (N = 40) that were entering their second or third lactation were enrolled in a trial to test the effectiveness of a novel vaccine candidate against *S. uberis*, i.e.: *Streptococcus uberis* adhesion molecule (SUAM) that involved a direct *S. uberis* challenge (USDA-NIFA AFRI 2011-67015-30168). Cows were in overall good health, pregnant, at the end of their first or second lactation, had quarter SCC <250,000 cells/mL, and were free of Johne's, brucellosis, tuberculosis, and bovine leukemia virus (BLV). Both *S. uberis* and SUAM baseline titers were assessed and cows with the lowest titers were enrolled. All cows were housed at the East Tennessee Education and Research Center (ETREC) Little River Dairy (Walland, TN). Upon arrival at ETREC, cows were acclimated to the facility for two weeks prior to the initiation of research protocols. Cows were housed separately from the core herd and cared for in accordance with ETREC Little River Dairy standard protocols. All animal based work was approved by and done in accordance with Institutional Care and Use Committee (IACUC) guidelines.

Challenge

Approximately one week prior to calving and at first milking post-calving, secretion samples were collected from each quarter to screen for bacterial presence to

select a challenge quarter free of pathogens. A frozen stock of *Streptococcus uberis* (UT888) was thawed in a 37°C water bath and an aliquot used to inoculate a blood agar plate. After 23-24 h of growth on the agar, 3 representative colonies were used to inoculate Todd Hewitt Broth for an additional 7 h of growth. The 1 mL aliquot of the resulting liquid culture was serially diluted in PBS to approximately 2,000 CFU for infusion. A total of 5 mL of the diluted culture was used for challenge, resulting in a total of 10,000 CFU delivered. Challenge was performed within 3 days post-calving immediately following milking. Prior to intramammary infusion, teat ends were cleaned thoroughly with sterile alcohol pads and the inoculum was infused using sterile disposable syringes with teat cannulas. The challenge inoculum was massaged upward into the teat and gland cisterns and plated (100µL) to determine the actual CFU/mL of *S. uberis* delivered which averaged 2122 ± 371 CFU/mL and ranged from 1185 to 3202 CFU/mL within the 5mL inoculation volume.

Sampling and Phenotype Description

Cows were closely monitored post-challenge for the development of mastitis with specific collections and observations taken at the following time points: immediately before challenge (0 d) then 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 11, 14, 18, 21, 25, and 28 d post-challenge. At each time point, a variety of measures were taken and used to develop novel phenotypes to test for associations with genomic regions.

Milk samples collected aseptically were used to determine the amount of *S. uberis* present, in CFU/mL, via culturing on blood agar plates as outlined in the NMC lab handbook (NMC 1999) and carried out by the Tennessee Quality Milk Lab (Knoxville, TN). Determined CFU/mL of *S. uberis* was used to develop four separate novel phenotypes. All cows became infected from challenge, based on *S. uberis* isolation from challenged quarter, thus, all cows had initially high *S. uberis* CFU/mL post-challenge. The number of days until *S. uberis* CFU/mL reached zero the first time post-challenge was used to establish the days until ***S. uberis* 1st clearance** phenotype. Some cows showed *S. uberis* CFU/mL increases following the initial clearance, the number of times this occurred in the 28 d sampling period was used to identify the number of ***S. uberis* shedding cycles**, which defined the second novel phenotype. The remaining two *S. uberis* phenotypes were developed from calculating the area under the curve (AUC) for *S. uberis* CFU/mL for 0-7 days (**7d AUC**) and 0-28 days (**28d AUC**) post-challenge from linear representations of CFU/mL over time for each cow using SAS (Cary, NC). AUCs were calculated in order to have a single value for each cow to test for genetic associations, since looking at effects over time is not possible with a GWAS approach.

Mammary and milk scores on the challenge quarter were evaluated on a 0 to 3 scale by ETREC Little River Staff to make 2 additional novel phenotypes. Mammary scores were defined as follows: 0 = normal, 1 = slight swelling, 2 = moderate swelling, 3 = edema. Milk scores were defined as follows: 0 = normal, 1 = small flakes or clots, 2 = large clots, 3 = stringy, off color, watery, or bloody. To focus on cows with a quick and severe inflammatory response to challenge, mammary and milk scores for the first 7 d post-challenge were summed to create the **severity score** phenotype. Contrastingly, to concentrate on cows with a more slow and steady, chronic response to challenge, mammary and milk scores were summed for the entire 28 d sampling period to create the **chronic score** phenotype.

Of note, select cows were sacrificed throughout the post-challenge sampling as part of the SUAM vaccine trial which resulted in incomplete data sets to generate phenotypes for some cows. As a result, the number of cows per phenotype differed as follows: *S. uberis* 1st clearance had N=34, *S. uberis* shedding cycles had N=34, 7d AUC had N=36, 28d AUC had N=35, severity score had N=36, and chronic score had N=34.

Association Analysis

Blood samples were collected from each cow via jugular puncture and DNA was isolated using UltraClean BloodSpin DNA Isolation Kit (Mo Bio, Carlsbad, CA). Resulting DNA was used to perform a 50K SNP chip analysis using the BovineSNP50 DNA Analysis BeadChip (Illumina, San Diego, CA) (Gunderson et al. 2005; Steemers et al. 2006). The SNP chip has 54,609 SNPs that span all 29 autosomes of the bovine genome with an average SNP spacing of 49.4kb and a median spacing of 36.9kb. Associations were tested between developed phenotypes and SNPs using Plink (Purcell et al. 2007). Effects of sex were excluded from the analysis because all individuals were female and SNPs did not represent sex chromosomes. The p-value cutoffs of $p < 1.95 \times 10^{-5}$ for highly significant SNPs and $p = 1.95 \times 10^{-5}$ to 9.34×10^{-5} for suggestively significant SNPs were set based on guidelines for reducing false positives and increasing reproducibility (Manly 2005). For identified SNPs to be reported, the minor allele had to be represented by more than one cow, unless the SNP aligned with prior evidence for mastitis or mastitis related traits. A complete index of identified SNPs, including those removed based on the minor allele criteria, is included as an appendix (Table A-1). To allow depiction of the numerous SNPs, their genome locations, and their variable association p-values, generated Plink results were visualized using Manhattan plots created with Integrated Genome Viewer (IGV) (Robinson et al. 2011). All reported gene and SNP locations were reported based on the *Bos taurus* UMD_3.1.1 (GCF_000003055.6) genome build.

Results

A total of 29 SNPs associated with unique novel phenotypes generated from *S. uberis* experimental challenge data were identified (Table 2). The identified SNPs span 18 chromosomes, with 2 SNPs each on BTAs 1, 3, 13, 14, 17, 18, 21, 22, and 28. The first *S. uberis* CFU based phenotype, *S. uberis* 1st clearance, enumerated the days until *S. uberis* CFU reached zero post-challenge. Days until *S. uberis* 1st clearance ranged from 4 to 47 days, and had a median of 23.5 days (Figure 6). Association analyses identified 3 suggestively significant ($p < 9.34 \times 10^{-5}$) SNPs associated with *S. uberis* 1st clearance time (Figure 7). One significant SNP each was identified on BTA 2 (rs41645172), 11 (rs41610542), and 28 (rs41653441), with the SNP on BTA 28 being located within the PSAP (prosaposin) gene.

The second *S. uberis* CFU based phenotype, *S. uberis* shedding cycles, counted the number of times *S. uberis* CFU reached zero within the 28 day post-challenge sampling period. The number of *S. uberis* cycles ranged from 1 to 4, with the majority of cows (47%) being represented by 2 cycles (Figure 8). Three suggestively significant SNPs ($p < 9.34 \times 10^{-5}$) were identified as associated with the *S. uberis* shedding cycles phenotype (Figure 9). One significant SNP each was identified on BTA 4 (rs41576873), 15 (rs109427689), and 17 (rs41570578). The SNP identified on BTA 4 is located within the SND1 (staphylococcal nuclease and tudor domain containing 1) gene and the SNP on BTA 17 is located in the HECTD4 (HECT domain E3 ubiquitin protein ligase 4) gene.

The last two *S. uberis* CFU based phenotypes were a quantitative representation of CFU AUC over time, either for the first 7 days post-challenge (7d AUC) or the entire 28 day sampling period (28d AUC). Areas under the curve for 7 days post-challenge ranged from 24.5 CFU*d to 1.27×10^8 CFU*d, with a median of 3.17×10^7 CFU*d (Figure 10). Association analyses identified 9 highly significant ($p < 1.95 \times 10^{-5}$) SNPs and 6 suggestively significant ($p = 1.95 \times 10^{-5}$ to 9.34×10^{-5}) SNPs associated with 7d AUC (Figure 11). The 7d AUC associated SNPs were located on BTAs 1 (rs110974478), 3 (rs43709346 and rs109880729), 5 (rs43423752), 9 (rs41665257), 12 (rs110866336), 13 (rs110804434), 17 (rs41831243), 18 (rs109290019 and rs109070561), 21 (rs111021268 and rs42651398), 22 (rs109747337 and rs110942101), and 26 (rs110747207). Five of these identified SNPs are within genes with known functions: rs110866336 is within HS6ST3 (heparan sulfate 6-O-sulfotransferase 3), rs110804434 is within TMEM189 (transmembrane protein 189), rs41831243 is within FBXW7 (F-box and WD repeat domain containing 7), rs111021268 is within FSD2 (fibronectin type III and SPRY domain containing 2), and rs110942101 is within CCR4 (C-C motif chemokine receptor 4).

Table 2: Identified loci associated with developed novel phenotypes.

Trait	rs	Locus (Chr: bp) ¹	SNP type	Allele Frequencies ²	Nearest gene(s)	Prior Evidence	P-value
Loci with prior evidence							
7d AUC	rs41665257	9: 4,581,947	5' upstream	C: 0.08 T: 0.92	LOC104968410	CM QTL (Lund et al., 2008)	6.56x10 ⁻⁶
7d AUC	rs43709346	3: 74,836,184	intergenic	A: 0.98 G: 0.02	PTGER3, CTH	CM QTL (Klungland et al., 2001)	7.64x10 ⁻⁶
7d AUC	rs43423752	5: 2,925,015	3' downstream	C: 0.98 T: 0.02	TRHDE	MY QTL (Daetwyler et al., 2008)	7.64x10 ⁻⁶
7d AUC	rs111021268	21: 23,530,389	Intron variant	C: 0.95 T: 0.05	FSD2	SCS QTL (Schulman et al., 2004)	7.64x10 ⁻⁶
7d AUC	rs109880729	3: 27,600,325	Intergenic	C: 0.05 T: 0.95	NHLH2, CASQ2	MY QTLs (Viitala et al., 2003, Ashwell et al., 2004)	3.10x10 ⁻⁵
Severity Score	rs110199901	14: 2,524,432	Intron variant	A: 0.48 G: 0.52	ZNF696	MY QTLs, SCS QTL (Viitala et al., 2003, Ashwell et al., 2004)	3.65x10 ⁻⁵
<i>S. uberis</i> 1 st clearance	rs41610542	11: 7,625,345	3' downstream	A: 0.78 G: 0.22	TMEM182	SCC QTL (Schnabel et al., 2005)	4.27x10 ⁻⁵
Severity Score	rs55617160	14: 4,468,478	Intron variant	A: 0.40 G: 0.60	TRAPPC9	MY QTL, SCS QTL (Kaupe et al., 2007)	7.04x10 ⁻⁵
<i>S. uberis</i> shedding cycles	rs109427689	15: 25,622,211	Intron variant	A: 0.50 C: 0.50	LOC101908225	SCS QTL (Boichard et al., 2003)	7.75x10 ⁻⁵
Newly identified loci							
7d AUC	rs41831243	17: 5,221,649	Intron variant	C: 0.08 T: 0.92	FBXW7	--	2.68x10 ⁻⁶
7d AUC	rs110804434 [†]	13: 78,901,415	Intron variant	A: 0.41 G: 0.59	TMEM189*	--	4.96x10 ⁻⁶
7d AUC	rs110866336	12: 77,746,796	Intron variant	A: 0.78 G: 0.22	HS6ST3	--	5.21x10 ⁻⁶
7d AUC	rs109747337	22: 58,790,629	Intergenic	A: 0.93 G: 0.07	CHCHD4, WNT7A	--	9.39x10 ⁻⁶

Table 2: Continued

Trait	rs	Locus (Chr: bp) ¹	SNP type	Allele Frequencies ²	Nearest gene(s)	Prior Evidence	P-value
7d AUC	rs109290019~	18: 53,260,898	Intergenic	A: 0.57 G: 0.43	TRAPPC6A, EXOC3L2	--	1.94x10 ⁻⁵
Chronic Score	rs43273037	1: 134,069,561	Intron variant	G: 0.59 T: 0.41	PCCB	--	2.00x10 ⁻⁵
Severity Score	rs109788678	13: 11,784,421	Intron variant	C: 0.67 T: 0.33	CAMK1D	--	2.40x10 ⁻⁵
7d AUC	rs42651398	21: 39,716,280	3' downstream	C: 0.49 T: 0.51	FOXG1	--	2.46x10 ⁻⁵
7d AUC	rs110942101	22: 7,385,040	Intron variant	G: 0.07 T: 0.93	CCR4	--	3.10x10 ⁻⁵
7d AUC	rs109070561	18: 53,281,931	Intron variant	A: 0.46 G: 0.54	LOC508455	--	3.96x10 ⁻⁵
28d AUC	rs110804434 [♦]	13: 78,901,415	Intron variant	A: 0.41 G: 0.59	TMEM189*	--	4.52x10 ⁻⁵
<i>S. uberis</i> 1 st clearance	rs41645172	2: 109,778,246	5' upstream	A: 0.69 G: 0.31	EPHA4	--	5.74x10 ⁻⁵
<i>S. uberis</i> shedding cycles	rs41576873	4: 92,742,789	Intron variant	A: 0.62 G: 0.38	SND1	--	6.12x10 ⁻⁵
7d AUC	rs110974478	1: 150,141,293	Intergenic	A: 0.09 G: 0.91	SETD4, CBR3	--	6.15x10 ⁻⁵
28d AUC	rs110693161	24: 48,872,488	Intron variant	A: 0.20 G: 0.80	CTIF	--	6.73x10 ⁻⁵
<i>S. uberis</i> 1 st clearance	rs41653441	28: 28,149,081	Intron variant	C: 0.51 T: 0.49	PSAP*	--	7.04x10 ⁻⁵
7d AUC	rs110747207	26: 50,933,887	Non-coding transcript variant	A: 0.17 G: 0.83	LOC100850437	--	7.14x10 ⁻⁵
<i>S. uberis</i> shedding cycles	rs41570578	17: 64,142,204	Intron variant	G: 0.70 T: 0.30	HECTD4	--	7.47x10 ⁻⁵
28d AUC	rs109290019~	18: 53,260,898	Intergenic	A: 0.57 G: 0.43	TRAPPC6A, EXOC3L2	--	8.91x10 ⁻⁵
Chronic Score	rs110599808	28: 247,173	Intron variant	A: 0.43 G: 0.57	RAB4A	--	9.70x10 ⁻⁵

¹Reported positions are based on the *Bos taurus* UMD_3.1.1 (GCF_000003055.6) genome build.

²Allele frequencies are representative of the current study population.

*Gene is a partial start transcript

[♦]~Indicate SNPs identified using more than one novel phenotype, each symbol is unique to one rs number.

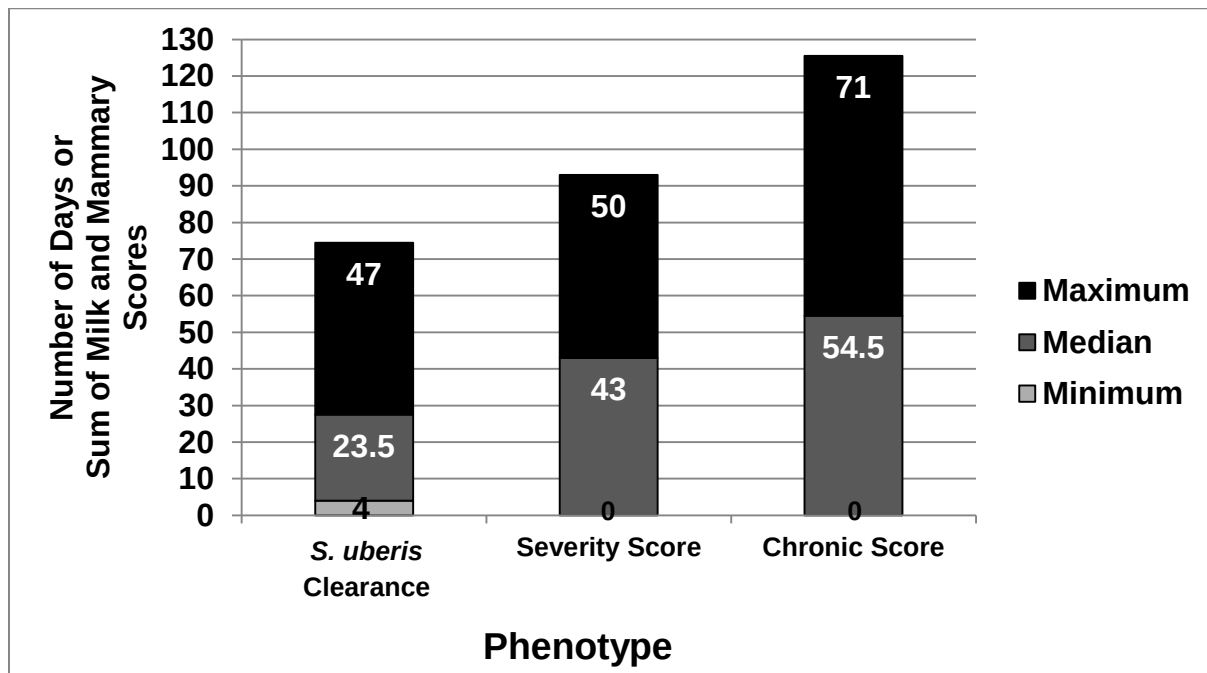


Figure 6: Minimum, median, and maximum for *S. uberis* 1st clearance, severity score, and chronic score phenotypes. *S. uberis* 1st clearance is represented by the number of days until *S. uberis* CFU reached zero based on milk culturing. Severity and chronic scores are represented by the sum of milk and mammary scores over the initial 7 d post-challenge or the entire 28 d challenge period, respectively. Minimums for both severity and chronic score were zero.

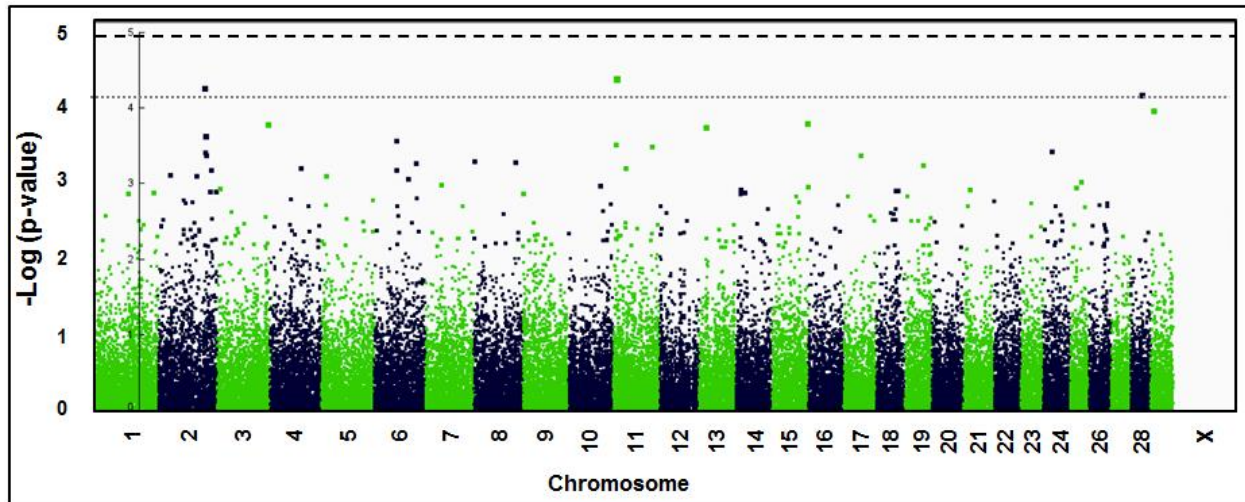


Figure 7: Manhattan plot of GWAS for *S. uberis* 1st clearance phenotype using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) in Holsteins (N = 34). The x-axis shows chromosome and SNP order within a chromosome. The y-axis show the $-\log_{10}(\text{P-value})$ for each SNP. P-value thresholds are indicated by dashed lines: top black line $p = 1.95 \times 10^{-5}$; bottom gray line $p = 9.34 \times 10^{-5}$.

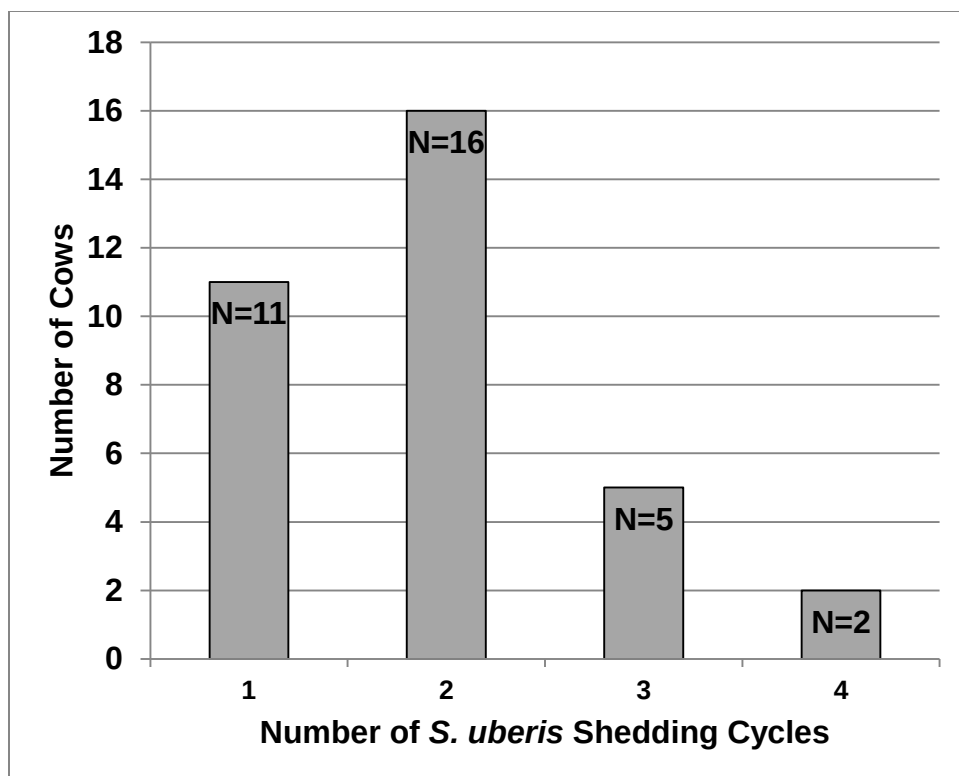


Figure 8: Number of cows per *S. uberis* shedding cycle category, where the x-axis indicates the number of times *S. uberis* CFU reached zero within the 28 day sampling period.

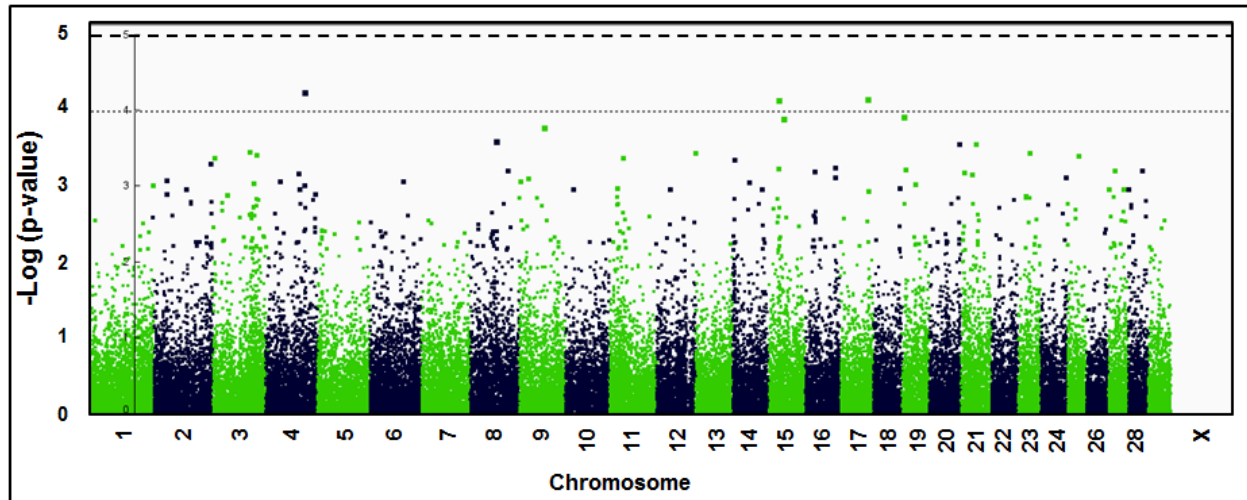


Figure 9: Manhattan plot of GWAS for *S. uberis* shedding cycles phenotype using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) in Holsteins (N = 34). The x-axis shows chromosome and SNP order within a chromosome. The y-axis show the $-\log_{10}(P\text{-value})$ for each SNP. P-value thresholds are indicated by dashed lines: top black line $p = 1.95 \times 10^{-5}$; bottom gray line $p = 9.34 \times 10^{-5}$.

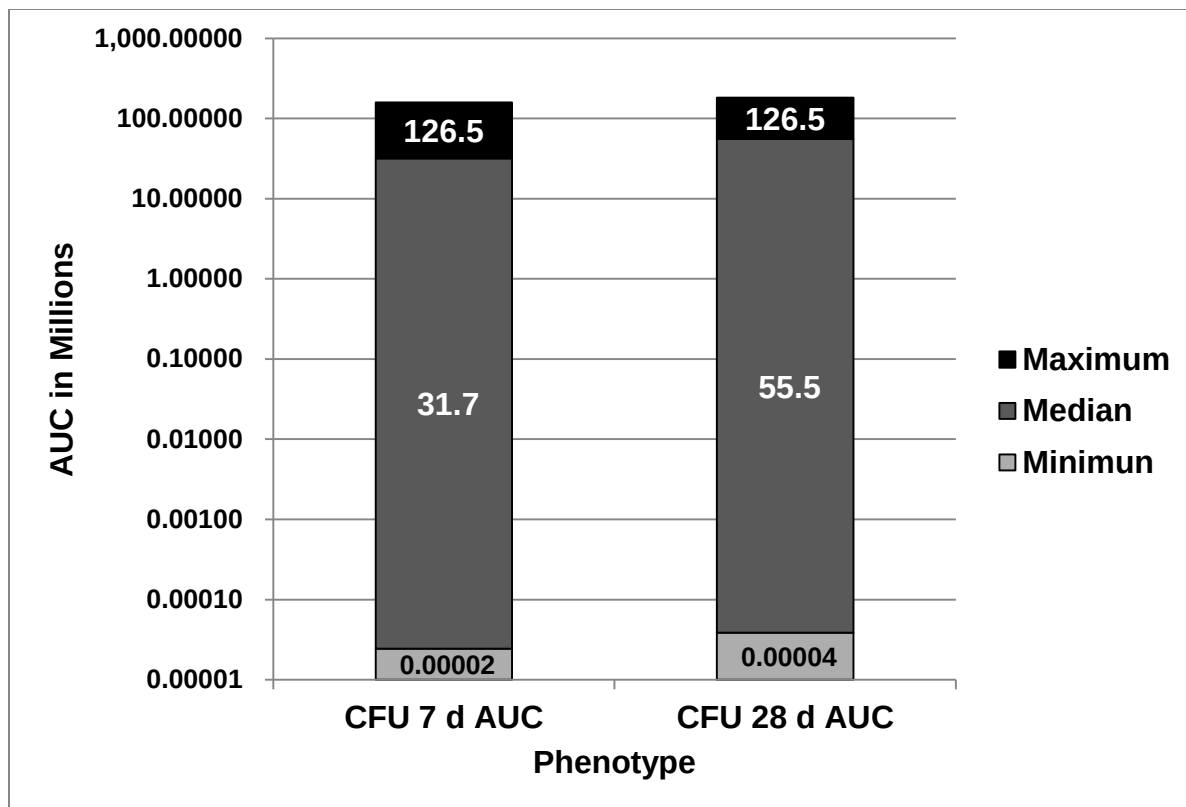


Figure 10: Minimum, median, and maximum *S. uberis* CFU AUC values for both 7 and 28 d phenotypes displayed in millions of CFU*d.

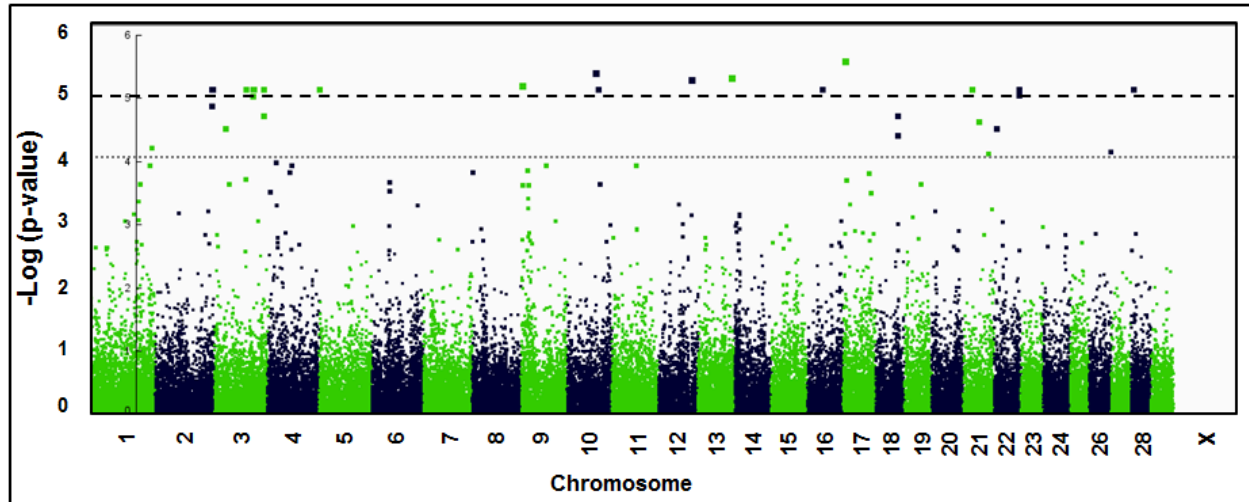


Figure 11: Manhattan plot of GWAS for 7 d *S. uberis* CFU AUC phenotype using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) in Holsteins (N = 36). The x-axis shows chromosome and SNP order within a chromosome. The y-axis show the $-\log_{10}(\text{P-value})$ for each SNP. P-value thresholds are indicated by dashed lines: top black line $p = 1.95 \times 10^{-5}$; bottom gray line $p = 9.34 \times 10^{-5}$.

Areas under the curve for the 28 day time period ranged from 38.5 CFU*d to 1.27×10^8 CFU*d, with a median of 5.55×10^7 CFU*d (Figure 10). Association analyses identified 3 suggestively significant ($p < 9.34 \times 10^{-5}$) associated with 28 d *S. uberis* AUC (Figure 12). Two of the identified SNPs (rs110804434 on BTA 15 and rs109290019 on BTA 18) were also identified using the 7d AUC phenotype. The remaining identified SNP (rs110693161) is on BTA 24 and located within the CTIF (cap binding complex dependent translation initiation factor) gene.

The next two phenotypes were generated using sums of mammary and milk scores to identify levels of inflammation. Severity score focused on cows displaying acute clinical signs post-challenge by summing their scores for the first 7 days after challenge. Severity scores ranged from 0 to 50, with a median of 43 (Figure 6). Analyses identified 3 suggestively significant ($p < 9.34 \times 10^{-5}$) SNPs associated with severity score located on BTAs 13 (rs109788678), 14 (rs110199901 and rs55617160) (Figure 13). All three SNPs are within genes with known functions. The BTA 13 SNP, rs109788678, is within the CAMK1D (calcium/calmodulin dependent protein kinase 1D) gene. The SNPs on BTA 14, rs110199901 and rs55617160 are within the ZNF696 (zinc finger protein 696) gene and the TRAPPC9 (trafficking protein particle complex 9) gene, respectively. Chronic score focused on chronically symptomatic cows by summing scores across the entire 28 day sampling period. Chronic scores ranged from 0 to 71 and had a median of 54.5 (Figure 6). Chronic score identified 2 suggestively significant SNPs ($p < 9.34 \times 10^{-5}$), with one SNP each on BTA 1 (rs43273037) and 28 (rs110599808) (Figure 14). Both SNPs are within known genes; rs43273037 is located within the PCCB (Propionyl CoA carboxylase, beta polypeptide) gene and rs110599808 is within the RAB4A (RAB4A, member RAS oncogene family) gene.

Discussion

This study successfully identified associations between SNPs in the bovine genome and novel phenotypes generated from data obtained using an *S. uberis* challenge model. Our prior study (Siebert Dissertation, Chapter 3) utilized the same study design to develop SCC based phenotypes and successfully identified 16 SNPs, where 4 SNPs were located directly in potential candidate genes, 3 more SNPs were in quantitative trait loci (QTLs) previously linked to mastitis or mastitis related phenotypes, and 2 more SNPs within candidate genes and previously identified QTLs. The success of our prior study helps lend strength to the results of the present study. Additionally, 9 of the SNPs identified in this study (rs41665257, rs43709346, rs43423752, rs111021268, rs109880729, rs110199901, rs41610542, rs55617160, and rs109427689) are located in genomic regions with prior evidence associated with mastitis or related phenotypes. Considering both our studies successfully identified genomic

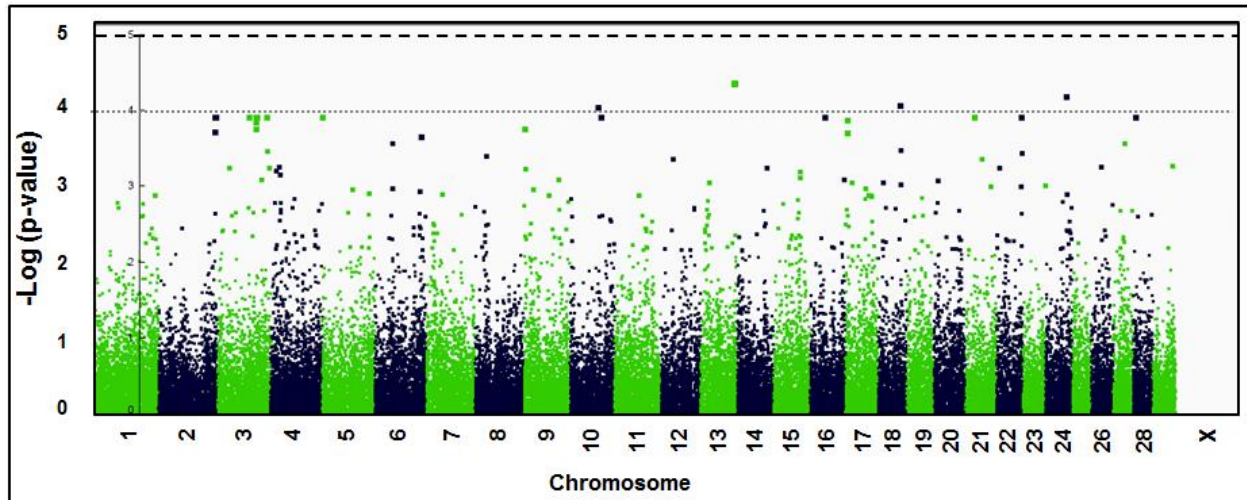


Figure 12: Manhattan plot of GWAS for 28 d *S. uberis* CFU AUC phenotype using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) in Holsteins (N = 35). The x-axis shows chromosome and SNP order within a chromosome. The y-axis show the $-\log_{10}(\text{P-value})$ for each SNP. P-value thresholds are indicated by dashed lines: top black line $p = 1.95 \times 10^{-5}$; bottom gray line $p = 9.34 \times 10^{-5}$.

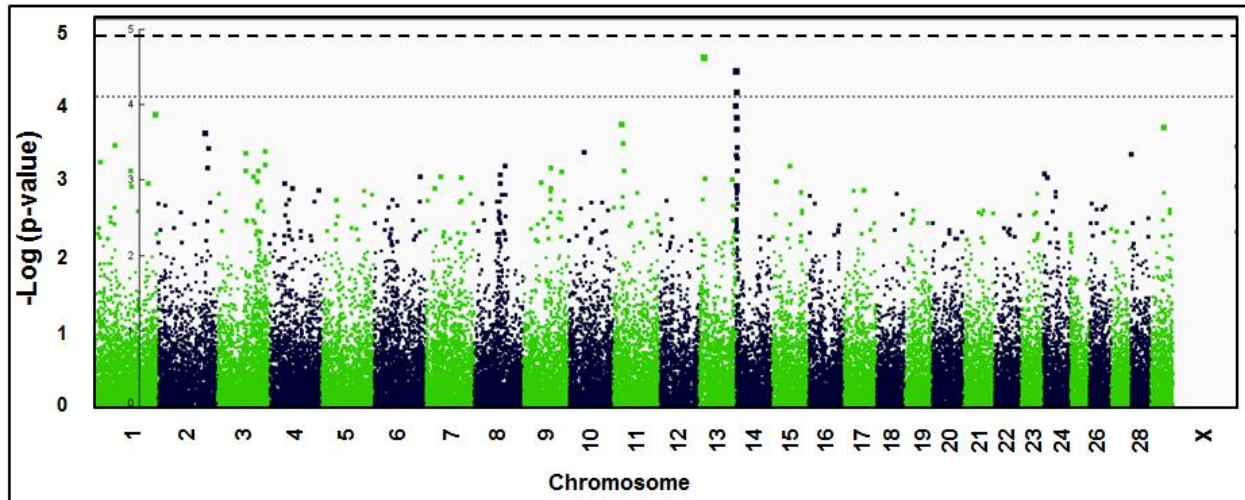


Figure 13: Manhattan plot of GWAS for severity score phenotype using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) in Holsteins (N = 36). The x-axis shows chromosome and SNP order within a chromosome. The y-axis show the $-\log_{10}(P\text{-value})$ for each SNP. P-value thresholds are indicated by dashed lines: top black line $p = 1.95 \times 10^{-5}$; bottom gray line $p = 9.34 \times 10^{-5}$.

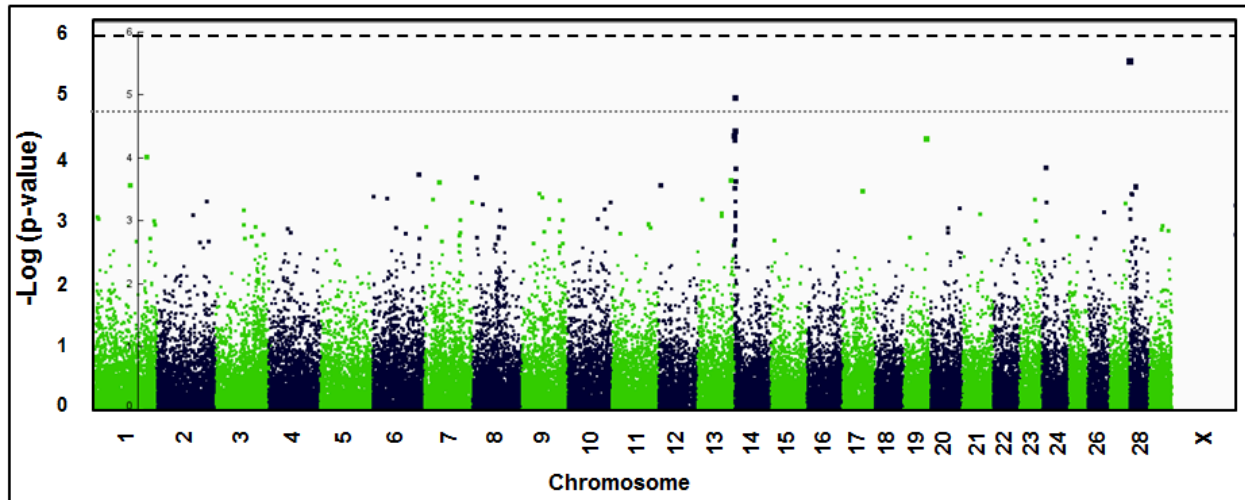


Figure 14: Manhattan plot of GWAS for chronic score phenotype using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) in Holsteins (N = 34). The x-axis shows chromosome and SNP order within a chromosome. The y-axis show the $-\log_{10}(P\text{-value})$ for each SNP. P-value thresholds are indicated by dashed lines: top black line $p = 1.95 \times 10^{-5}$; bottom gray line $p = 9.34 \times 10^{-5}$.

regions with prior evidence linked to mastitis or traits associated with mastitis, the use of novel phenotypes derived from *S. uberis* challenge data to identify genomic regions of interest for *S. uberis* mastitis may be a viable new method for identifying candidate genes.

Of the identified SNPs in the present study, two SNPs associated with the 7d AUC phenotype (rs41665257 and rs43709346) were located within QTLs for clinical mastitis (CM) (Klungland et al. 2001; Lund et al. 2008). Clinical mastitis is hallmarked by mammary gland inflammation and altered milk production. The inflammation associated with CM typically indicates a strong immune response correlating to higher recruitment of SCC. Higher SCC could lead to faster clearance of *S. uberis* from the gland and lead to lower observed 7d AUC phenotypes. Another 5 SNPs significantly associated with the phenotypes *S. uberis* first clearance (rs41610542), *S. uberis* shedding cycles phenotype (rs109427687), 7d AUC (rs111021268), and severity score (rs110199901 and rs55617160) were located within QTLs associated with SCC and SCS (Ashwell et al. 2004; Boichard et al. 2003; Kaupe et al. 2007; Schnabel et al. 2005; Schulman et al. 2004). More elevated SCC/SCS would indicate more immune cells to fight the infection and would likely correlate to faster bacteria clearance post-challenge linking SCC/SCS directly to our *S. uberis* 1st clearance and 7d AUC phenotypes. The more elevated SCC/SCS and subsequent faster bacteria clearance would also lead to more inflammation observed and higher mammary and milk scores which ties SCS to our severity score phenotype. Conversely, lower SCC/SCS would indicate fewer immune cells to fight the infection and potentially allow for the infection to persist, or even demonstrate cycling of bacteria clearance when the immune cell populations are depleted and more haven't yet reached the gland, which links SCS to our *S. uberis* cycles phenotype. Of interest, SNPs associated with 7d AUC (rs43423752 and rs109880729) and severity score (rs110199901 and rs55617160) phenotypes were located in regions previously identified as associated with milk yield (MY) QTLs (Ashwell et al. 2004; Daetwyler et al. 2008; Kaupe et al. 2007; Viitala et al. 2003). The association of mastitis incidence and milk yield has been known for decades (Meredith et al. 2012; Shaw and Beam 1935; Wilson et al. 1997) since fighting infections sequesters energy that would otherwise be partitioned to milk production. Since, higher *S. uberis* CFU would lead to higher SCC (used as an indicator of mastitis) and higher SCC causes higher mammary and milk scores, both 7d AUC and severity score could impact MY.

While the identified SNPs themselves may not be causal but rather in linkage disequilibrium is the causal variant, they do provide several interesting candidate genes based on the genes they are located within. Of the SNPs identified with prior evidence,

3 are also within known genes. First, rs111021268, identified with the 7d AUC phenotype, is located within the FSD2 (fibronectin type III and SPRY domain containing 2) gene. While there is not a specific known function for FSD2, the domains it contains (FN3 and SPRY) are components of many other proteins. The FN3 domain is an extracellular free-acting domain that provides a scaffold for the binding of a variety of ligands (Bloom and Calabro 2009; Campbell and Spitzfaden 1994). The SPRY domain is intracellular and is also involved in protein interactions, including those that regulate the innate and adaptive immune systems through transcriptional regulation and production of reactive compounds (D'Cruz et al. 2013). Inabilities to properly interact with ligand and/or mount an effective immune response would permit *S. uberis* growth in the mammary gland and lead to higher 7d AUC phenotypes observed.

The other two SNPs with prior evidence and within genes were identified using the severity score phenotype. Within the ZNF696 (zinc finger protein 696) gene is the rs110199901SNP. While no specific function is attributed to ZNF696, zinc finger proteins are frequently involved in DNA/RNA binding and subsequent regulation of transcription/translation (Laity et al. 2001). Lastly, rs55617160 is within the TRAPPC9 (trafficking protein particle complex 9). TRAPPC9 has been linked to increased cytokine-mediated NFκB activation through phosphorylation of the IKK complex (Hu et al. 2005). NFκB is a transcriptional activator of many genes including many that are critical for inflammation and immunity (Bonizzi and Karin 2004). The inability to properly regulate transcriptional and/or translational changes following an infection would likely result in lower observed inflammation and reduced immune responses despite bacteria presence, leading to the lower severity score phenotypes observed.

The *S. uberis* 1st clearance phenotype identified 2 novel SNPs, one SNP (rs41645172) on BTA 2 at 109.8Mb, and another SNP (rs41653441) on BTA 28 at 28.1Mb that is within the PSAP (prosaposin) gene. Prosaposin is the precursor that generates saposin A, B, C, and D through limited proteolytic processing (O'Brien and Kishimoto 1991). Saposins are involved in lipid degradation within the lysosome and also play a role in loading lipid antigens onto CD1 molecules (Barral and Brenner 2007). Antigen presentation by CD1 activates several sub-sets of T-cells, including CD8+ T-cells which have demonstrated importance in for killing *S. uberis* in the mammary gland (Barral and Brenner 2007; Denis et al. 2011). Additionally, the un-cleaved prosaposin precursor may play a role in regulating cell proliferation and apoptosis (Sun et al. 2010). Thus, PSAP could play a role eliciting a more efficient response to and clearance of invading pathogens and lead to lower observed *S. uberis* 1st clearance phenotype.

The *S. uberis* cycles phenotype identified two other novel SNPs in genes, one SNP (rs41576873) on BTA 4 at 92.7Mb and another SNP (rs41570578) on BTA 17 at

64.1Mb. The BTA 4 SNP is within the SND1 (staphylococcal nuclease and tudor domain containing 1) gene. SND1 is a transcriptional coactivator that interacts with many other transcriptional activators including STAT5 and STAT6 (Yang et al. 2002; Yang et al. 2007). STAT5 plays a role in preventing apoptosis of lymphocytes (Mui et al. 1995). STAT6 is involved in IL-4 signal transduction to differentiate naïve t-helper cells into Th2 cells that are essential for activating the humoral (B-cell) response (Takeda et al. 1997). Reduced lymphocyte survival and humoral response could result in fewer immune cells to fight infection. Fewer immune cells to fight infection could lead to a cycling of bacteria clearance when immune cell populations are depleted or unprimed with antigen leading to higher numbers of *S. uberis* cycles observed. The BTA 17 SNP is within the HECTD4 (HECT domain E3 ubiquitin protein ligase 4) gene. The HECT domain subfamily of a ubiquitin ligase protein helps determine the type of ubiquitin chain formed during ubiquitination (Sheng et al. 2012). Ubiquitination helps determine the biological fates of protein by targeting them for degradation, altering their cellular location, or affecting their interactions/ activity (Glickman and Ciechanover 2002; Schnell and Hicke 2003). Ubiquitination has been demonstrated to control the immune response through NFκB regulation, thus alterations in HECT domains could result in altered NFκB regulation and subsequent altered immune response (Magnani et al. 2000). Alterations in immune response could alter *S. uberis*'s ability to proliferate within the gland and possibly allow for bacterial clearance cycling as other aspects of the immune system attempt to compensate for those ubiquitination may affect.

The 7d AUC phenotype identified the most SNPs when compared to other phenotypes, with 10 SNPs being novel. Four of these novel SNPs are located in genes with known functions. One SNP, rs41831243, on BTA 17 at 5.2Mb, is in the FBXW7 (F-box and WD repeat domain containing 7) gene. FBXW7 is a subunit of the SCF ubiquitin protein ligase complex that targets mediators of the cell cycle and cell division using phosphorylation-dependent ubiquitination. Specifically, FBXW7 is critical for immature CD4+CD8+ T-cells to exit the cell cycle and for the proliferation of mature T-cells (Onoyama et al. 2007). Since CD8+ T-cells have demonstrated particular importance in for killing *S. uberis* in the mammary gland (Barral and Brenner 2007; Denis et al. 2011), inadequate T-cell maturation and proliferation could permit *S. uberis* to persist in the gland post-challenge and lead to higher the 7d AUC phenotypes observed. Another SNP on BTA 13 at 78.9Mb, rs110804434, also identified with the 28d AUC phenotype, is within the TMEM189 (transmembrane protein 189) gene. TMEM189 is closely downstream to UBE2V1 (ubiquitin-conjugating enzyme E2 variant 1) and the two genes are often expressed as a conjoined gene (Prakash et al. 2010). While there is no known function tied to the conjoined genes, it is hypothesized they play a role in gene regulation (Prakash et al. 2010). Additionally, overexpression of this

particular conjoined gene has been linked to slower degradation of other (EGF) receptor complexes via endocytosis (Duex et al. 2010). Therefore, TMEM189 could help regulate gene expression and receptor-ligand complex degradation during the infection process and lead to higher or lower observed 7d and 28d AUC observed phenotypes depending on the effect. The third SNP identified with 7d AUC, rs110866336, at 77.7Mb on BTA 12, is in the HS6ST3 (heparan sulfate 6-O-sulfotransferase 3) gene. HS6ST3 modifies heparan sulfate, found in heparan sulfate proteoglycans on the cell surface and in the extracellular matrix, to generate a variety of structures needed for protein interactions (Habuchi et al. 1995; Yanagishita and Hascall 1992). Such protein interactions have demonstrated importance for both inflammation and bacteria adherence/infection (Rostand and Esko 1997; Tanaka et al. 1993). Therefore, a reduction of modified heparan sulfates would likely lead to less inflammation and reduced bacterial adherence/infection and probably correlate with lower 7d AUC phenotypes. The final SNP (rs110942101) within a gene for 7d AUC is on BTA 22 at 7.4Mb and is within the CCR4 (C-C motif chemokine receptor) gene. CCR4 is the G-protein coupled receptor for chemokine CCL22, which is released by macrophages upon the initiation of immune responses (Imai et al. 1998). CCR4 is expressed on both CD4+ and CD8+ T-cells and its expression is upregulated after activation (i.e.: ligand binding) (D'Ambrosio et al. 1998). Receptor-ligand complexes help increase extravasation, cell migration, and effector T-cell positioning at the recruitment site (Chang et al. 1997; D'Ambrosio et al. 1998). During mastitis, significant increases in T-cell recruitment to the gland have been observed and CD8+ T-cells have demonstrated specific and substantial killing ability towards *S. uberis* (Denis et al. 2011; Mehrzad et al. 2008). Thus, alterations in CCR4 could correlate to faster clearance of invading bacteria in the mammary gland and lower observed CFU leading to lower 7d AUC phenotypes observed.

The 28d AUC phenotype identified 3 novel SNPs. Both rs109290019 and rs110804434 were also identified by 7d AUC phenotype and were discussed previously. The third SNP, rs100693161, located on BTA 24 at 48.9Mb, is located within the CTIF (cap binding complex dependent translation initiation factor) gene. CTIF is a component of the CBP80/20-dependent translation initiation complex that co-transcriptionally binds the cap end of nascent mRNAs (Kim et al. 2009). The CBP80/20 complex recognizes premature termination codons (PTCs) during translation which helps minimize expression of truncated proteins by targeting them for decay (Chang et al. 2007). A mutation in CTIF could permit a breakdown in this editing process allowing a higher expression of truncated proteins that are either non-functional or toxic to host cells, which depending on the effected proteins, could permit more growth of *S. uberis* and lead to higher observed 28d AUC phenotypes.

The severity score phenotype identified only one novel SNP located on BTA 13 at 11.8Mb and within the CAMK1D (calcium/calmodulin dependent protein kinase 1D) gene. CAMK1D is a component of the calcium-regulated calmodulin-dependent protein kinase cascade which has been implicated in the migration of neutrophils following chemoattractant stimulation (Verploegen et al. 2002). Inefficiencies in stimulating neutrophil migration would lead to lower levels of inflammation observed at the site of infection and result in the observation of lower severity score phenotypes.

Both of the novel SNPs identified by the chronic score phenotype are located within genes. Within the PCCB (propionyl CoA carboxylase, beta polypeptide) gene, on BTA 1 at 134.1Mb, is rs43273037. PCCB is a subunit of propionyl CoA carboxylase which is involved in the metabolism of propionic acid that has no definitive connections to the chronic score phenotype (Tietz and Ochoa 1959). However, propionic acid is one of the main energy sources used by dairy cattle and energy status has demonstrated importance for susceptibility to mastitis (Knegsel et al. 2005; Rezamand et al. 2007). Thus, reduced ability to metabolize propionic acid could decrease available energy and influence susceptibility to mastitis and potentially allow for infections to persist and cause increased chronic scores observed. Lastly, rs110599808 is on BTA 28 at 0.2Mb and is in the RAB4A (RAB4A, member RAS oncogene family) gene. RAB4A is a member of the RAS superfamily of small GTPases that are involved in regulating membrane trafficking. Specifically, RAB4A regulates the secretion of α -granules from platelets (Shirakawa et al. 2000). Platelet α -granules contain P-selectin which is a vital cell adhesion molecule needed for monocyte and neutrophil extravasation which is required for cell recruitment to the site of infection (Furie et al. 2001). Improper extravasation would result in fewer immune cells at the infection site and result in lower levels of inflammation and lower chronic scores observed.

In conclusion, this study was successful in identifying SNPs associated with novel phenotypes generated from data collected following direct *S. uberis* experimental challenge. Considering 31% of the SNPs identified using our novel phenotypes are in regions with prior evidence linked to mastitis, the probability that our remaining identified loci are valid associations is high. The majority of the identified SNPs are also located within known genes where 17% have functions linked to immune functions and 28% have roles in the regulation of gene or protein expression. Associations with immune function and gene regulation are logical since recognition of *S. uberis* within the mammary gland would cause the upregulation of many genes, especially those linked to immune processes. Thus, the investigation of these candidate genes could potentially uncover novel treatment or preventative strategies for *S. uberis* mastitis or serve as selection markers for cows less susceptible to *S. uberis* mastitis. Furthermore, the

possibility that these genes could be associated with other causative pathogens exists since the identified genes seem to have links to more generic immune processes rather than to *S. uberis*'s unique pathology.

Acknowledgements

The authors would like to thank Dr. Arnold Saxton of the department of Animal Science at the University of Tennessee for his assistance in converting the data format. The challenge portion of this study was supported by USDA-NIFA (2011-67015-30168). Additional support was obtained from the UT AgResearch AgInnovation Fund for SNP chip analyses.

References

- Ashwell, M. S., Heyen, D. W., Sonstegard, T. S., Van Tassell, C. P., Da, Y., VanRaden, P. M., Ron, M., Weller, J. I., and Lewin, H. A. (2004). "Detection of quantitative trait loci affecting milk production, health, and reproductive traits in Holstein cattle." *Journal of Dairy Science*, 87(2), 468-475.
- Barral, D. C., and Brenner, M. B. (2007). "CD1 antigen presentation: How it works." *Nature Reviews Immunology*, 7(12), 929-941.
- Bloom, L., and Calabro, V. (2009). "FN3: A new protein scaffold reaches the clinic." *Drug Discovery Today*, 14(19–20), 949-955.
- Boichard, D., Grohs, C., Bourgeois, F., Cerqueira, F., Faugeras, R., Neau, A., Rupp, R., Amigues, Y., Boscher, M. Y., and Levéziel, H. (2003). "Detection of genes influencing economic traits in three French dairy cattle breeds." *Genetics, Selection, Evolution : GSE*, 35(1), 77-101.
- Bonizzi, G., and Karin, M. (2004). "The two NF-kappaB activation pathways and their role in innate and adaptive immunity." *Trends in Immunology*, 25(6), 280-8.
- Campbell, I. D., and Spitzfaden, C. (1994). "Building proteins with fibronectin type III modules." *Structure*, 2(5), 333-337.
- Chang, M., McNinch, J., Elias, C., Manthey, C. L., Grosshans, D., Meng, T., Boone, T., and Andrew, D. P. (1997). "Molecular cloning and functional characterization of a novel cc chemokine, stimulated T cell chemotactic protein (STCP-1) that specifically acts on activated T lymphocytes." *Journal of Biological Chemistry*, 272(40), 25229-25237.
- Chang, Y. F., Imam, J. S., and Wilkinson, M. F. (2007). "The nonsense-mediated decay RNA surveillance pathway." *Annual Review Biochemistry*, 76, 51-74.
- D'Cruz, A. A., Babon, J. J., Norton, R. S., Nicola, N. A., and Nicholson, S. E. (2013). "Structure and function of the SPRY/B30.2 domain proteins involved in innate immunity." *Protein Science*, 22(1), 1-10.
- D'Ambrosio, D., Iellem, A., Bonecchi, R., Mazzeo, D., Sozzani, S., Mantovani, A., and Sinigaglia, F. (1998). "Cutting edge: selective up-regulation of chemokine receptors CCR4 and CCR8 upon activation of polarized human type 2 Th cells." *The Journal of Immunology*, 161(10), 5111-5115.
- Daetwyler, H. D., Schenkel, F. S., Sargolzaei, M., and Robinson, J. A. B. (2008). "A genome scan to detect quantitative trait loci for economically important traits in holstein cattle using two methods and a dense single nucleotide polymorphism map." *Journal of Dairy Science*, 91(8), 3225-3236.

- Denis, M., Lacy-Hulbert, S. J., Buddle, B. M., Williamson, J. H., and Wedlock, D. N. (2011). "*Streptococcus uberis*-specific T cells are present in mammary gland secretions of cows and can be activated to kill *S. uberis*." *Veterinary Research Communications*, 35(3), 145-156.
- Duex, J. E., Mullins, M. R., and Sorkin, A. (2010). "Recruitment of Uev1B to Hrs-containing endosomes and its effect on endosomal trafficking." *Experimental Cell Research*, 316(13), 2136-2151.
- Furie, B., Furie, B. C., and Flaumenhaft, R. (2001). "A journey with platelet P-selectin: the molecular basis of granule secretion, signalling and cell adhesion." *Thrombosis and Haemostasis*, 86(1), 214-221.
- Glickman, M. H., and Ciechanover, A. (2002). "The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction." *Physiological Reviews*, 82(2), 373-428.
- Habuchi, H., Habuchi, O., and Kimata, K. (1995). "Purification and characterization of heparan sulfate 6-sulfotransferase from the culture medium of Chinese hamster ovary cells." *Journal of Biological Chemistry*, 270(8), 4172-4179.
- Hu, W. H., Pendergast, J. S., Mo, X. M., Brambilla, R., Bracchi-Ricard, V., Li, F., Walters, W. M., Blits, B., He, L., Schaal, S. M., and Bethea, J. R. (2005). "NIBP, a novel NIK and IKK(beta)-binding protein that enhances NF-(kappa)B activation." *Journal of Biological Chemistry*, 280(32), 29233-41.
- Imai, T., Chantry, D., Raport, C. J., Wood, C. L., Nishimura, M., Godiska, R., Yoshie, O., and Gray, P. W. (1998). "Macrophage-derived chemokine is a functional ligand for the CC chemokine receptor 4." *Journal of Biological Chemistry*, 273(3), 1764-1768.
- Kaupe, B., Brandt, H., Prinzenberg, E. M., and Erhardt, G. (2007). "Joint analysis of the influence of CYP11B1 and DGAT1 genetic variation on milk production, somatic cell score, conformation, reproduction, and productive lifespan in German Holstein cattle." *Journal of Animal Science*, 85(1), 11-21.
- Kim, K. M., Cho, H., Choi, K., Kim, J., Kim, B. W., Ko, Y. G., Jang, S. K., and Kim, Y. K. (2009). "A new MIF4G domain-containing protein, CTIF, directs nuclear cap-binding protein CBP80/20-dependent translation." *Genes & Development* 23(17), 2033-45.
- Klungland, H., Sabry, A., Heringstad, B., Olsen, H. G., Gomez-Raya, L., Våge, D. I., Olsaker, I., Ødegård, J., Klemetsdal, G., Schulman, N., Vilkki, J., Ruane, J., Aasland, M., Rønningen, K., and Lien, S. (2001). "Quantitative trait loci affecting clinical mastitis and somatic cell count in dairy cattle." *Mammalian Genome*, 12(11), 837-842.

- Knegsel, A. T. M. v., Brand, H. v. d., Dijkstra, J., Tamminga, S., and Kemp, B. (2005). "Effect of dietary energy source on energy balance, production, metabolic disorders and reproduction in lactating dairy cattle." *Reproduction Nutrition Development*, 45(6), 665-688.
- Laity, J. H., Lee, B. M., and Wright, P. E. (2001). "Zinc finger proteins: new insights into structural and functional diversity." *Current Opinion in Structural Biology*, 11(1), 39-46.
- Lund, M. S., Guldbrandtsen, B., Buitenhuis, A. J., Thomsen, B., and Bendixen, C. (2008). "Detection of quantitative trait loci in Danish Holstein cattle affecting clinical mastitis, somatic cell score, udder conformation traits, and assessment of associated effects on milk yield." *Journal of Dairy Science*, 91(10), 4028-4036.
- Magnani, M., Crinelli, R., Bianchi, M., and Antonelli, A. (2000). "The ubiquitin-dependent proteolytic system and other potential targets for the modulation of nuclear factor- κ B (NF- κ B)." *Current Drug Targets*, 1(4), 387-399.
- Mehrzhad, J., Janssen, D., Duchateau, L., and Burvenich, C. (2008). "Increase in *Escherichia coli* inoculum dose accelerates CD8⁺ T-cell trafficking in the primiparous bovine mammary gland." *Journal of Dairy Science*, 91(1), 193-201.
- Meredith, B. K., Kearney, F. J., Finlay, E. K., Bradley, D. G., Fahey, A. G., Berry, D. P., and Lynn, D. J. (2012). "Genome-wide associations for milk production and somatic cell score in Holstein-Friesian cattle in Ireland." *BMC Genetics*, 13(1), 1-11.
- Mui, A. L., Wakao, H., Harada, N., O'Farrell, A. M., and Miyajima, A. (1995). "Interleukin-3, granulocyte-macrophage colony-stimulating factor, and interleukin-5 transduce signals through two forms of STAT5." *Journal of Leukocyte Biology*, 57(5), 799-803.
- O'Brien, J. S., and Kishimoto, Y. (1991). "Saposin proteins: Structure, function, and role in human lysosomal storage disorders." *The FASEB Journal*, 5(3), 301-8.
- Onoyama, I., Tsunematsu, R., Matsumoto, A., Kimura, T., de Alborán, I. M., Nakayama, K., and Nakayama, K. I. (2007). "Conditional inactivation of Fbxw7 impairs cell-cycle exit during T cell differentiation and results in lymphomatogenesis." *The Journal of Experimental Medicine*, 204(12), 2875-2888.
- Prakash, T., Sharma, V. K., Adati, N., Ozawa, R., Kumar, N., Nishida, Y., Fujikake, T., Takeda, T., and Taylor, T. D. (2010). "Expression of conjoined genes: Another mechanism for gene regulation in eukaryotes." *PLOS ONE*, 5(10), e13284.
- Rezamand, P., Hoagland, T. A., Moyes, K. M., Silbart, L. K., and Andrew, S. M. (2007). "Energy status, lipid-soluble vitamins, and acute phase proteins in periparturient

- Holstein and Jersey dairy cows with or without subclinical mastitis." *Journal of Dairy Science*, 90(11), 5097-5107.
- Rostand, K. S., and Esko, J. D. (1997). "Microbial adherence to and invasion through proteoglycans." *Infection and Immunity*, 65(1), 1.
- Schnabel, R. D., Sonstegard, T. S., Taylor, J. F., and Ashwell, M. S. (2005). "Whole-genome scan to detect QTL for milk production, conformation, fertility and functional traits in two US Holstein families." *Animal Genetics*, 36(5), 408-416.
- Schnell, J. D., and Hicke, L. (2003). "Non-traditional functions of ubiquitin and ubiquitin-binding proteins." *Journal of Biological Chemistry*, 278(38), 35857-35860.
- Schulman, N. F., Viitala, S. M., de Koning, D. J., Virta, J., Mäki-Tanila, A., and Vilkki, J. H. (2004). "Quantitative trait loci for health traits in Finnish Ayrshire cattle." *Journal of Dairy Science*, 87(2), 443-449.
- Shaw, A. O., and Beam, A. L. (1935). "The effect of mastitis upon milk production." *Journal of Dairy Science*, 18(6), 353-357.
- Sheng, Y., Hong, J. H., Doherty, R., Srikumar, T., Shloush, J., Avvakumov, G. V., Walker, J. R., Xue, S., Neculai, D., Wan, J. W., Kim, S. K., Arrowsmith, C. H., Raught, B., and Dhe-Paganon, S. (2012). "A human ubiquitin conjugating enzyme (E2)-HECT E3 ligase structure-function screen." *Molecular & Cellular Proteomics*, 11(8), 329-341.
- Shirakawa, R., Yoshioka, A., Horiuchi, H., Nishioka, H., Tabuchi, A., and Kita, T. (2000). "Small GTPase Rab4 regulates Ca²⁺-induced α -granule secretion in platelets." *Journal of Biological Chemistry*, 275(43), 33844-33849.
- Sun, L., Wang, S., Hu, C., and Zhang, X. (2010). "Regulation of cell proliferation and apoptosis through fibrocystin–prosaposin interaction." *Archives of Biochemistry and Biophysics*, 502(2), 130-136.
- Takeda, K., Kishimoto, T., and Akira, S. (1997). "STAT6: its role in interleukin 4-mediated biological functions." *Journal of Molecular Medicine*, 75(5), 317-326.
- Tanaka, Y., Adams, D. H., Hubscher, S., Hirano, H., Siebenlist, U., and Shaw, S. (1993). "T-cell adhesion induced by proteoglycan-immobilized cytokine MIP-1 beta." *Nature*, 361(6407), 79.
- Tietz, A., and Ochoa, S. (1959). "Metabolism of propionic acid in animal tissues: V. Purification and properties of propionyl carboxylase." *Journal of Biological Chemistry*, 234(6), 1394-1400.

- Verploegen, S., van Leeuwen, C. M., van Deutekom, H. W. M., Lammers, J. W. J., Koenderman, L., and Coffey, P. J. (2002). "Role of Ca²⁺/calmodulin regulated signaling pathways in chemoattractant induced neutrophil effector functions." *European Journal of Biochemistry*, 269(18), 4625-4634.
- Viitala, S. M., Schulman, N. F., de Koning, D. J., Elo, K., Kinoshita, R., Virta, A., Virta, J., Mäki-Tanila, A., and Vilkki, J. H. (2003). "Quantitative trait loci affecting milk production traits in Finnish Ayrshire dairy cattle." *Journal of Dairy Science*, 86(5), 1828-36.
- Wilson, D. J., Gonzalez, R. N., and Das, H. H. (1997). "Bovine mastitis pathogens in New York and Pennsylvania: Prevalence and effects on somatic cell count and milk production." *Journal of Dairy Science*, 80(10), 2592-2598.
- Yanagishita, M., and Hascall, V. C. (1992). "Cell surface heparan sulfate proteoglycans." *Journal of Biological Chemistry*, 267.
- Yang, J., Aittomäki, S., Pesu, M., Carter, K., Saarinen, J., Kalkkinen, N., Kieff, E., and Silvennoinen, O. (2002). "Identification of p100 as a coactivator for STAT6 that bridges STAT6 with RNA polymerase II." *The EMBO Journal*, 21(18), 4950-4958.
- Yang, J., Välineva, T., Hong, J., Bu, T., Yao, Z., Jensen, O. N., Frilander, M. J., and Silvennoinen, O. (2007). "Transcriptional co-activator protein p100 interacts with snRNP proteins and facilitates the assembly of the spliceosome." *Nucleic Acids Research*, 35(13), 4485-4494.

Chapter V: Investigation of *Bos taurus* CXCR1 gene identifies loci with potential mastitis implications using GWAS and linkage analyses

Investigation of *Bos taurus* CXCR1 gene identifies loci with potential mastitis implications using GWAS and linkage analyses

L. Siebert¹, M.E. Staton², S.P. Oliver^{1, 3}, and G.M. Pighetti¹

¹The University of Tennessee, Knoxville, TN, Department of Animal Science

²The University of Tennessee, Knoxville, TN, Department of Entomology and Plant Pathology

³The University of Tennessee, Knoxville, TN, AgResearch

Abstract

Mastitis is a detrimental disease in the dairy industry that costs upwards of \$2 billion annually and decreases milk quality. Often, mastitis results from bacteria entering the gland through the teat end. A common mastitis causing pathogen is *S. uberis*, which is responsible for a high percentage of subclinical and clinical mastitis cases. Following an intramammary experimental challenge with *S. uberis* on Holstein cows (N=40), differing phenotypes with respect to *S. uberis* counts, inflammation indicators, and antibiotic therapies were observed depending on the CXCR1 haplotype combination of the cow. However, phenotypic differences are suspected to have other genetic influences. To test this both linkage and association tests were performed to identify other candidate genes. Linkage analysis was performed in Haploview and identified that the CXCR1 gene is within 2 haplotype blocks, blocks 413 and 414. Block 413 contains only the 3' end of CXCR1 (rs43323012 and 43323013). Block 414, which contains the 5' end of the CXCR1 haplotype combinations (rs211042414 and rs208795699) also contains another 21 genes that the CXCR1 SNPs could be serving as markers for. To identify loci of interest a 50K SNP chip analysis was performed using the BovineSNP50 v2 DNA BeadChip from Illumina. Associations were tested using Plink and identified 7 significant ($p < 9.34 \times 10^{-5}$) SNPs outside of the CXCR1 haplotype blocks. One of the identified SNPs is within a gene with known function and three more are within a previously identified QTL for productive life. The identified loci should be further investigated to potentially identify causation behind the observed phenotypes. Such investigations could lead to novel treatment and prevention compounds/protocols for *S. uberis* mastitis or genetic selection methods for cows with greater potential to resist *S. uberis* infection.

Introduction

Mastitis is a devastating disease to the dairy industry that affects 20-50% of cows, causing increases in cow morbidity/mortality, decreased milk quality, and decreased farm revenues (Akers and Nickerson 2011; Barbano et al. 1991; Bradley et

al. 2007; Klei et al. 1998; Ma et al. 2000; Piepers et al. 2007). Invasion of bacteria into the mammary gland through the teat end is typically what causes mastitis. One of the most common mastitis causing pathogens is *Streptococcus uberis*. *S. uberis* causes roughly 14-26% of all the subclinical and clinical incidences of mastitis and is the most commonly isolated pathogen during the dry (non-lactating) period (Bradley et al. 2007; Petrovski et al. 2011; Verbeke et al. 2014). *S. uberis* is especially problematic because, while it typically originates as an environmental pathogen, it can spread throughout the herd in a contagious manner once introduced (Zadoks 2007). Additionally, *S. uberis* has some unique pathology characteristics that can make clearing the infection from the mammary gland more problematic. Once *S. uberis* enters the mammary gland, it binds lactoferrin, which is a common milk protein (Fang and Oliver 1999). Lactoferrin can then act as a bridging molecule and allow binding of *S. uberis* to the mammary epithelial cells (Fang et al. 2000). After binding, *S. uberis* then invades the host cells via caveolae dependent endocytosis (Almeida and Oliver 2006; Matthews et al. 1994). Once internalized, *S. uberis* can exploit host cell transduction and persist in the cytoplasm for hours (Almeida et al. 2000; Tamilselvam et al. 2006) effectively evading the host immune system. Since *S. uberis* mastitis is a common problem dairy producers combat and *S. uberis* has mechanisms to evade the immune response, identifying specific genes involved in *S. uberis* pathogenesis could prove beneficial for the control and prevention of *S. uberis* mastitis on dairy farms.

Once in the gland, bacteria encounter several cell types including alveolar and ductal epithelial cells that are responsible for synthesizing and secreting milk, fibroblasts, and resident immune cells (i.e.: macrophages, dendritic cells, neutrophils). The encountered host cell types release immune mediators in response to recognition of bacterial presence. Immune mediators can include, but are not limited to: growth related oncogene (GRO)- γ , epithelial-derived neutrophil-activating peptide (ENA)-78, and interleukin (IL)-8 (Baggiolini and Clark-Lewis 1992; Rainard et al. 2008; Walz et al. 1991). The released immune mediators initiate the visual signs associated with mastitis, such as udder inflammation and altered milk production (i.e.: reduced volume, clots, etc.). Specifically, IL-8, upon binding its receptors CXCR1 and CXCR2, can induce migration of immune cells to the site of infection, cause increased cell survival, and modify chemokine/cytokine production (Grob et al. 1990; Kettritz et al. 1998; Shamaladevi et al. 2009; Takata et al. 2004). These functions influence the rate of response to infection and are critical to resolution of infections like mastitis. Furthermore, associations between single nucleotide polymorphisms (SNPs) in the IL-8 receptor, CXCR1 gene and mastitis have been identified (Galvão et al. 2011; Grosse et al. 1999; Leyva-Baca et al. 2008; Verbeke et al. 2012; Youngerman et al. 2004). Some of the identified SNPs have also been associated with diminished migration, binding of

secondary mediators, and receptor internalization following IL-8 binding (Fan et al. 2002; Fan et al. 2000; Raman et al. 2010; Rambeaud et al. 2006; Rambeaud and Pighetti 2007; Rambeaud and Pighetti 2005; Yang et al. 1999). Within the coding sequence of the CXCR1 gene, there are 5 SNPs (+365 (rs211042414), +621 (rs135751317), +735 (rs208795699), +980 (rs43323012), and +995 (rs43323013)) that allowed for the identification of three dominant amino acid haplotypes (VWHKH, VWHRR, AWQRR) (Pighetti et al. 2012).

A recent *S. uberis* challenge study demonstrated that Holstein cows (N=40) displayed differing phenotypes with respect to *S. uberis* counts, inflammation indicators, and antibiotic therapies depending on the CXCR1 haplotype combination of the cow (Siebert 2013). However, we suspect that the observed differences in phenotype following challenge could have other genetic influences. Other genetic influences could include loci in linkage disequilibrium with the previously identified CXCR1 SNPs or loci throughout the genome that work together with CXCR1 pleiotropically. To test for genetic influences, we will identify the haplotype block surrounding the bovine CXCR1 gene and perform an association analysis using a phenotype which includes genetic CXCR1 haplotype combination with antibiotic therapy grouping following intramammary *S. uberis* experimental challenge. This approach will identify genes that are in linkage disequilibrium with CXCR1 and genes that could be working pleiotropically with CXCR1 throughout the genome that could be influencing the different phenotypes observed following *S. uberis* challenge in our prior study (Siebert 2013). Both approaches will identify other candidate genes that can be investigated for potential novel preventative or treatment compounds or improved genetic selection methods for cows less susceptible to *S. uberis* mastitis.

Materials and Methods

Animal Selection

Holstein dairy cows (N = 40) that were entering their second or third lactation were enrolled in a trial to test the effectiveness of a novel vaccine candidate against *S. uberis*, i.e.: *Streptococcus uberis* adhesion molecule (SUAM) that involved a direct *S. uberis* challenge (USDA-NIFA AFRI 2011-67015-30168). Cows were in overall good health, pregnant, at the end of their first or second lactation, had quarter SCC <250,000 cells/mL, and were free of Johne's, brucellosis, tuberculosis, and bovine leukemia virus (BLV). Both *S. uberis* and SUAM baseline titers were assessed and cows with the lowest titers were enrolled. All cows were housed at the East Tennessee Education and Research Center (ETREC) Little River Dairy (Walland, TN). Upon arrival at

ETREC, cows were acclimated to the facility for two weeks prior to the initiation of research protocols. Cows were housed separately from the core herd and cared for in accordance with ETREC Little River Dairy standard protocols. All animal based work was approved by and done in accordance with Institutional Care and Use Committee (IACUC) guidelines.

Haplotyping

Blood samples (n=1 for each cow), collected during the dry period via jugular puncture, were used to isolate DNA with UltraClean BloodSpin DNA Isolation Kit (Mo Bio, Carlsbad, CA). The SNP containing portion of the CXCR1 gene was amplified as follows: 1-2 ng of bovine genomic DNA was used as template in a 25 µl reaction containing specific primers (Forward: 5'-CTTCATC-TTCCGGTGAGGCCTATCAAC-3'; Reverse: 5'-AAGGGAAGGGGACTTTCCTGGCTG-3') and Eppendorf HotMaster Mix (Eppendorf North America; Westbury, NY) according to manufacturer's guidelines. DNA concentration was confirmed via spectrophotometry using NanoDrop 1000 (Thermo Fisher Scientific; Wilmington, DE). Conditions for amplification were: an initial hot-start denaturation at 94°C for 2 min, followed by 50 cycles of 94°C denaturation for 15 s, 62°C annealing for 20 s, and 72°C extension for 60 s. After the last cycle, a 10 min final extension step at 72°C was added before reactions were chilled to 4°C. Amplified products were purified to remove primer and excess nucleotides. The resulting purified CXCR1 DNA was sequenced using bovine CXCR1 specific primers (Forward: 5'-CTTCATC-TTCCGGTGAGGCCTATCAAC-3'; Reverse: 5'-AAGGGAAGGGGACTTTCCTGGCTG-3') and the CXCR1 haplotype, comprised of positions +365 (rs211042414), +621 (rs135751317), +735 (rs208795699), +980 (rs43323012), and +995 (rs43323013), was determined based on nucleotide expression at tagSNP positions +621 (rs135751317), +735 (rs208795699), +816 (rs210501501), and +980 (rs43323012) in the CXCR1 gene previously identified using phasing on a larger population (Pighetti et al. 2012; Stephens and Scheet 2005; Stephens et al. 2001). Within the haplotype combination groups: VWHRR x VWHRR cows did not share any parentage back three generations; VWHKH x VWHRR, AWQRR x VWHRR, and AWQRR x VWHKH groups each had two cows sharing a grandsire, which was the same grandsire for each haplotype combination; and AWQRR x AWQRR cows were the most related group, with two pairs of cows sharing sires, one pair of cows sharing a maternal grandsire, and one set of three cows sharing a paternal grandsire.

Challenge

Approximately one week prior to calving and at first milking post-calving, secretion samples were collected from each quarter to screen for bacterial presence to

select a challenge quarter free of pathogens. A frozen stock of *Streptococcus uberis* (UT888) was thawed in a 37°C water bath and an aliquot used to inoculate a blood agar plate. After 23-24 h of growth on the agar, 3 representative colonies were used to inoculate Todd Hewitt Broth for an additional 7 h of growth. The 1 mL aliquot of the resulting liquid culture was serially diluted in PBS to approximately 2,000 CFU for infusion. A total of 5 mL of the diluted culture was used for challenge, resulting in a total of 10,000 CFU delivered. Challenge was performed within 3 days post-calving immediately following milking. Prior to intramammary infusion, teat ends were cleaned thoroughly with sterile alcohol pads and the inoculum was infused using sterile disposable syringes with teat cannulas. The challenge inoculum was massaged upward into the teat and gland cisterns and plated (100µL) to determine the actual CFU/mL of *S. uberis* delivered which averaged 2122 ± 371 CFU/mL and ranged from 1185 to 3202 CFU/mL.

Sampling and Phenotype Description

Cows were closely monitored post-challenge for the development of mastitis for 28 d post-challenge. Mammary and milk scores were evaluated on a 0 to 3 scale by ETREC Little River Staff twice daily. Mammary scores were defined as follows: 0 = normal, 1 = slight swelling, 2 = moderate swelling, 3 = edema. Milk scores were defined as follows: 0 = normal, 1 = small flakes or clots, 2 = large clots, 3 = stringy, off color, watery, or bloody. Cows with 3 consecutive sets of severe scores (a 2 in both mammary and milk or a 3 in either) were treated with antibiotics. Alternatively, if cows were still shedding *S. uberis* 14 d post-challenge, antibiotics were administered to clear the infection, since the end goal was to return healthy cows back to the general milking herd. Following the 28 d sampling period, the percentage of cows within each CXCR1 haplotype combination that required antibiotics to clear the infection was determined. This resulted in 3 percentage groups used to test for associations: 100% antibiotic therapy (VWHRR x VWHRR and AWQRR x VWHRR), 75% antibiotic therapy (AWQRR x VWHKH and AWQRR x AWQRR), and 33% antibiotic therapy (VWHKH x VWHRR). The haplotypes of two cows, 222 and 237, could not be determined and were not included in response variable analyses, resulting in N=38 cows.

Association and Linkage Analysis

Blood samples were collected from each cow via jugular puncture and DNA was isolated using UltraClean BloodSpin DNA Isolation Kit (Mo Bio, Carlsbad, CA). Resulting DNA was used to perform a 50K SNP chip analysis using the BovineSNP50 DNA Analysis BeadChip (Illumina, San Diego, CA) (Gunderson et al. 2005; Steemers et al. 2006). The SNP chip has 54,609 SNPs that span all 29 autosomes of the bovine

genome with an average SNP spacing of 49.4kb and a median spacing of 36.9kb. Associations were tested between developed the phenotype and SNPs using Plink (Purcell et al. 2007). Effects of sex were excluded from the analysis because all individuals were female and SNPs did not represent sex chromosomes. The p-value cutoffs of $p < 1.95 \times 10^{-5}$ for highly significant SNPs and $p = 1.95 \times 10^{-5}$ to 9.34×10^{-5} for suggestively significant SNPs were set based on guidelines for reducing false positives and increasing reproducibility (Manly 2005). For identified SNPs to be reported, the minor allele had to be represented by more than one cow, unless the SNP aligned with prior evidence for mastitis or mastitis related traits. A complete index of identified SNPs, including those removed based on the minor allele criteria, is included as an appendix (Table A-1). To allow depiction of the numerous SNPs, their genome locations, and their variable association p-values, generated Plink results were visualized using Manhattan plots created with Integrated Genome Viewer (IGV) (Robinson et al. 2011).

For linkage analysis a subset of the 50K SNP chip data was created to include only SNPs on BTA 2, the chromosome where the CXCR1 gene is located. Genotypes for each of the 5 CXCR1 SNPs were added to the SNP chip dataset based on the haplotyping results for each cow. Additionally the locations of the CXCR1 SNPs were added to the SNP map file. Linkage analyses were performed using Haploview (Barrett et al. 2005). SNPs were excluded from the linkage analysis when the minor allele was not observed in the genotyped cows, which resulted in the removal of the CXCR1 +621 (rs135751317). Additionally, SNPs where less than 50% of cows were genotyped due to poor reads were removed from analysis. After SNP editing, a total of 692 SNPs were removed, resulting in 2,142 SNPs on BTA being included in the linkage analysis. Haplotype blocks surrounding the CXCR1 SNPs were determined using the solid spine algorithm such that the first and last marker in each block is in strong linkage disequilibrium with all the intermediate markers. All reported gene and SNP locations for both analyses were reported based on the *Bos taurus* UMD_3.1.1 (GCF_000003055.6) genome build.

Results

The linkage analysis identified a total of 529 haplotype blocks on BTA 2, which is 137.24 Mb in length, using solid spine linkage analysis. The CXCR1 SNPs were included within 2 blocks, block 413 and block 414 (Table 3, Figure 15, Figure 16). Block 413 ranges from rs43323008 to rs43323012 which is 106,924,044 – 106,937,600 bp (13kb) on BTA 2. The only gene coded within block 413 is the 3' end of CXCR1 which includes SNPs rs43323012 and 43323013 (Table 4). Block 414 ranges from rs208795699 to rs110289171 which is 106,937,845 – 107,378,666 bp (440kb) on BTA 2

Table 3: SNPs identified to be in linkage disequilibrium with CXCR1 SNPs using Haploview based on BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) positions.

SNP Name	rs	BTA 2 Location
Block 413 (13 kb)		
ARS-BFGL-NGS-3684	rs43323008	106,924,044
CXCR1c995	rs43323013	106,937,585
CXCR1c980	rs43323012	106,937,600
Block 414 (440 kb)		
CXCR1c735	rs208795699	106,937,845
CXCR1c621	rs135751317	106,937,962
CXCR1c365	rs211042414	106,938,215
BTA-48880-no-rs	rs41585646	106,952,479
BTA-48881-no-rs	rs41585647	106,984,473
Hapmap60413-ss46526389	rs41255527	107,005,982
Hapmap31750-BTA-134395	rs109464198	107,168,923
ARS-BFGL-NGS-116634	rs109230217	107,318,080
ARS-BFGL-NGS-5566	rs110289171	107,378,666

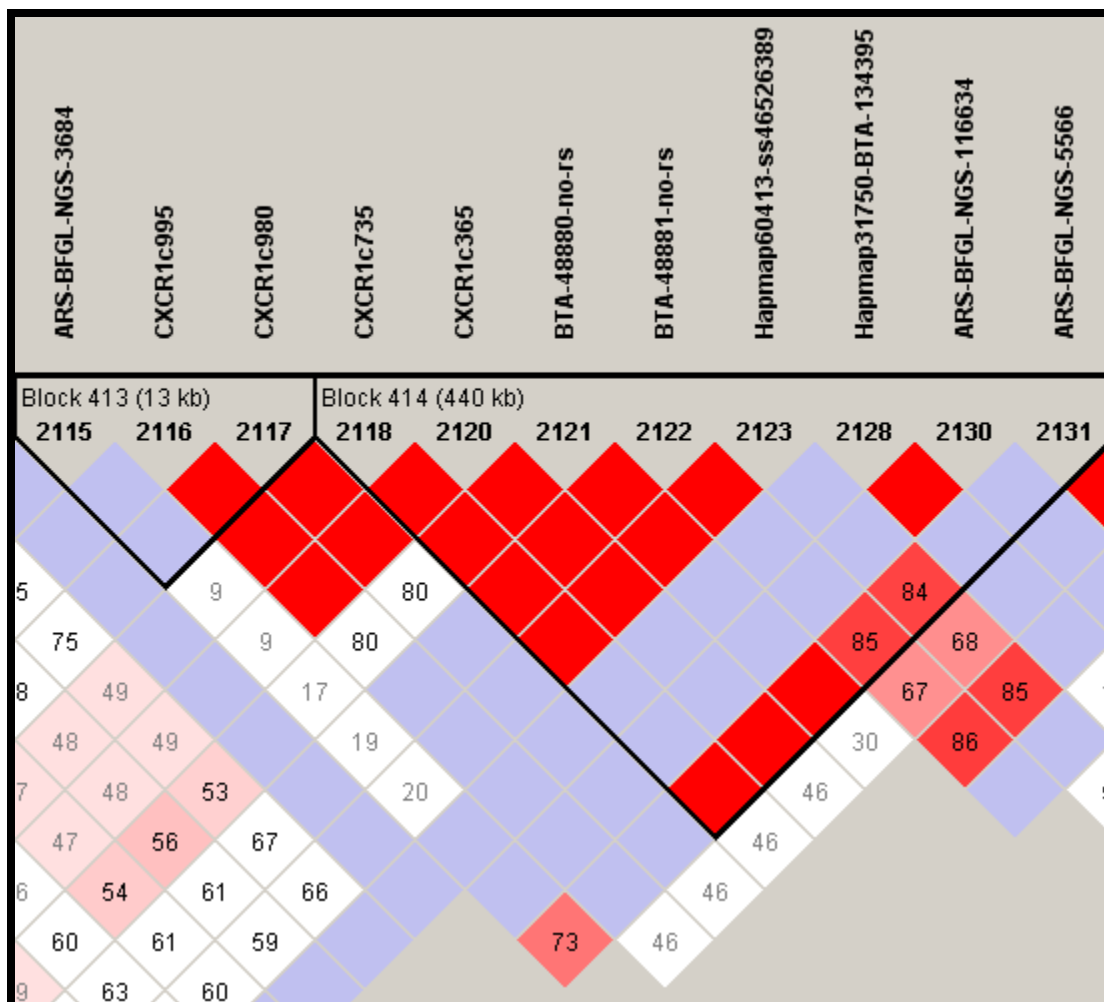


Figure 15: Linkage analysis identifying SNPs on BTA 2 in linkage disequilibrium with CXCR1 SNPs using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA). Blocks are outlined in thick black lines. Colors indicate likelihood of linkage disequilibrium (LD) as calculated by logarithm of odds (LOD; $\log_{10}(\text{likelihood of LD}/\text{likelihood of equilibrium})$ where higher values indicate more likely LD) and D' (disequilibrium coefficient normalized for the observed allele frequencies; values closer to 1 or 100% indicate higher LD; numbers in squares). Red squares = $\text{LOD} \geq 2$, $D' = 1$. Shades of pink = $\text{LOD} \geq 2$, $D' < 1$. Blue = $\text{LOD} < 2$, $D' = 1$. White = $\text{LOD} < 2$, $D' < 1$; strong evidence of recombination.

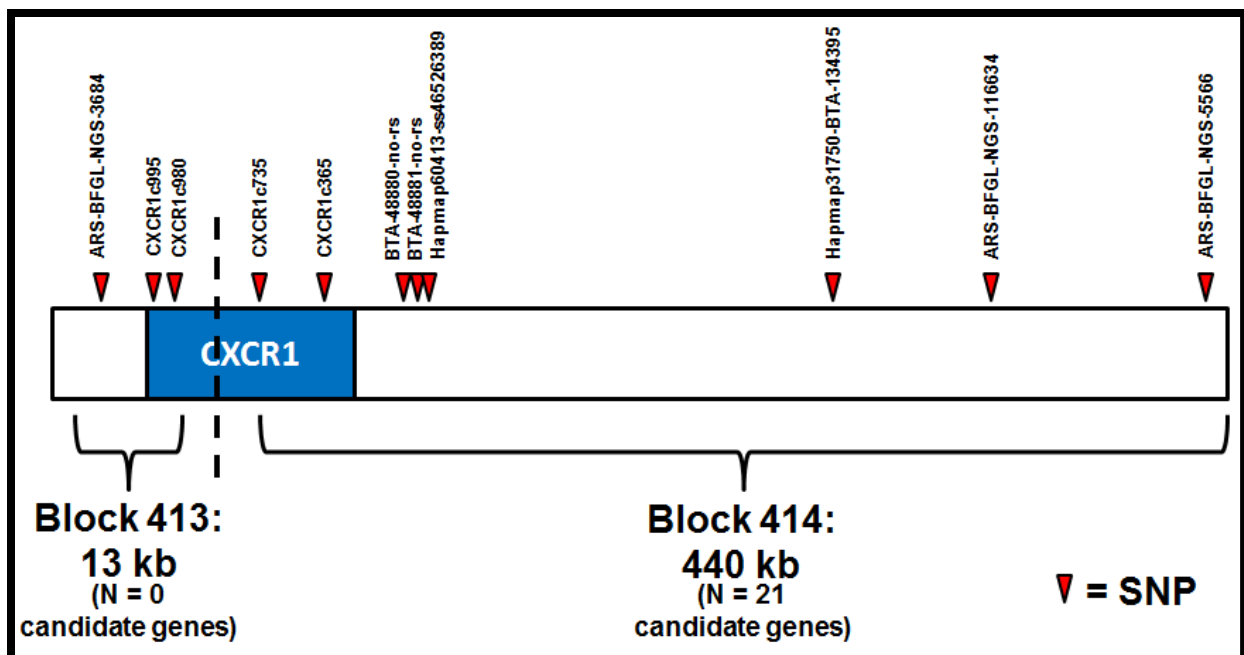


Figure 16: Spatial diagram SNPs identified to be in linkage disequilibrium with CXCR1 SNPs using Haploview based on BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) positions.

Table 4: Genes and RNAs contained within the haplotype blocks surrounding CXCR1 SNPs based on the *Bos taurus* UMD_3.1.1 (GCF_000003055.6) genome build.

Bp Range	Gene Abbreviation	Gene Name	NCBI ID
Block 413 (106,924,044-106,937,600 bp)			
106,936,878-106,937,600	CXCR1	C-X-C motif receptor 1	281863
Block 414 (106,937,845-107,387,666 bp)			
106,937,845-106,938,583	CXCR1	C-X-C motif receptor 1	281863
106,943,692-106,949,081	LOC107131616	non-coding RNA	107131616
106,978,526-107,006,143	ARPC2	actin related protein complex, subunit 2	540838
107,011,567-107,015,000	GPBAR1	g-protein coupled bile acid receptor 1	317756
107,015,357-107,020,300	AAMP	angio-associated migratory cell protein	767919
107,020,657-107,093,550	PNKD	paroxysmal non-kinesigenic dyskinesia	616561
107,023,823-107,041,372	TMBIM1	transmembrane BAX inhibitor motif containing 1	404134
107,082,988-107,087,346	LOC104971306	non-coding RNA	104971306
107,099,749-107,106,176	CATIP	ciliogenesis associated TTC17 interacting protein	514380
107,103,742-107,105,782	LOC104971307	non-coding RNA	104971307
107,114,929-107,125,854	SLC11A1	solute carrier family 11 (proton-coupled divalent metal ion transporter)	282470
107,130,426-107,139,000	CTDSP1	CTD (carboxy-terminal domain) small phosphatase 1	516199
107,133,394-107,133,478	MIR26B	microRNA 26b	MI0004745
107,145,511-107,169,204	VIL1	villin 1	281573
107,170,031-107,255,579	USP37	ubiquitin specific peptidase 37	407168
107,256,013-107,289,222	RQCD1	RCD1 required for cell differentiation 1 homolog (<i>S. pombe</i>)	536537
107,294,531-107,321,327	PLCD4	phospholipase C, delta 4	540771
107,322,027-107,341,857	ZNF142	zinc finger protein 142	511573
107,341,872-107,345,831	BCS1L	BCS1-like (<i>S. cerevisiae</i>)	539713
107,346,275-107,355,558	RNF25	ring finger protein 25	511571
107,355,281-107,387,666	STK36	serine/threonine kinase 36	508408

and includes the 5' end of the CXCR1 gene (rs211042414 and rs208795699). There are a total of 17 protein coding genes, 3 non-coding RNAs, and 1 micro RNA coded within block 414 (Table 4).

For the antibiotic therapy phenotype developed from both CXCR1 haplotype combination and the necessity of antibiotics following challenge (Table 5, Figure 17), the 70-75% antibiotic therapy phenotype was most common, representing N=18 cows with both AWQRR x VWHKH and AWQRR x AWQRR haplotype combinations. The second most common phenotype was the 100% antibiotic therapy, with N=10 cows from both VWHRR x VWHRR and AWQRR x VWHRR haplotype combinations. The 33.3% antibiotic therapy phenotype was the least common, representing only VWHKH x VWHRR cows with N=6 cows. Association analyses identified a total of 7 SNPs associated with the antibiotic therapy phenotype (Table 6 and Figure 18). Three of the identified SNPs were highly significant ($p < 1.95 \times 10^{-5}$) and 4 were suggestively significant ($p = 1.95 \times 10^{-5}$ to 9.34×10^{-5}). The identified SNPs spanned 4 chromosomes, with 1 SNP on each BTA 7 (rs41612889) and BTA 8 (rs109792349), 3 SNPs on BTA 16 (rs41660125, rs41576666, rs41596554) and 2 SNPs on BTA 2 (rs43317360, rs41644229). One SNP on BTA 2, rs43317360, is located within the CRYBA2 (crystallin beta A2) gene.

Discussion

The goal of this study was to identify genes in linkage disequilibrium with SNPs in CXCR1 and could be serving as markers for and to identify genes that could be working pleiotropically with CXCR1 located throughout the genome. These loci of interest could help explain the different phenotypes observed with respect to inflammation indicators, *S. uberis* CFU, and the need for antibiotics to clear the infection from our prior study by providing candidate genes for further investigation (Siebert 2013). Both approaches to identify genes were successful and a total of 21 candidate genes were identified and are discussed below.

Candidate Loci in LD with CXCR1:

The linkage analysis identified 2 separate blocks encompassing the previously identified CXCR1 SNPs on BTA 2. Block 413 did not include any candidate genes that were in linkage disequilibrium (LD) with the CXCR1 SNPs. However, block 414 included 22 candidate genes that are in LD with two CXCR1 SNPs. Since these genes are in LD with the CXCR1 3' SNPs rs211042414 and rs208795699 previously associated with phenotypic differences following *S. uberis* challenge, the CXCR1 SNPs could be markers for the functions of these candidate genes. Several of these

Table 5: Frequency of haplotype combinations observed in our sample population and their corresponding antibiotic therapy groupings.

Haplotype Combination	Frequency	Antibiotic Therapy Phenotype
VWHRR x VWHRR	5	100%
VWHKH x VWHRR	7	33.3%
AWQRR x VWHRR	6	100%
AWQRR x VWHKH	9	70-75%
AWQRR x AWQRR	11	70-75%
Total	38	

Table 6: Identified loci associated with the CXCR1 antibiotic therapy phenotype.

rs	Locus (chr: bp) ¹	SNP type	Allele frequencies ²	Nearest gene(s)	Prior evidence	P-value
Loci with prior evidence						
rs41660125	16: 10389306	intergenic	A: 0.45 G: 0.55	LOC100848559, CDC73	PL QTL (Zhang et al., 1998)	1.30x10 ⁻⁵
rs41576666	16: 9941492	intergenic	C: 0.61 T: 0.39	LOC100848559, CDC73	PL QTL (Zhang et al., 1998)	3.70x10 ⁻⁵
rs41596554	16: 10740538	intergenic	G: 0.50 T: 0.50	LOC100848559, CDC73	PL QTL (Zhang et al., 1998)	9.36x10 ⁻⁵
Newly identified loci						
rs43317360	2: 107657675	intron variant	A: 0.70 C: 0.30	CRYBA2	--	2.20x10 ⁻⁶
rs41644229	2: 107681700	5' upstream	C: 0.65 T: 0.35	LOC101906201	--	1.54x10 ⁻⁵
rs41612889	7: 107527276	5' upstream	C: 0.45 T: 0.55	EFNA5	--	2.75x10 ⁻⁵
rs109792349	8: 8618532	intergenic	A: 0.61 T: 0.39	LOC783359, PRSS55	--	3.74x10 ⁻⁵

¹Reported positions are based on the *Bos taurus* UMD_3.1.1 (GCF_000003055.6) genome build.

²Allele frequencies are representative of the current study population.

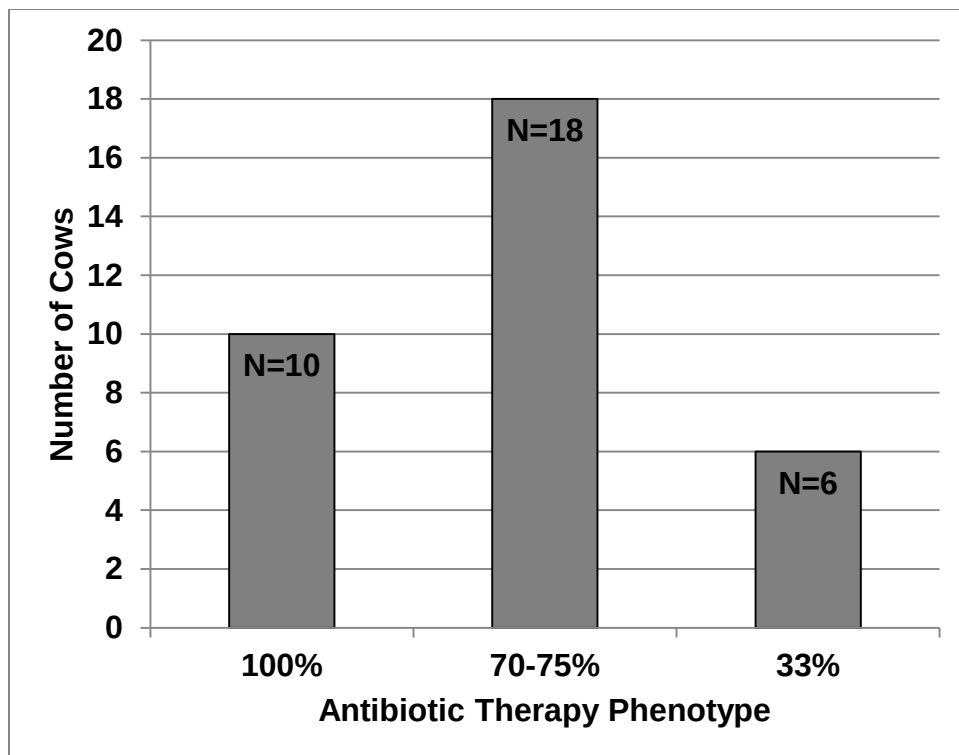


Figure 17: Number of cows per antibiotic therapy phenotype grouping.

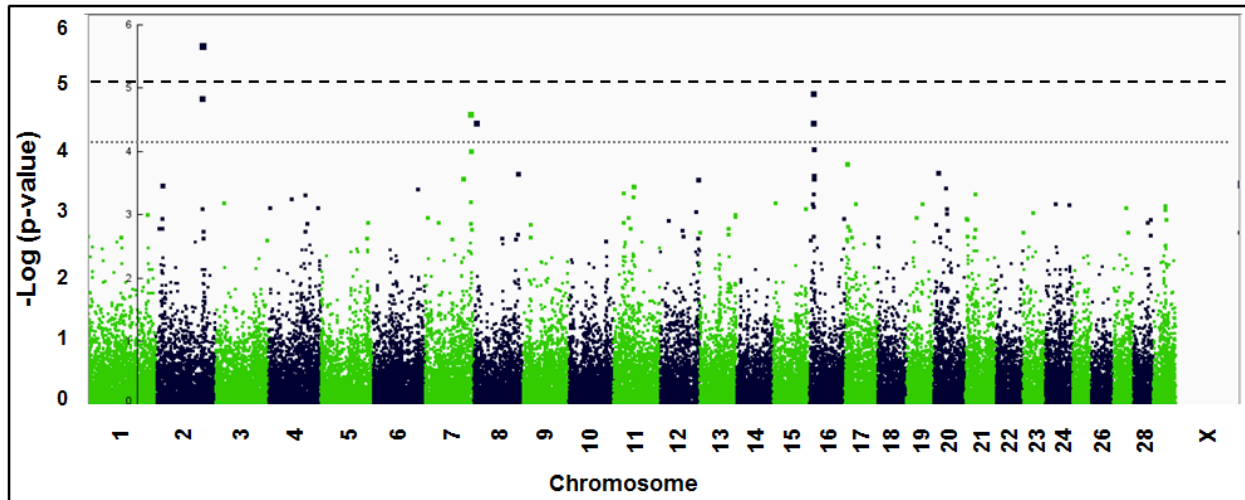


Figure 18: Manhattan plot of GWAS for antibiotic therapy phenotype using BovineSNP50 v2 BeadChip (Illumina, San Diego, CA) in Holsteins (N = 35). The x-axis shows chromosome and SNP order within a chromosome. The y-axis show the $-\log_{10}(P\text{-value})$ for each SNP. P-value thresholds are indicated by dashed lines: top black line $p = 1.95 \times 10^{-5}$; bottom gray line $p = 9.34 \times 10^{-5}$.

candidate genes have functions that could be linked to *S. uberis* pathogenesis or immune processes.

The ARPC2 (actin related protein 2/3 complex, subunit 2) gene is a subunit of the ARP complex. The ARP complex is implicated in the control of actin polymerization within cells, which is critical for a variety of cellular processes including cell motility and cell-to-cell adhesion (DeMali et al. 2002; Goley and Welch 2006; Pollard and Borisy 2003; Weiner et al. 1999). Alterations in a cow's cell's ability to migrate and adhere to other cells could alter its ability to respond to an infection state since the recruitment of neutrophils to the site of infection is heavily dependent on both actions and could explain the differences in phenotype observed following our *S. uberis* challenge.

Another gene in LD with CXCR1 3' SNPs rs211042414 and rs208795699 is GPBAR1 (g-protein coupled bile acid receptor 1) which is the cell surface receptor for bile acids. Links between GPBAR1 and macrophage function have been established by applying agonists where such agonists have resulted in reduced monocyte adhesion to endothelial cells and decreased production of pro-inflammatory cytokines shifting monocyte conversion to primarily M2 macrophage-like anti-inflammatory cells (Högenauer et al. 2014; Kida et al. 2013). Alterations in monocyte and macrophage function in an infection state could alter a cow's ability to clear an infection and influence the intensity of the inflammatory response initiated in response to bacteria presence as observed following our *S. uberis* challenge.

AAMP (angio-associated migratory cell protein) is another gene in LD with CXCR1 3' SNPs rs211042414 and rs208795699 that is part of the immunoglobulin superfamily (Beckner et al. 1995). AAMP plays a role in the formation of endothelial tube formation and is associated with angiogenesis, but can also activate NFκB (Beckner et al. 2002; Beckner and Liotta 1996; Bielig et al. 2009). NFκB is a transcriptional activator of numerous genes including many that are critical for inflammation and immunity (Bonizzi and Karin 2004). Decreased transcriptional activation of genes critical for inflammation and immunity could result in a diminished response to bacteria presence and allow for *S. uberis* continued growth as observed in the varying responses to our challenge.

TMBIM1 (transmembrane BAX inhibitor motif containing 1) is also in LD with CXCR1 3' SNPs rs211042414 and rs208795699 and is associated with extrinsic apoptosis regulation and negative regulation of MMP9 (Xu and Reed 1998; Zhao et al. 2006). MMP9 is required for macrophage migration necessary for invading pathogen elimination (Gong et al. 2008). Reductions in extrinsic apoptosis and macrophage migration could result in a diminished immune response to invading pathogens leading

to lower inflammation observed and allowing continued pathogen growth as observed in the responses to our *S. uberis* challenge (Siebert 2013).

Another gene in LD with CXCR1 3' SNPs rs211042414 and rs208795699 is SLC11A1 (solute carrier family 11 – proton-coupled divalent metal ion transporter). SLC11A1 is a member of the proton-coupled divalent metal ion transporters that transports iron, zinc, and manganese cations that is localized to late lysosomes and endosomes (Goswami et al. 2001; Gruenheid et al. 1997; Searle et al. 1998). Additionally, SLC11A1 has been implicated in disease resistance through cation control in macrophages which protects macrophages from their own ROS generation and sequesters cations that pathogens could use to produce protective enzymes as reviewed by Blackwell et al. (2001). Thus, SLC11A1 could play a role in a cow's ability to properly clear *S. uberis* following challenge through proper macrophage function.

The VIL1 (villin 1) gene is also in LD with CXCR1 3' SNPs rs211042414 and rs208795699 and is a calcium regulated actin binding protein within the cytoskeleton (Bretscher and Weber 1980). Expression of VIL1 has been observed in the leading edge of lamellipodia implying a regulating role in actin remodeling needed for cell migration and it has been demonstrated to regulate apoptosis by maintaining mitochondrial integrity (Nusrat et al. 1992; Wang et al. 2008). Thus, alterations in VIL1 function could influence the immune response to invading pathogens, specifically the ability of immune cells to migrate to the site of infection and survive long enough to eliminate the pathogen which could explain the differences we observed post *S. uberis* challenge.

PLCD4 (phospholipase C delta 4), another gene in LD with CXCR1 3' SNPs rs211042414 and rs208795699, is a member of the delta class of phospholipase C enzymes which hydrolyzes PIP2 (phosphatidylinositol 4,5-bisphosphate) to make DAG (diacylglycerol) and IP3 (inositol 1,4,5-triphosphate) (Ochocka and Pawelczyk 2003). The production of these secondary messengers is the first step of the arachidonic acid pathway that is responsible for the production of leukotrienes, thromboxanes, prostaglandins, and lipoxins which have an array of roles in inflammation, where alterations in their production could result in different responses to invading pathogens.

The last gene in LD with CXCR1 3' SNPs rs211042414 and rs208795699 is RNF25 (ring finger protein 25) which interacts with the p65 subunit of NFκB to support transcription (Asamitsu et al. 2003). As discussed with AAMP above, decreased transcriptional activation of inflammation and immunity genes could result in diminished response to invading bacteria allowing growth to continue as observed in response to our *S. uberis* challenge. Thus, the identification of the haplotype block surrounding the

bovine CXCR1 gene has provided several potential candidate genes that the CXCR1 3' SNPs rs211042414 and rs208795699 could be serving as markers for and could be influencing the phenotypic differences observed following our *S. uberis* challenge.

Furthermore, there are an additional 13 genes in this block that are either uncharacterized or do not have a documented link to inflammation, immunity, or transcriptional regulation that should be further investigated to determine if they could play a role in *S. uberis* pathogenesis.

Candidate Loci Associated with Antibiotic Therapy Groupings:

To determine if there were other genes throughout the bovine genome that could be working pleiotropically with our CXCR1 SNPs or the loci in LD with the CXCR1 SNPs to cause the phenotypic differences observed following our *S. uberis* challenge, we performed an association analysis. The phenotype used grouped cows based on both CXCR1 haplotype combination and that haplotype combination's need for antibiotics to clear *S. uberis* from the gland. This association analysis identified a total of 7 SNPs. None of these identified SNPs are within the CXCR1 haplotype blocks discussed above, but one SNP (rs43317360) was identified on BTA 2 just 0.28 Mb outside of the CXCR1 haplotype block 414. While the identified SNPs themselves may not be causal but rather in linkage disequilibrium is the causal variant, they do provide several interesting candidate genes based on the genes they are located within. This SNP (rs43317360) is an intron variant in the CRYBA2 (crystallin beta A2) gene. Crystallins are the major structural component of the eye ocular lens but are also implicated in retinal ganglion cell axonal regeneration following injury (Piri et al. 2013). Additionally, another crystallin, CRYBB2, has been implicated in promoting neurite activity, demonstrating roles in tissues besides eyes (Liedtke et al. 2007). The possibility exists that CRYBA2 could play a role in tissue regeneration following injury in other tissues such as the mammary gland and contribute to resolving the infection state following *S. uberis* challenge. This could allow cows to clear the infection more readily leading to the demonstration of reduced mammary and milk scores. Since highly inflammatory mammary and milk scores were used as the basis for antibiotic therapy administration, and CRYBA2 could help explain the variation observed in antibiotic therapy groupings based on CXCR1 haplotype combination following our *S. uberis* challenge.

Also of interest, the 3 associated SNPs on BTA 16 (rs41660125, rs41576666, rs41596554) are within a productive life QTL (Zhang et al. 1998). Productive life refers to the length of time a cow remains in the herd producing milk, which can be affected by many factors including health, fertility, and milk yield (Roxström and Strandberg 2002). Cows with phenotypes indicating they can mount an effective immune response

to invading pathogens and clear the infection without the aid of antibiotic therapy (i.e.: 33% antibiotic therapy phenotype; VWHKH x VWHRR cows) would be more likely to be retained in the herd, barring other issues, because they would be less labor intensive and cost producers less money in treatment than cows with the other phenotypes.

In conclusion, this study successfully identified an array of candidate genes that could help explain the phenotypic differences observed following *S. uberis* challenge either because they are in LD with the CXCR1 SNPs or because they are associated with the antibiotic therapy groupings developed with the cow's CXCR1 genetic backgrounds. Many of these candidate genes have roles in immunity, inflammation, or the regulation of gene expression. Further investigation of these candidate genes should determine their expression levels in the mammary gland and could help identify potential preventative or novel treatment compounds for *S. uberis* mastitis. Additionally, these results could discover better defined or more accurate genetic markers for selecting cows with less susceptibility to *S. uberis* mastitis.

References

- Asamitsu, K., Tetsuka, T., Kanazawa, S., and Okamoto, T. (2003). "Ring finger protein AO7 supports NF- κ B-mediated transcription by interacting with the transactivation domain of the p65 subunit." *Journal of Biological Chemistry*, 278(29), 26879-26887.
- Beckner, M. E., Jagannathan, S., and Peterson, V. A. (2002). "Extracellular angio-associated migratory cell protein plays a positive role in angiogenesis and is regulated by astrocytes in coculture." *Microvascular Research*, 63(3), 259-269.
- Beckner, M. E., Kruttsch, H. C., Stracke, M. L., Williams, S. T., Gallardo, J. A., and Liotta, L. A. (1995). "Identification of a new immunoglobulin superfamily protein expressed in blood vessels with a heparin-binding consensus sequence." *Cancer Research*, 55(10), 2140-2149.
- Beckner, M. E., and Liotta, L. A. (1996). "AAMP, a conserved protein with immunoglobulin and WD40 domains, regulates endothelial tube formation in vitro." *Laboratory Investigation*, 75(1), 97-107.
- Bielig, H., Zurek, B., Kutsch, A., Menning, M., Philpott, D., Sansonetti, P., and Kufer, T. (2009). "A function for AAMP in Nod2-mediated NF- κ B activation." *Molecular immunology*, 46(13), 2647-2654.
- Blackwell, J. M., Goswami, T., Evans, C. A. W., Sibthorpe, D., Papo, N., White, J. K., Searle, S., Miller, E. N., Peacock, C. S., and Mohammed, H. (2001). "SLC11A1 (formerly NRAMP1) and disease resistance." *Cellular Microbiology*, 3(12), 773-784.
- Bonizzi, G., and Karin, M. (2004). "The two NF- κ B activation pathways and their role in innate and adaptive immunity." *Trends in Immunology*, 25(6), 280-8.
- Bretscher, A., and Weber, K. (1980). "Villin is a major protein of the microvillus cytoskeleton which binds both G and F actin in a calcium-dependent manner." *Cell*, 20(3), 839-847.
- DeMali, K. A., Barlow, C. A., and Burridge, K. (2002). "Recruitment of the Arp2/3 complex to vinculin coupling membrane protrusion to matrix adhesion." *The Journal of Cell Biology*, 159(5), 881-891.
- Goley, E. D., and Welch, M. D. (2006). "The ARP2/3 complex: an actin nucleator comes of age." *Nature Reviews Molecular Cell Biology*, 7(10), 713-726.
- Gong, Y., Hart, E., Shchurin, A., and Hoover-Plow, J. (2008). "Inflammatory macrophage migration requires MMP-9 activation by plasminogen in mice." *The Journal of Clinical Investigation*, 118(9), 3012-3024.

- Goswami, T., Bhattacharjee, A., Babal, P., Searle, S., Moore, E., Ming, L., and Jenefer, M. (2001). "Natural-resistance-associated macrophage protein 1 is an H⁺/bivalent cation antiporter." *Biochemical Journal*, 354(3), 511-519.
- Gruenheid, S., Pinner, E., Desjardins, M., and Gros, P. (1997). "Natural resistance to infection with intracellular pathogens: the Nramp1 protein is recruited to the membrane of the phagosome." *Journal of Experimental Medicine*, 185(4), 717-730.
- Högenauer, K., Arista, L., Schmiedeberg, N., Werner, G., Jaksche, H., Bouhelal, R., Nguyen, D. G., Bhat, B. G., Raad, L., Rauld, C., and Carballido, J. M. (2014). "G-Protein-coupled bile acid receptor 1 (GPBAR1, TGR5) agonists reduce the production of proinflammatory cytokines and stabilize the alternative macrophage phenotype." *Journal of Medicinal Chemistry*, 57(24), 10343-10354.
- Kida, T., Tsubosaka, Y., Hori, M., Ozaki, H., and Murata, T. (2013). "Bile acid receptor TGR5 agonism induces NO production and reduces monocyte adhesion in vascular endothelial cells." *Arteriosclerosis, Thrombosis, and Vascular Biology*, 33(7), 1663-1669.
- Liedtke, T., Schwamborn, J. C., Schröer, U., and Thanos, S. (2007). "Elongation of axons during regeneration involves retinal crystallin β b2 (crybb2)." *Molecular & Cellular Proteomics*, 6(5), 895-907.
- Nusrat, A., Delp, C., and Madara, J. L. (1992). "Intestinal epithelial restitution. Characterization of a cell culture model and mapping of cytoskeletal elements in migrating cells." *Journal of Clinical Investigation*, 89(5), 1501-1511.
- Ochocka, A.-M., and Pawelczyk, T. (2003). "Isozymes delta of phosphoinositide-specific phospholipase C and their role in signal transduction in the cell." *Acta Biochimica Polonica-English Edition*, 50(4), 1097-1110.
- Piri, N., Kwong, J. M. K., and Caprioli, J. (2013). "Crystallins in retinal ganglion cell survival and regeneration." *Molecular Neurobiology*, 48(3), 819-828.
- Pollard, T. D., and Borisy, G. G. (2003). "Cellular motility driven by assembly and disassembly of actin filaments." *Cell*, 112(4), 453-465.
- Roxström, A., and Strandberg, E. (2002). "Genetic analysis of functional, fertility-, mastitis-, and production-determined length of productive life in Swedish dairy cattle." *Livestock Production Science*, 74(2), 125-135.
- Searle, S., Bright, N. A., Roach, T. I., Atkinson, P. G., Barton, C. H., Meloen, R. H., and Blackwell, J. M. (1998). "Localisation of Nramp1 in macrophages: modulation with activation and infection." *Journal of Cell Science*, 111(19), 2855-2866.

- Siebert, L. J. (2013). "Bovine mammary interleukin-8 receptor expression and genetic association with *Streptococcus uberis* based mastitis."
- Wang, Y., Srinivasan, K., Siddiqui, M. R., George, S. P., Tomar, A., and Khurana, S. (2008). "A novel role for villin in intestinal epithelial cell survival and homeostasis." *Journal of Biological Chemistry*, 283(14), 9454-9464.
- Weiner, O. D., Servant, G., Welch, M. D., Mitchison, T. J., Sedat, J. W., and Bourne, H. R. (1999). "Spatial control of actin polymerization during neutrophil chemotaxis." *Nature Cell Biology*, 1(2), 75-81.
- Xu, Q., and Reed, J. C. (1998). "Bax inhibitor-1, a mammalian apoptosis suppressor identified by functional screening in yeast." *Molecular Cell*, 1(3), 337-346.
- Zhang, Q., Boichard, D., Hoeschele, I., Ernst, C., Eggen, A., Murkve, B., Pfister-Genskow, M., Witte, L. A., Grignola, F. E., Uimari, P., Thaller, G., and Bishop, M. D. (1998). "Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree." *Genetics*, 149(4), 1959-1973.
- Zhao, H., Ito, A., Sakai, N., Matsuzawa, Y., Yamashita, S., and Nojima, H. (2006). "RECS1 is a negative regulator of matrix metalloproteinase-9 production and aged RECS1 knockout mice are prone to aortic dilation." *Circulation Journal : Official Journal Of The Japanese Circulation Society*, 70(5), 615-624.

Chapter VI: Determining growth potential of *Streptococcus uberis* in aseptic milk samples obtained from cows with differing CXCR1 genetic backgrounds

Determining growth potential of *Streptococcus uberis* in aseptic milk samples obtained from cows with differing CXCR1 genetic backgrounds

L. Siebert¹, S.P. Oliver^{1, 2}, and G.M. Pighetti¹

¹The University of Tennessee, Knoxville, TN, Department of Animal Science

²The University of Tennessee, Knoxville, TN, AgResearch

Abstract

Mastitis, inflammation of the mammary gland, is a devastating disease for the dairy industry causing billions in lost revenues annually and decreasing milk quality. Due to prevalence and impacts of mastitis, improved prevention and treatment protocols are a consistent goal of the industry. *S. uberis* is a common mastitis pathogen whose growth in milk is influenced by stage of lactation. The IL-8 receptor, CXCR1, has demonstrated importance for mastitis and SNPs within CXCR1 have been associated with mastitis and neutrophil migration. A prior *S. uberis* intramammary challenge study demonstrated Holsteins with different CXCR1 haplotype combinations display varying phenotypes with respect to inflammation indicators and *S. uberis* counts suggesting CXCR1 haplotype combination could influence the ability or rate of *S. uberis* to grow in milk and the recruitment of leukocytes to the mammary gland. To test this, we collected aseptic milk samples from cows with differing CXCR1 genetic backgrounds within 7 days post-calving and inoculated them in vitro with 3 strains of *S. uberis*. Additionally, leukocyte profiles were evaluated in the aseptic milk samples using the QScout® platform to determine if CXCR1 haplotype combination influenced total leukocyte recruitment or type of leukocytes recruited to the mammary gland. Results indicate that both CXCR1 haplotype combination and *S. uberis* strain influence the rate of *S. uberis* in milk ($p < 0.0001$). VWHRR x VWHRR milk had the highest rate of *S. uberis* growth (slope = 0.58) and AWQRR x AWQRR milk had the lowest rate of *S. uberis* growth (slope = 0.098). The highest rate of *S. uberis* growth was observed from strain UT366 (slope = 0.58) and the lowest rate of growth was observed from strain UT387 (slope = 0.17). CXCR1 haplotype combination did not influence total leukocyte counts or percentages of leukocyte subpopulations observed in milk ($p > 0.05$). These results indicate that both CXCR1 haplotype combination and *S. uberis* strain virulence influence *S. uberis* growth rate in milk which could indicate that concentration of an inhibitory or permissive factor towards *S. uberis* growth in milk differs depending on cow CXCR1 haplotype combination. Future proteomics studies would be beneficial in identifying any potential factors of interest and lead to the development of novel preventative or treatment compounds towards *S. uberis* mastitis.

Introduction

Mastitis, the inflammation of the mammary gland, is typically caused by bacteria invading the mammary gland. Impacts of mastitis include dramatic decreases in milk quality and farm revenues (Fetrow 2000; Klei et al. 1998; Ma et al. 2000).

Streptococcus uberis is a common mastitis pathogen which is commonly isolated from the mammary gland during the non-lactating period and accounts for a high percentage of clinical and subclinical mastitis (Bradley et al. 2007; Petrovski et al. 2011; Verbeke et al. 2014). *S. uberis* originates in the environment of the cow (i.e.: soils), but can be passed from cow-to-cow in a contagious manner (Zadoks 2007), making it a large proportion of the mastitis problem dairy producers combat.

In order for mastitis to occur, invading pathogenic bacteria must be able to survive and proliferate within the mammary gland by overcoming the host defenses. The host defenses in the mammary gland are numerous and milk contains a variety of anti-microbial factors that inhibit pathogen growth (Oliver and Calvinho 1995). Prior work on growth of *S. uberis* in mammary secretions has focused stage of lactation. Periods of intense physiological change demonstrate greater susceptibility to *S. uberis* (Jayarao et al. 1999). Secretions obtained during any stage of mammary gland involution (the period when the mammary gland undergoes shrinkage and tissue remodeling after lactation cessation) or the dry (non-lactating) period support *S. uberis* growth better than secretions obtained during lactation (Oliver 1991; Todhunter et al. 1985). The differences in growth of *S. uberis* observed during periods of intense physiological change were attributed to concentrations of antimicrobial factors (i.e.: lactoperoxidase) and or essential nutrients (i.e.: plasminogen degradation by-products) needed for *S. uberis* growth, where secretions obtained during lactation had more antimicrobial factors or less essential nutrients and supported less *S. uberis* growth (Kliem and Hillerton 2002; Leigh 1993; Todhunter et al. 1985). However, when considering only phases when the mammary gland is actively lactating, early lactation better permits *S. uberis* growth than mid or late lactation (Rambeaud et al. 2004). Early lactation being a highly susceptible period for *S. uberis* mastitis was supported by the observation of a 100% infection rate in N=40 Holstein cows in an intramammary experimental challenge administered within three days post-calving (Siebert 2013). Early lactation supporting *S. uberis* growth more than mid or late lactation contradicts that essential nutrients could be driving differences in *S. uberis* growth. Plasminogen concentrations increase as lactation progresses meaning more available nutrients to promote *S. uberis* growth, which would dictate more *S. uberis* growth in later lactation not less (Benslimane et al. 1990). This suggests there is another mechanism that is capable of influencing *S. uberis* growth in milk during lactation.

Studies have demonstrated that the IL-8 receptor, CXCR1, plays a critical role in the defense of the mammary gland by influencing neutrophil migration, survival, calcium signaling, and reactive oxygen species (ROS) generation (Rambeaud et al. 2006; Rambeaud and Pighetti 2007; Rambeaud and Pighetti 2005). Additionally, polymorphisms in the CXCR1 gene sequence have been linked to mastitis incidence (Galvão et al. 2011; Grosse et al. 1999; Leyva-Baca et al. 2008; Verbeke et al. 2012; Youngerman et al. 2004). A recent *S. uberis* intramammary challenge study demonstrated that Holstein cows (N=40) with different CXCR1 haplotype combinations (Pighetti et al. 2012) displayed varying phenotypes with respect to *S. uberis* counts, inflammation indicators, and antibiotic therapies (Siebert 2013). The differing phenotypes observed following intramammary challenge could indicate that a cow's CXCR1 haplotype combination influences the ability or rate of *S. uberis* to grow in milk and/or the recruitment of leukocytes to the mammary gland. To test this, we collected aseptic whole milk from non-infected cows with different CXCR1 haplotype combinations and inoculated them with three strains of *S. uberis* obtained from mastitic cows to monitor growth rates for 12 hours in a laboratory setting. Additionally, leukocyte profiles were evaluated in the aseptic milk samples using the QScout® platform to determine if CXCR1 haplotype combination influenced total leukocyte recruitment or type of leukocytes recruited to the mammary gland.

Materials and Methods

Haplotyping

Blood samples (N=1 for each cow), collected during the dry period via coccygeal vein puncture, were obtained from all cows expected to calve at Little River ETREC dairy. Blood was used to isolate DNA with UltraClean BloodSpin DNA Isolation Kit (Mo Bio, Carlsbad, CA). The SNP containing portion of the CXCR1 gene was amplified as follows: 1-2 ng of bovine genomic DNA was used as template in a 25 µl reaction containing specific primers (Forward: 5'-CTTCATC-TTCCGGTGAGGCCTATCAAC-3'; Reverse: 5'-AAGGGAAGGGGACTTTCCTGGCTG-3') and Eppendorf HotMaster Mix (Eppendorf North America; Westbury, NY) according to manufacturer's guidelines. DNA concentration was confirmed via spectrophotometry using NanoDrop 1000 (Thermo Fisher Scientific; Wilmington, DE). Conditions for amplification were: an initial hot-start denaturation at 94°C for 2 min, followed by 50 cycles of 94°C denaturation for 15 s, 62°C annealing for 20 s, and 72°C extension for 60 s. After the last cycle, a 10 min final extension step at 72°C was added before reactions were chilled to 4°C. Amplified products were purified to remove primer and excess nucleotides. The resulting purified CXCR1 DNA was sequenced using bovine CXCR1 specific primers (Forward: 5'-

CTTCATC-TTCCGGTGAGGCCTATCAAC-3'; Reverse: 5'-AAGGGAAGGGGACTTTCCTGGCTG-3') and the CXCR1 haplotype, comprised of positions +365 (rs211042414), +621 (rs135751317), +735 (rs208795699), +980 (rs43323012), and +995 (rs43323013), was determined based on nucleotide expression at tagSNP positions +621 (rs135751317), +735 (rs208795699), +816 (rs210501501), and +980 (rs43323012) in the CXCR1 gene previously identified using phasing on a larger population (Pighetti et al. 2012; Stephens and Scheet 2005; Stephens et al. 2001). Not all haplotypes combinations were enrolled in the study because of infection status prior to milk inoculation.

Milk Sampling and Collection

Between two and three days post-calving, aseptic quarter samples were taken from each cow. Samples were kept on ice and transported back to the lab. One mL of each sample was plated on an individual blood agar plate using flame sterilized hockey sticks under a sterile hood. Plates were allowed to completely dry under the hood, then inverted and place in an incubator at 37°C, 5%CO₂ for 48 hrs. Cows with two or more plates showing no bacteria growth after incubation were chosen to proceed to the milk inoculation. Composite, aseptic, 100 mL inoculation samples were collected between 4 and 6 days post-calving. An aliquot of the aseptic milk sample was used to determine the types of leukocytes present in milk using two methods. First, overall somatic cell counts (SCC) were determined using a benchtop DeLaval Cell Counter DCC (DeLaval, Kansas City, Missouri). The DeLaval SCC was used as a final screening measure, where cows with SCC above 200,000 cells/mL (industry standard division point indicating an intramammary infection), were excluded from milk inoculation protocols even though their initial microbiology screening two days prior indicated they did not have an infection. Samples that passed the SCC screening were kept on ice and transported back to the lab. Prior to inoculation, samples were partitioned into 4 20mL aliquots (1 for each of the 3 strains and a negative) and warmed to 37°C for 30 mins. After pre-warming, 1mL of the negative sample was plated on a blood agar plate to determine if any infection occurred between screening and inoculation sample collection or if contamination occurred.

Evaluation of Leukocytes in Milk

The aliquot of the aseptic milk used for inoculation was used to determine the types of leukocytes present in the milk using QScout® MLD platform (Advanced Animal Diagnostics, Morrisville, NC). QScout® MLD utilizes special cartridges, where 10µL of the milk sample is loaded onto the cartridge which is inserted into the machine. The resulting data includes a total leukocyte count (TLC) which is equivalent to a SCC, as

well as counts and percentages for 3 specific leukocyte populations per microliter (neutrophils, lymphocytes, and macrophages).

***S. uberis* Cultures**

Frozen stock cultures of 3 strains of *S. uberis* (UT366, UT387, and UT888) were obtained from the Tennessee Quality Milk Lab. Stocks were used to inoculate 25 mL of Todd Hewitt Broth (THB) and cultures were grown at 37°C, 5%CO₂ for 24 hrs. Resulting cultures were used to make 50µL aliquots that were frozen at -80°C. One fresh aliquot for each strain was used for each inoculation.

Two days prior to milk inoculation, the 50µL aliquots are thawed to room temperature then plated on a blood agar plate, one plate per strain, using streak technique. Streak plates were grown at 37°C, 5%CO₂ for 24 hrs. After incubation, for each strain, a sterile loop was used to select 3 individual colonies and then inoculate a pre-warmed, 15mL aliquot of THB which was incubated overnight at 37°C, 5%CO₂. The next morning, 3 hours prior to milk inoculation, overnight cultures were removed from the incubator, vortexed, and then 500µL were used to inoculate a pre-warmed 5mL THB aliquot. New aliquots were incubated at 37°C, 5%CO₂ for 3 hrs, and then serially diluted to approximately 1,000 CFU/mL. Four hundred µL of the diluted culture for each strain was then used to inoculate the 20mL, pre-warmed milk aliquot for each cow. Additionally, 100µL of the diluted cultures was plated on Todd Hewitt agar (THA) to determine actual CFU inoculated for each strain, which was used as the time zero CFU for each strain's growth curve.

Inoculating Milk with *S. uberis*

The diluted *S. uberis* culture for each strain (400µL) was added to the corresponding pre-warmed 20mL milk sample. Milk cultures were then vortexed and placed in the incubator at 37°C, 5%CO₂ for 12 hrs. At the start of each hour, samples were removed from the incubator, vortexed, and 1mL was plated using a flame sterilized hockey stick to determine how strain growth progressed. After time point, milk cultures were also serially diluted to ensure colonies were sparse enough to count. Inoculated milk samples were plated on THA and the negative sample was plated on blood agar. Plates were allowed to dry completely in the hood, inverted, and then placed in an incubator at 37°C, 5%CO₂. After 48 hrs of incubation, plates were removed and colonies of *S. uberis* were counted and used to determine total CFU present in milk at each time point for each strain.

Statistical Analysis

A mixed model analysis of variance tested the fixed effects of CXCR1 haplotype combination, strain, and time and their interactions. The random effect of cow nested within CXCR1 haplotype combination was also included to account for unknown variation that an individual's background and physiology may have introduced into the model. Time was declared as repeated measures since time point samples were collected from the same inoculated milk sample. Missing time points were due to colonies on plates being too numerous to count or because plate contamination was observed that made counting colonies of *S. uberis* impossible. Time points 7, 11, and 12 were removed due to > 50% missing data points. The same model was also used to determine if slopes of *S. uberis* growth over time differed between CXCR1 haplotype combinations and strains of *S. uberis* by including significant interaction terms as fixed effects directly. Data was $\log_{10}$ transformed in order to maintain normality. Additionally, QScout® MLD data was utilized to determine if leukocyte populations differed dependent on CXCR1 haplotype combination, where the model tested the fixed effect of CXCR1 haplotype combination only, but still included the random effect of cow nested within CXCR1 haplotype combination to account for unknown variation that an individual's background and physiology may have introduced. Analyses were performed in SAS (v 9.4, Cary, NC) where significance was declared at $p < 0.05$, and a trend was declared at $p = 0.05$ to $p = 0.1$.

Results

The dominant haplotype combination of the inoculated cows was AWQRR x VWHRR, which represented 42.9% of cows. The second most frequent haplotype was homozygous AWQRR representing 33.3% of cows. The remaining two haplotype combinations (VWHRR x VWHRR, VWHKH x VWHRR) represented much smaller percentages of the inoculated cows (Table 7). Cows with haplotype combination VWHKH x AWQRR were not observed.

S. uberis growth was significantly influenced by strain ($p < 0.0001$) (Figure 19). Strain 888 grew the most (200.1 ± 178.5) but was statistically similar to strain 366 (70.8 ± 61.4 CFU). Strain 366 was also similar to strain 387 (11.7 ± 10.2 CFU). All strains were statistically unique from the negative (0.07 ± 0.06 CFU). Additionally, the interaction of strain and time was significant ($p < 0.0001$) indicating the rate of growth over time differed across the strains (Figure 20). Strain 366 had the greatest slope, 0.58 CFU/hr. indicating the quickest rate of growth over time. Strain 888 had the second largest slope, 0.31 CFU/hr., followed by strain 387 with a slope of 0.17 CFU/hr. The smallest slope was observed by the negative, and it actually showed a decrease in

Table 7: Frequency of haplotype combinations observed in cows whose milk was inoculated with the 3 strains of *S. uberis*.

Haplotype Combination	Frequency	Percent
VWHRR x VWHRR	2	9.5
VWHKH x VWHRR	3	14.3
AWQRR x VWHRR	9	42.9
AWQRR x AWQRR	7	33.3
Total	21	100

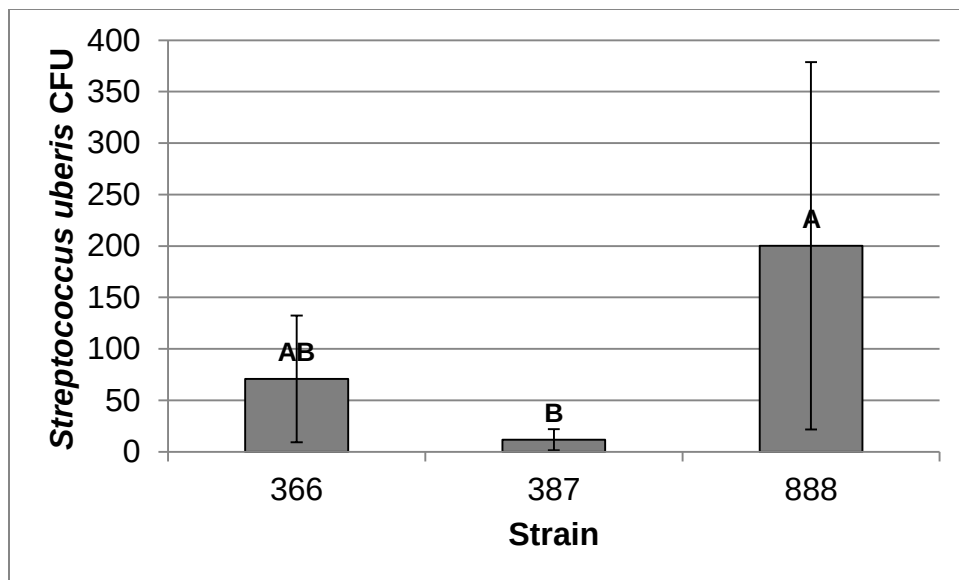


Figure 19: Growth of three strains of *S. uberis*, represented as back transformed least squares means from $\log_{10}$, in aseptic whole milk obtained from Holstein cows within 7 days post-calving. Letters above the bars indicated means that differ ($p < 0.0001$).

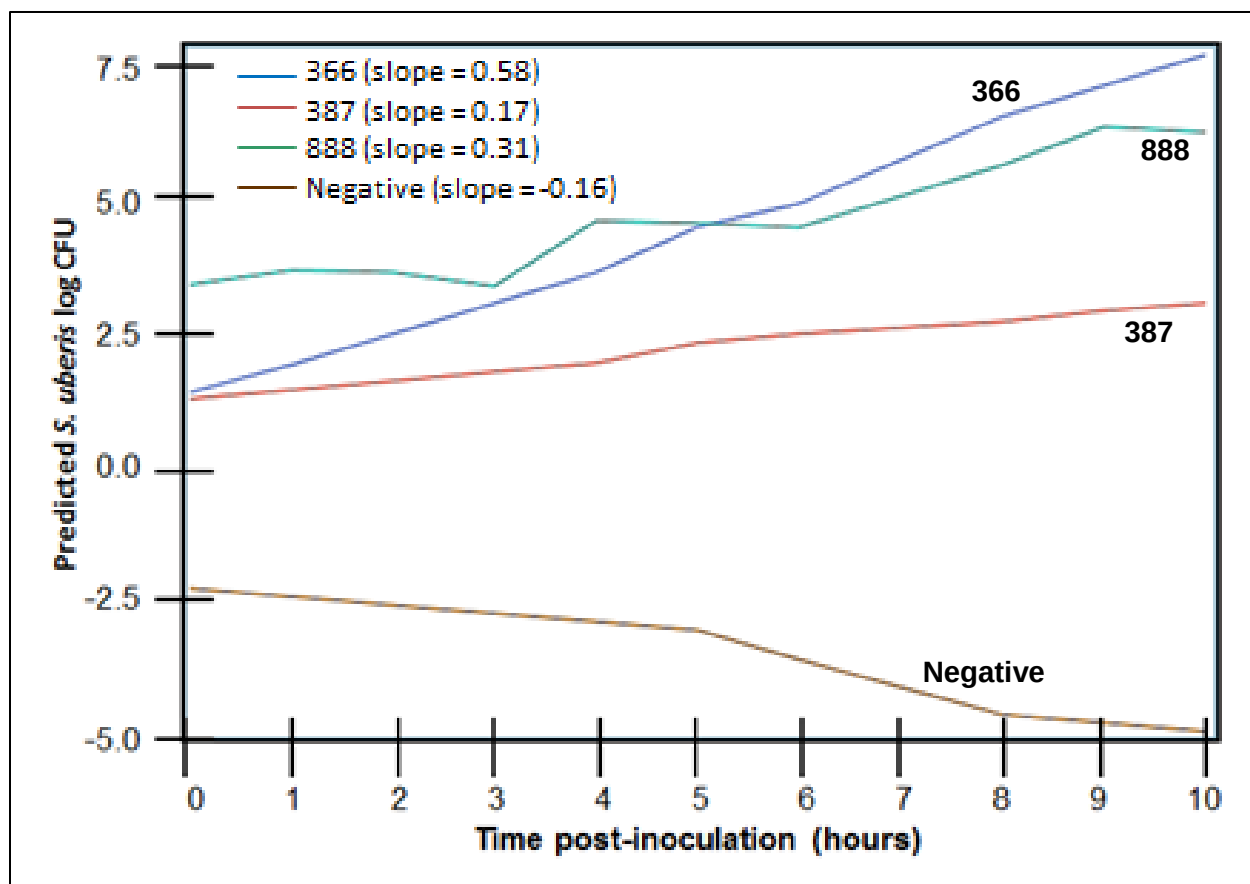


Figure 20: The effect of *S. uberis* strain on growth over time in aseptic whole milk obtained from Holstein cows within 7 days post-calving. All slopes differed from each other significantly over time ($p < 0.0001$).

growth over time with a slope of -0.16.

S. uberis growth was not influenced by haplotype combination ($p = 0.668$), nor was a strain by haplotype combination interaction observed ($p = 7879$). However, the interaction of haplotype combination and time was significant ($p < 0.0001$) indicating the rate of *S. uberis* growth over time differed across haplotype combinations (Figure 21). VWHRR x VWHRR milk samples grew the fastest with a slope of 0.58 CFU/hr. AWQRR x AWQRR milk samples grew the slowest with a slope of only 0.098 CFU/hr. VWHKH x VWHRR and AWQRR x VWHRR milk samples had intermediate slopes at 0.41 and 0.21 CFU/hr. respectively.

Haplotype combination did not influence total leukocyte counts collected using the QScout® MLD platform ($p = 0.6424$) (Figure 22). Haplotype combination also did not influence the percentages of neutrophils ($p = 0.1509$), lymphocytes ($p = 0.7863$), or macrophages ($p = 0.4903$) observed in the aseptic milk collected (Figure 23). However, for all haplotype combinations, neutrophils were the majority of leukocytes present (60-50%) with lymphocytes and macrophages each representing smaller fractions (15-30% each) (Figure 23).

Discussion

The present study determined that CXCR1 haplotype combination influences the rate that *S. uberis* growth in whole milk. These results are consistent with results from our prior *S. uberis* intramammary challenge study that demonstrated cows with differing CXCR1 haplotype combinations showed differing *S. uberis* CFU in the gland post-challenge (Siebert 2013). The VWHRR x VWHRR displayed the highest overall *S. uberis* CFU in the challenge study and had the highest rate of *S. uberis* growth in the present study suggesting *S. uberis* most easily grows in VWHRR x VWHRR milk. The challenge study identified AQWRR x VWHRR as the haplotype combination with the least amount of *S. uberis* growth and in the present study this combination had the second smallest rate of *S. uberis* growth in milk. Combined this suggests that the AWQRR allele better permits *S. uberis* growth than the VWHRR allele. This is further confirmed by the AWQRR x AWQRR milk displaying the overall lowest *S. uberis* rate of growth in the present study. The VWHKH x VWHRR milk displayed an intermediate rate of *S. uberis* growth in the present study and had overall *S. uberis* CFU similar to all other haplotype combination in the challenge study suggesting the VWHKH allele supports growth of *S. uberis* better than AWQRR but not as well as VWHRR.

The differences in *S. uberis* growth based on CXCR1 haplotype combination

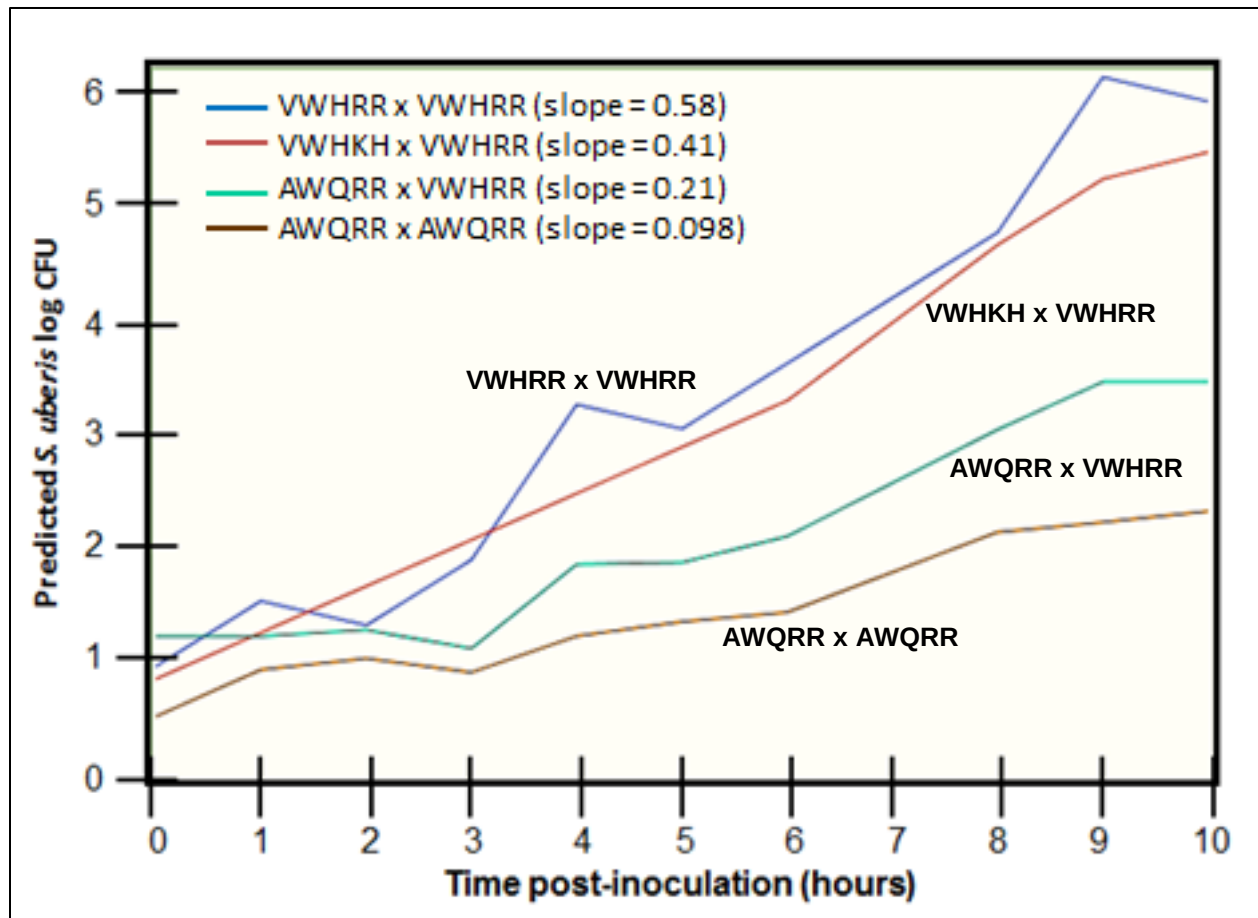


Figure 21: The effect of haplotype combination on *S. uberis* growth over time in aseptic whole milk obtained from Holstein cows within 7 days post-calving. All slopes differed from each other significantly over time ($p < 0.0001$).

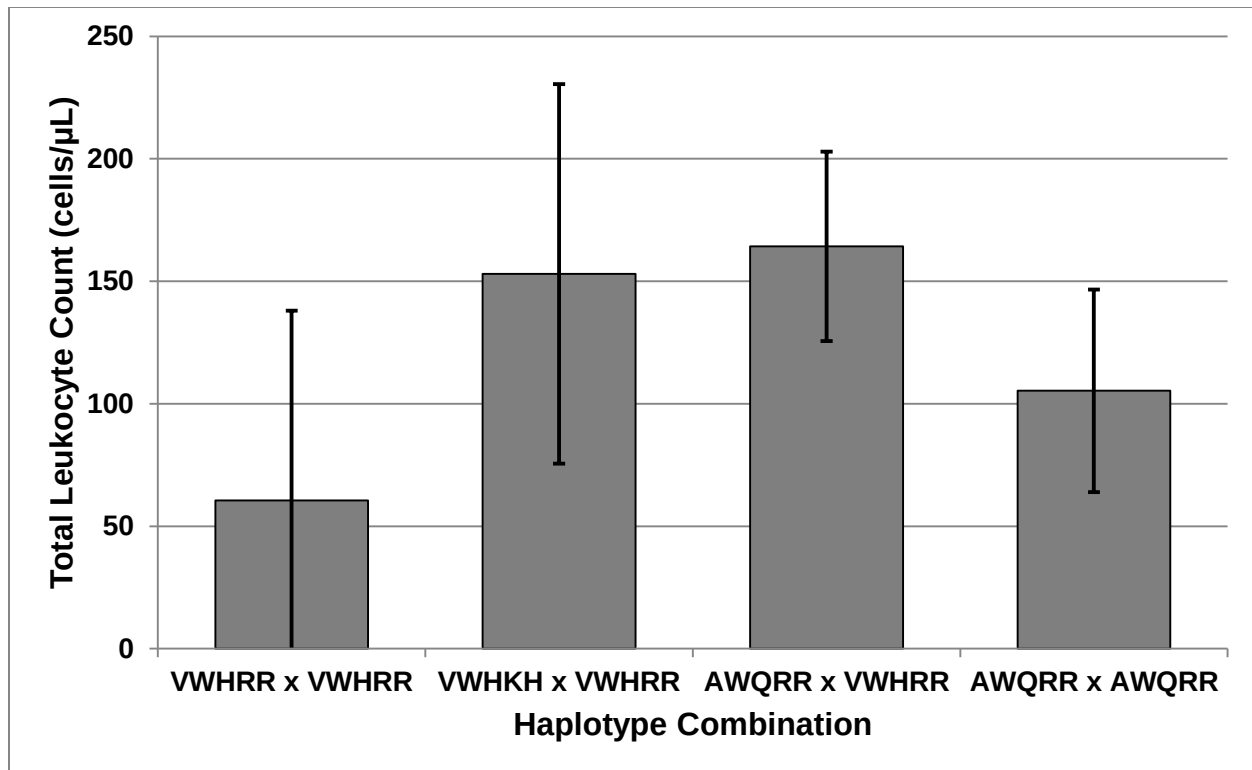


Figure 22: Total leukocyte counts observed for each haplotype combination based on QScout analysis performed on aseptic whole milk obtained from Holstein cows within 7 days post-calving. Haplotype combination did not influence total leukocyte counts $p = 6424$.

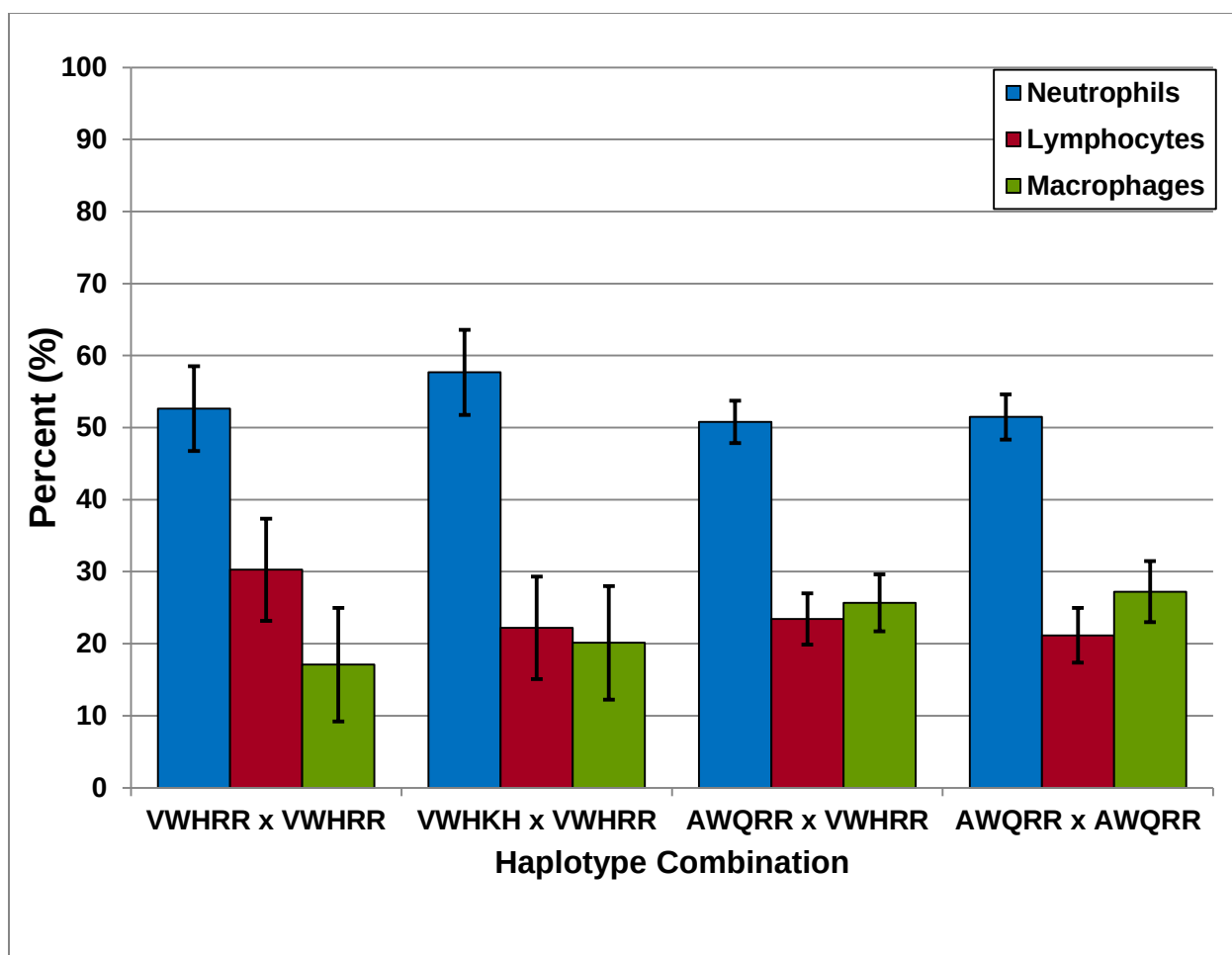


Figure 23: Percentages of leukocyte sub-populations observed for each haplotype combination based on QScout analysis performed on aseptic whole milk obtained from Holstein cows within 7 days post-calving. Haplotype combination did not influence the percentages of neutrophils ($p = 0.1509$), lymphocytes ($p = 0.7863$), or macrophages ($p = 0.4903$) observed.

could be due to differences in neutrophil ability. The middle position of the haplotype (amino acid position 245, SNP 735, rs208795699) is associated with the ability of neutrophils to migrate and survive as well as internal calcium signaling and reactive oxygen species (ROS) generation (Rambeaud et al. 2006; Rambeaud and Pighetti 2007; Rambeaud and Pighetti 2005). Where VWHRR x VWHRR cows would be less capable of recruiting neutrophils to the site of infection and have better killing ability compared to AWQRR x AWQRR cows. However, the QScout® MLD analysis conducted in the present study did not show that the percentages of neutrophils differed across haplotype combinations. Furthermore, *S. uberis* has demonstrated the ability to evade neutrophil killing in vitro, but is effectively killed by macrophages suggesting the percentage of macrophages could be more important for *S. uberis* clearance from the mammary gland (Oliver et al. 1998a). But the present study also demonstrated that the percentage of macrophages present in milk did not differ based on haplotype combination, suggesting other mechanisms may be at work. Of interest, since the present study was conducted in vitro where introduction of *S. uberis* occurred once milk was removed from the cow, differences in *S. uberis* clearance by neutrophils and macrophages would rely on leukocytes present in the milk at the time of collection since the ability to recruit more leukocytes is irrelevant.

Variation in proteomic profiles of milk could influence a cow's ability to clear *S. uberis* from the mammary gland. Proteomics studies of bovine milk and its fractions (i.e.: exosomes, whey, milk fat globule membranes) have identified many components that can influence the immune response to invading pathogens (Ibeagha-Awemu et al. 2010; Reinhardt and Lippolis 2006; Reinhardt and Lippolis 2008; Reinhardt et al. 2012; Reinhardt et al. 2013; Smolenski et al. 2014). Milk fat globule membranes and exosome milk fractions contain immune components that are essential for the detection of infection and for proper immune response to infection (Reinhardt and Lippolis 2006; Reinhardt et al. 2012). Additionally, following *Staphylococcus aureus* infection, there is an increase in the components of neutrophil NETs seen in milk fat globule membranes (Reinhardt et al. 2013). Neutrophil NETs function to trap and bacteria to prevent further pathogenesis (Brinkmann et al. 2004). Thus, alterations in concentration of any of these immune components could influence the growth of *S. uberis* in milk. If levels of these components essential for fighting off invading pathogens differ based on CXCR1 haplotype combination, they could explain the differences in *S. uberis* growth rate observed in the present study. Of note, while these results may prove relevant for the results we observed, they were reported based on individual milk fractions whereas our study was conducted using whole milk. Thus, the interactions between all the different fractions and their components could have an effect on the ability of *S. uberis* to grow. Also of interest, a prior study suggested that plasminogen concentration could influence

growth rate of *S. uberis* because its degradation is essential for *S. uberis* to get the essential nutrients it needs for growth (Kliem and Hillerton 2002). If plasminogen, or other *S. uberis* nutrient sources, vary with CXCR1 haplotype combination, haplotype combinations with higher nutrient source concentrations would better permit *S. uberis* growth.

The present study also observed differences in the ability and rate of *S. uberis* growth based on *S. uberis* strain. These findings were consistent with prior findings by Rambeaud et al. (2004) and are likely due to differences in virulence between *S. uberis* strains. Both UT366 and UT888 demonstrated the two highest overall *S. uberis* growth in the present study and are known to cause clinical mastitis. Strain UT366 is associated with the rapid onset of acute clinical systemic symptoms (Doane et al. 1987) which is supported by its steep rate of *S. uberis* growth in the present study. The UT888 strain is associated with milder clinical mastitis and a lag in *S. uberis* growth while it adapts to the mammary gland environment (Oliver et al. 2003; Rambeaud et al. 2003) which is supported by its rate of *S. uberis* growth in the present study being lower than UT366 but higher than UT387. Furthermore, both UT366 and UT888 have demonstrated the ability to internalize into host cells via binding lactoferrin, a common milk protein, and exploiting caveolae mediated endocytosis (Almeida et al. 2010). However, UT888 has enhanced ability to bind lactoferrin and has higher transcytosis when compared to UT366 (Almeida et al. 2010), suggesting UT888 has a better ability to internalize and hide from the immune system than UT366. Better ability to internalize and hide from the immune system would facilitate lower growth observed in milk as observed in the present study and the ability of UT888 infections to sometimes persist for more than one lactation (Oliver et al. 1998b). UT387 displays the lowest rate of *S. uberis* growth in the present study likely because it expresses a capsule making growth and proliferation more time intensive than the other two strains studied which do not express capsules.

In conclusion, this study provides evidence that *S. uberis* growth in milk during early lactation is influenced by both the cow CXCR1 haplotype combination and *S. uberis* strain virulence. This could indicate that the concentration of an inhibitory or permissive factor towards *S. uberis* growth in milk differs depending on cow CXCR1 haplotype combination. Conducting a proteomics survey of milk from cows with differing CXCR1 haplotype combinations could identify any potential compounds of interest and lead to the development of novel preventative or treatment compounds towards *S. uberis* mastitis.

Acknowledgements

The authors would like to thank the Little River Dairy staff for all their assistance in collecting blood and milk samples for this project. Additionally, the authors would like to thank Drs. Stephen P. Oliver and Raul A. Almeida of the Tennessee Milk Quality Lab for the donation of the strains used in this project.

References

- Almeida, R. A., Dunlap, J. R., and Oliver, S. P. (2010). "Binding of host factors influences internalization and intracellular trafficking of *Streptococcus uberis* in bovine mammary epithelial cells." *Veterinary Medicine International*, 319192(10), 3.
- Benslimane, S., Dognin-Bergeret, M. J., Berdague, J.-L., and Gaudemer, Y. (1990). "Variation with season and lactation of plasmin and plasminogen concentrations in Montbeliard cows' milk." *Journal of Dairy Research*, 57(04), 423-435.
- Bradley, A. J., Leach, K. A., Breen, J. E., Green, L. E., and Green, M. J. (2007). "Survey of the incidence and aetiology of mastitis on dairy farms in England and Wales." *The Veterinary Record*, 160, 253-258.
- Brinkmann, V., Reichard, U., Goosmann, C., Fauler, B., Uhlemann, Y., Weiss, D. S., Weinrauch, Y., and Zychlinsky, A. (2004). "Neutrophil extracellular traps kill bacteria." *Science*, 303(5663), 1532-5.
- Doane, R. M., Oliver, S. P., Walker, R. D., and Shull, E. P. (1987). "Experimental infection of lactating bovine mammary glands with *Streptococcus uberis* in quarters colonized by *Corynebacterium bovis*." *American Journal of Veterinary Research*, 48(5), 749-754.
- Fetrow, J. "Mastitis: An economic consideration." *Presented at Council, N.M. (ed.), 39th Annual Meeting of the National Mastitis Council*.
- Galvão, K. N., Pighetti, G. M., Cheong, S. H., Nydam, D. V., and Gilbert, R. O. (2011). "Association between interleukin-8 receptor- α (CXCR1) polymorphism and disease incidence, production, reproduction, and survival in Holstein cows." *Journal of Dairy Science*, 94(4), 2083-2091.
- Grosse, W. M., Kappes, S. M., Laegreid, W. W., Keele, J. W., Chitko-McKown, C. G., and Heaton, M. P. (1999). "Single nucleotide polymorphism (SNP) discovery and linkage mapping of bovine cytokine genes." *Mammalian Genome*, 10(11), 1062-1069.
- Ibeagha-Awemu, E. M., Ibeagha, A. E., Messier, S., and Zhao, X. (2010). "Proteomics, genomics, and pathway analyses of *Escherichia coli* and *Staphylococcus aureus* infected milk whey reveal molecular pathways and networks involved in mastitis." *Journal of Proteome Research*, 9(9), 4604-4619.
- Jayarao, B. M., Gillespie, B. E., Lewis, M. J., Dowlen, H. H., and Oliver, S. P. (1999). "Epidemiology of *Streptococcus uberis* intramammary infections in a dairy herd." *Journal of Veterinary Medicine, Series B*, 46(7), 433-442.

- Klei, L., Yun, J., Sapru, A., Lynch, J., Barbano, D., Sears, P., and Galton, D. (1998). "Effects of milk somatic cell count on cottage cheese yield and quality." *Journal of Dairy Science*, 81(5), 1205-1213.
- Kliem, K. E., and Hillerton, J. E. (2002). "Possible labile inhibition of the growth of *Streptococcus uberis* in milk from cows free from mastitis." *Journal of Dairy Research*, 69(03), 375-382.
- Leigh, J. A. (1993). "Activation of bovine plasminogen by *Streptococcus uberis*." *FEMS Microbiology Letters*, 114(1), 67-71.
- Leyva-Baca, I., Schenkel, F., Martin, J., and Karrow, N. A. (2008). "Polymorphisms in the 5' upstream region of the CXCR1 chemokine receptor gene and their association with somatic cell score in Holstein cattle in Canada." *Journal of Dairy Science*, 91(1), 407-417.
- Ma, Y., Ryan, C., Barbano, D. M., Galton, D. M., Rudan, M. A., and Boor, K. J. (2000). "Effects of somatic cell count on quality and shelf-life of pasteurized fluid milk." *Journal of Dairy Science*, 83(2), 264-274.
- Oliver, S. P. (1991). "Growth of *Staphylococcus aureus* and *Streptococcus* species in bovine mammary secretions during the nonlactating and peripartum periods following intramammary infusion of lipopolysaccharide at cessation of milking." *Zoonoses and Public Health*, 38(1-10), 538-544.
- Oliver, S. P., Almeida, R. A., and Calvinho, L. F. (1998a). "Virulence factors of *Streptococcus uberis* isolated from cows with mastitis." *Zentralbl Veterinarmed B*, 45(8), 461-71.
- Oliver, S. P., Almeida, R. A., Gillespie, B. E., Ivey, S. J., Moorehead, H., Lunn, P., Dowlen, H. H., Johnson, D. L., and Lamar, K. C. (2003). "Efficacy of extended pirlimycin therapy for treatment of experimentally induced *Streptococcus uberis* intramammary infections in lactating dairy cattle." *Veterinary Therapeutics*, 4(3), 299-308.
- Oliver, S. P., and Calvinho, L. F. (1995). "Influence of inflammation on mammary gland metabolism and milk composition." *Journal of Animal Science*, 73(suppl_2), 18-33.
- Oliver, S. P., Gillespie, B. E., and Jayarao, B. M. (1998b). "Detection of new and persistent *Streptococcus uberis* and *Streptococcus dysgalactiae* intramammary infections by polymerase chain reaction-based DNA fingerprinting." *FEMS Microbiology Letters*, 160(1), 69-73.
- Petrovski, K. R., Williamson, N. B., Lopez-Villalobos, N., Parkinson, T. J., and Tucker, I. G. (2011). "Culture results from milk samples submitted to veterinary diagnostic

- laboratories from August 2003 to December 2006 in New Zealand." *New Zealand Veterinary Journal*, 59(6), 317-322.
- Pighetti, G. M., Kojima, C. J., Wojakiewicz, L., and Rambeaud, M. (2012). "The bovine CXCR1 gene is highly polymorphic." *Veterinary Immunology and Immunopathology*, 145(1–2), 464-470.
- Rambeaud, M., Almeida, R. A., and Oliver, S. P. (2004). "Growth of *Streptococcus uberis* in skim milk obtained from holstein and jersey dairy cows during different stages of lactation." *Journal of Veterinary Medicine, Series B*, 51(3), 143-145.
- Rambeaud, M., Almeida, R. A., Pighetti, G. M., and Oliver, S. P. (2003). "Dynamics of leukocytes and cytokines during experimentally induced *Streptococcus uberis* mastitis." *Veterinary Immunology and Immunopathology*, 96(3–4), 193-205.
- Rambeaud, M., Clift, R., and Pighetti, G. M. (2006). "Association of a bovine CXCR2 gene polymorphism with neutrophil survival and killing ability." *Veterinary Immunology and Immunopathology*, 111(3–4), 231-238.
- Rambeaud, M., and Pighetti, G. (2007). "Differential calcium signaling in dairy cows with specific CXCR1 genotypes potentially related to interleukin-8 receptor functionality." *Immunogenetics*, 59(1), 53-58.
- Rambeaud, M., and Pighetti, G. M. (2005). "Impaired neutrophil migration associated with specific bovine CXCR2 genotypes." *Infection and Immunity*, 73(8), 4955-4959.
- Reinhardt, T. A., and Lippolis, J. D. (2006). "Bovine milk fat globule membrane proteome." *Journal of Dairy Research*, 73(04), 406-416.
- Reinhardt, T. A., and Lippolis, J. D. (2008). "Developmental changes in the milk fat globule membrane proteome during the transition from colostrum to milk." *Journal of Dairy Science*, 91(6), 2307-2318.
- Reinhardt, T. A., Lippolis, J. D., Nonnecke, B. J., and Sacco, R. E. (2012). "Bovine milk exosome proteome." *Journal of Proteomics*, 75(5), 1486-1492.
- Reinhardt, T. A., Sacco, R. E., Nonnecke, B. J., and Lippolis, J. D. (2013). "Bovine milk proteome: quantitative changes in normal milk exosomes, milk fat globule membranes and whey proteomes resulting from *Staphylococcus aureus* mastitis." *Journal of Proteomics*, 82, 141-154.
- Siebert, L. J. (2013). "Bovine mammary interleukin-8 receptor expression and genetic association with *Streptococcus uberis* based mastitis."

- Smolenski, G. A., Broadhurst, M. K., Stelwagen, K., Haigh, B. J., and Wheeler, T. T. (2014). "Host defence related responses in bovine milk during an experimentally induced *Streptococcus uberis* infection." *Proteome Science*, 12(1), 19.
- Stephens, M., and Scheet, P. (2005). "Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation." *The American Journal of Human Genetics*, 76(3), 449-462.
- Stephens, M., Smith, N. J., and Donnelly, P. (2001). "A new statistical method for haplotype reconstruction from population data." *The American Journal of Human Genetics*, 68(4), 978-989.
- Todhunter, D. A., Larry Smith, K., and Schoenberger, P. S. (1985). "In vitro growth of mastitis associated *Streptococci* in bovine mammary secretions 1, 2." *Journal of Dairy Science*, 68(9), 2337-2346.
- Verbeke, J., Piepers, S., Peelman, L., Van Poucke, M., and De Vlieghe, S. (2012). "Pathogen-group specific association between CXCR1 polymorphisms and subclinical mastitis in dairy heifers." *Journal of Dairy Research*, 79(03), 341-351.
- Verbeke, J., Piepers, S., Supré, K., and De Vlieghe, S. (2014). "Pathogen-specific incidence rate of clinical mastitis in Flemish dairy herds, severity, and association with herd hygiene." *Journal of Dairy Science*, 97(11), 6926-6934.
- Youngerman, S. M., Saxton, A. M., Oliver, S. P., and Pighetti, G. M. (2004). "Association of CXCR2 polymorphisms with subclinical and clinical mastitis in dairy cattle." *Journal of Dairy Science*, 87(8), 2442-2448.
- Zadoks, R. N. (2007). "Sources and epidemiology of *Streptococcus uberis*, with special emphasis on mastitis in dairy cattle." *Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*, 2(30), 15.

Chapter VII: Conclusion

The objective of this dissertation was to identify genetic associations with unique mastitis phenotypes developed using data from an experimental intramammary *S. uberis* challenge. Data collected following challenge included somatic cell counts (SCC), *S. uberis* CFU, mammary scores, milk scores, and antibiotic therapy use (Siebert 2013), which were used to develop 10 novel phenotypes directly related to *S. uberis* mastitis. Association analyses for these novel phenotypes identified 17 SNPs that are within previously identified quantitative trait loci (QTLs) for mastitis or phenotypes with links to mastitis (i.e.: productive life, milk yield). Identifying associations with SNPs in regions with prior evidence linked to mastitis helps establish the validity of using our novel phenotypes in response to direct, experimental *S. uberis* challenge for GWAS. Furthermore, genomic studies with pathogen specificity have been limited (Sørensen et al. 2008), which has limited the identification of candidate loci with functions linked to individual pathogens, a gap our *S. uberis* specific study may help fill. Additionally, because we identified SNPs in regions with prior evidence for mastitis or related phenotypes that were not pathogen specific, these SNPs could also be associated with other causative pathogens. Besides SNPs in regions with prior evidence, our novel phenotypes were associated an additional 35 SNPs across the bovine genome.

Identified SNPs have provided an array of candidate genes whose functions could be linked to the variation in phenotypes observed following *S. uberis* intramammary experimental challenge. We identified a total of 23 possible candidate genes where the majority of identified candidate genes fall into one of two categories: inflammation/immune response or gene expression regulation. Once bacteria invade the mammary gland and are recognized by host cells, inflammation is initiated. Inflammatory cytokines/chemokines then initiate the immune response. The immune response is controlled by many aspects, but a key regulator is gene transcription, where bacterial recognition and pro-inflammatory cytokines induce the transcription of other genes needed for a proper immune response. Therefore, candidate genes associated with these processes are critical to a cow's ability to combat invading pathogens within the mammary gland such as *S. uberis*.

Of particular interest are two candidate genes: TIAM2 (t-cell invasion and metastasis-2) identified with the SCC cure phenotype and TRAPPC9 (trafficking protein particle complex 9) identified with the severity score phenotype. TIAM2 is a guanine nucleotide exchange factor that functions within small GTPases that has been implicated in cell invasion, migration, adhesion, and apoptosis in cancers (Chen et al. 2012; Rooney et al. 2010). Since TIAM2 was identified with the SCC cure phenotype, which grouped cows based on how long post-challenge their SCC returned to a healthy

level indicating clearance of infection, TIAM2 could play a role in ensuring cells migrate to the site of infection efficiently following infection allowing for fast resolution of infection. TRAPPC9 plays a role in phosphorylating the IKK complex which controls activation of NFκB (Hu et al. 2005). NFκB is a transcriptional activator of many genes including many that are critical for inflammation and immunity including pro-inflammatory cytokine/chemokine production and adhesion molecule expression (Bonizzi and Karin 2004). TRAPPC9 was identified as associated with the severity score phenotype which summed milk and mammary score for the first 7 days post-challenge. Cows with higher severity scores would have demonstrated more inflammation following challenge indicating a strong immune response and strong activation of NFκB. Further support of this is that a QTL for SCS was identified in this same region (Kaupe et al. 2007). Strong activation of NFκB would likely result in better recruitment of SCC to the infection site because of increased cytokine and adhesion molecule expression.

Furthermore, prior studies have identified CXCR1 as a gene important for mastitis (Rambeaud et al. 2006; Rambeaud and Pighetti 2007; Rambeaud and Pighetti 2005; Siebert 2013; Youngerman et al. 2004a; Youngerman et al. 2004b). Additionally, CXCR1 genetic background has been demonstrated to influence cow response to *S. uberis* intramammary challenge (Siebert 2013). However, observed variation in response to challenge could be attributed to other genes in LD with CXCR1. We identified the haplotype block surrounding the CXCR1 gene to identify other genes in LD with the polymorphisms in CXCR1 that are potential candidate genes. Linkage analysis identified that the CXCR1 SNPs are part of 2 separate haplotype blocks which also contain 21 candidate genes.

Of particular interest of the candidate genes identified from linkage analysis is the SLC11A1 (solute carrier family 11 – proton-coupled divalent metal ion transporter; NRAMP1) located from 107,114,929 – 107,125,854 bp on BTA 2. SLC11A1 is a member of the proton-coupled divalent metal ion transporters that transports iron, zinc, and manganese cations that is localized to late lysosomes and endosomes (Goswami et al. 2001; Gruenheid et al. 1997; Searle et al. 1998). Additionally, SLC11A1 has been implicated in disease resistance through cation control in macrophages which protects macrophages from their own ROS generation and sequesters cations that pathogens could use to produce protective enzymes as reviewed by Blackwell et al. (2001). Proper macrophage function and killing ability are particularly important for *S. uberis* infections *S. uberis* evades neutrophils, which are the primary phagocytic cell present following infection of the mammary gland (Oliver et al. 1998). Furthermore, SLC11A1 has been associated with resistance to intracellular pathogens by prohibiting their

replication and *S. uberis* has the ability to internalize into host cells (Almeida et al. 2000; Gruenheid et al. 1997; Tamilselvam et al. 2006). Thus, SLC11A1 could play a role in a cow's ability to properly clear *S. uberis* following challenge.

Lastly, we demonstrated that CXCR1 haplotype combination influences the rate of *S. uberis* growth in milk in vitro. Where VWHRR x VWHRR milk supported *S. uberis* growth the most and AWQRR x AWQRR milk supported *S. uberis* growth the least regardless of *S. uberis* strain used for inoculation. These two haplotype combinations differ only at positions 365 and 735 which are the positions demonstrated to be in LD with SLC11A1. This provides further support that SLC11A1 could be a valid candidate gene for explaining variation in responses following *S. uberis* intramammary challenge and that variation in divalent metal ions, which are controlled by SLC11A1, may influence the growth rate of *S. uberis* in milk.

Overall, we have identified many candidate genes, whose functions are primarily linked to inflammation/immunity and gene regulation. We also have identified that *S. uberis* growth rate in milk differs depending on the CXCR1 haplotype combination expressed by the cow, suggesting concentration of an inhibitory or permissive factor towards *S. uberis* growth differs with CXCR1 haplotype combination. Considered together, our results imply that variation in one of the identified candidate genes could be influencing the rate of *S. uberis* growth and thus *S. uberis* mastitis. Furthermore, using pathogen specific phenotypes like those generated here could provide an ideal method for identifying host genes involved in mediating the virulence of *S. uberis* which we also demonstrated to be an important factor in the rate of *S. uberis* growth in milk. Further investigation of identified candidate genes could lead to the establishment of more effective or better defined genetic selection methods or uncover novel preventative or treatment compounds that could be used against *S. uberis* mastitis.

References

- Almeida, R. A., Calvinho, L. F., and Oliver, S. P. (2000). "Influence of protein kinase inhibitors on *Streptococcus uberis* internalization into bovine mammary epithelial cells." *Microbial Pathogenesis*, 28(1), 9-16.
- Blackwell, J. M., Goswami, T., Evans, C. A. W., Sibthorpe, D., Papo, N., White, J. K., Searle, S., Miller, E. N., Peacock, C. S., and Mohammed, H. (2001). "SLC11A1 (formerly NRAMP1) and disease resistance." *Cellular Microbiology*, 3(12), 773-784.
- Bonizzi, G., and Karin, M. (2004). "The two NF-kappaB activation pathways and their role in innate and adaptive immunity." *Trends in Immunology*, 25(6), 280-8.
- Chen, J. S., Su, I. J., Leu, Y. W., Young, K. C., and Sun, H. S. (2012). "Expression of T-cell lymphoma invasion and metastasis 2 (TIAM2) promotes proliferation and invasion of liver cancer." *International Journal of Cancer*, 130(6), 1302-1313.
- Goswami, T., Bhattacharjee, A., Babal, P., Searle, S., Moore, E., Ming, L., and Jenefer, M. (2001). "Natural-resistance-associated macrophage protein 1 is an H⁺/bivalent cation antiporter." *Biochemical Journal*, 354(3), 511-519.
- Gruenheid, S., Pinner, E., Desjardins, M., and Gros, P. (1997). "Natural resistance to infection with intracellular pathogens: the Nramp1 protein is recruited to the membrane of the phagosome." *Journal of Experimental Medicine*, 185(4), 717-730.
- Hu, W. H., Pendergast, J. S., Mo, X. M., Brambilla, R., Bracchi-Ricard, V., Li, F., Walters, W. M., Blits, B., He, L., Schaal, S. M., and Bethea, J. R. (2005). "NIBP, a novel NIK and IKK(beta)-binding protein that enhances NF-(kappa)B activation." *Journal of Biological Chemistry*, 280(32), 29233-41.
- Kaupe, B., Brandt, H., Prinzenberg, E. M., and Erhardt, G. (2007). "Joint analysis of the influence of CYP11B1 and DGAT1 genetic variation on milk production, somatic cell score, conformation, reproduction, and productive lifespan in German Holstein cattle." *Journal of Animal Science*, 85(1), 11-21.
- Oliver, S. P., Almeida, R. A., and Calvinho, L. F. (1998). "Virulence factors of *Streptococcus uberis* isolated from cows with mastitis." *Zentralbl Veterinarmed B*, 45(8), 461-71.
- Rambeaud, M., Clift, R., and Pighetti, G. M. (2006). "Association of a bovine CXCR2 gene polymorphism with neutrophil survival and killing ability." *Veterinary Immunology and Immunopathology*, 111(3-4), 231-238.

- Rambeaud, M., and Pighetti, G. (2007). "Differential calcium signaling in dairy cows with specific CXCR1 genotypes potentially related to interleukin-8 receptor functionality." *Immunogenetics*, 59(1), 53-58.
- Rambeaud, M., and Pighetti, G. M. (2005). "Impaired neutrophil migration associated with specific bovine CXCR2 genotypes." *Infection and Immunity*, 73(8), 4955-4959.
- Rooney, C., White, G., Nazgiewicz, A., Woodcock, S. A., Anderson, K. I., Ballestrem, C., and Malliri, A. (2010). "The Rac activator STEF (Tiam2) regulates cell migration by microtubule-mediated focal adhesion disassembly." *EMBO Reports*, 11(4), 292-298.
- Searle, S., Bright, N. A., Roach, T. I., Atkinson, P. G., Barton, C. H., Meloen, R. H., and Blackwell, J. M. (1998). "Localisation of Nramp1 in macrophages: modulation with activation and infection." *Journal of Cell Science*, 111(19), 2855-2866.
- Siebert, L. J. (2013). "Bovine mammary interleukin-8 receptor expression and genetic association with *Streptococcus uberis* based mastitis."
- Sørensen, L. P., Guldbrandtsen, B., Thomasen, J. R., and Lund, M. S. (2008). "Pathogen-specific effects of quantitative trait loci affecting clinical mastitis and somatic cell count in Danish Holstein cattle." *Journal of Dairy Science*, 91(6), 2493-2500.
- Tamilselvam, B., Almeida, R. A., Dunlap, J. R., and Oliver, S. P. (2006). "*Streptococcus uberis* internalizes and persists in bovine mammary epithelial cells." *Microbial Pathogenesis*, 40(6), 279-285.
- Youngerman, S. M., Saxton, A. M., Oliver, S. P., and Pighetti, G. M. (2004a). "Association of CXCR2 polymorphisms with subclinical and clinical mastitis in dairy cattle." *Journal of Dairy Science*, 87(8), 2442-2448.
- Youngerman, S. M., Saxton, A. M., and Pighetti, G. M. (2004b). "Novel single nucleotide polymorphisms and haplotypes within the bovine CXCR2 gene." *Immunogenetics*, 56(5), 355-359.

Appendix

Table A- 1: Comprehensive list of SNPs identified using novel phenotypes with distance to neighboring SNPs and genotype frequencies indicated. Final column indicates status of the SNP following SNP pruning based on minor allele frequencies.

SNP (rs)	Locus (Chr: bp)	Dist. to 3'SNP (bp)	Dist. To 5' SNP (bp)	# Genotypes	Genotypes Freqs.	# Missing	Missing Freq.	Status
rs43273037	1: 134,069,561	38,757	20,683	GG: 7 GT: 16 TT: 0	GG: 0.30 GT: 0.70 TT: 0.0	17	0.43	Keep
rs110974478	1: 150,141,293	26,075	27,008	AA: 0 AG: 4 GG: 36	AA: 0.0 AG: 0.10 GG: 0.90	0	0.00	Keep
rs43224998	1: 34,903,570	39,073	30,946	AA: 25 AG: 14 GG: 1	AA: 0.63 AG: 0.35 GG: 0.03	0	0.00	Keep
rs43317360	2: 107,657,675	40,772	24,025	AA: 22 AC: 17 CC: 0	AA: 0.56 AC: 0.44 CC: 0.00	1	0.03	Keep
rs41644229	2: 107,681,700	24,025	24,019	CC: 17 CT: 17 TT: 1	CC: 0.49 CT: 0.49 TT: 0.02	5	0.13	Keep
rs41645172	2: 109,778,246	37,402	34,097	AA: 22 AG: 18 GG: 0	AA: 0.55 AG: 0.45 GG: 0.0	0	0.00	Keep
rs41619984	2: 132,148,018	54,001	22,137	AA: 39 AC: 1 CC: 0	AA: 0.97 AC: 0.03 CC: 0.0	0	0.00	Remove
rs110345656	2: 133,861,562	44,754	27,677	AA: 0 AG: 1 GG: 39	AA: 0.0 AG: 0.03 GG: 0.97	0	0.00	Remove
rs41637469	2: 132,404,449	118,720	22,533	AA: 0 AG: 1 GG: 37	AA: 0.0 AG: 0.03 GG: 0.97	2	0.05	Remove
rs43499732	3: 6,480,313	36,955	30,339	GG: 3 GT: 6 TT: 29	GG: 0.08 GT: 0.16 TT: 0.76	2	0.05	Keep
rs109880729	3: 27,600,325	45,685	41,976	CC: 0 CT: 2 TT: 38	CC: 0.0 CT: 0.05 TT: 0.95	0	0.00	Keep
rs109513885	3: 34,263,481	35,397	21,762	CC: 4 CT: 20 TT: 15	CC: 0.10 CT: 0.51 TT: 0.39	1	0.03	Keep
rs43709346	3: 74,836,184	91,656	61,282	AA: 39 AG: 1 GG: 0	AA: 0.97 AG: 0.03 GG: 0.0	0	0.00	Keep
rs43352862	3: 89,649,212	27,834	26,390	GG: 33 GT: 1 TT: 0	GG: 0.97 GT: 0.03 TT: 0.00	6	0.15	Remove
rs43161729	3: 91,379,427	26,464	22,825	GG: 38 GT: 1 TT: 0	GG: 0.97 GT: 0.03 TT: 0.00	1	0.03	Remove
rs41615858	3: 91,402,252	22,825	31,407	CC: 39 CG: 1 GG: 0	CC: 0.97 CG: 0.03 GG: 0.0	0	0.00	Remove
rs41677930	3: 93,543,745	20,056	45,499	CC: 39 CT: 1 TT: 0	CC: 0.97 CT: 0.03 TT: 0.0	0	0.00	Remove
rs41630751	3: 93,589,244	45,499	58,424	AA: 39 AC: 1 CC: 0	AA: 0.97 AC: 0.03 CC: 0.0	0	0.00	Remove
rs43715308	3: 115,115,735	22,332	25,008	GG: 0 GT: 1 TT: 39	GG: 0.0 GT: 0.03 TT: 0.97	0	0.00	Remove

Table A- 1: Continued

SNP (rs)	Locus (Chr: bp)	Dist. to 3'SNP (bp)	Dist. To 5' SNP (bp)	# Genotypes	Genotypes Freqs.	# Missing	Missing Freq.	Status
rs109567359	3: 116,110,288	48,493	21,672	CC: 31 CT: 1 TT: 0	CC: 0.97 CT: 0.03 TT: 0.0	8	0.20	Remove
rs41576873	4: 92,742,789	37,598	22,337	AA: 16 AG: 18 GG: 3	AA: 0.43 AG: 0.49 GG: 0.08	3	0.08	Keep
rs43423752	5: 2,925,015	79,943	43,556	CC: 39 CT: 1 TT: 0	CC: 0.97 CT: 0.03 TT: 0.0	0	0.00	Keep
rs41655875	5: 26,621,180	59,518	39,863	CC: 0 CT: 2 TT: 24	CC: 0.00 CT: 0.08 TT: 0.92	14	0.35	Keep
rs41612889	7: 107,527,276	26,324	74,705	CC: 4 CT: 18 TT: 9	CC: 0.13 CT: 0.58 TT: 0.29	9	0.23	Keep
rs110556992	7: 65,047,126	36,604	24,768	CC: 10 CT: 20 TT: 4	CC: 0.29 CT: 0.59 TT: 0.12	6	0.15	Keep
rs109792349	8: 8,618,532	78,759	33,274	AA: 17 AT: 16 TT: 5	AA: 0.45 AT: 0.42 TT: 0.13	2	0.05	Keep
rs42488893	8: 19,967,642	26,627	51,187	CC: 12 CT: 16 TT: 1	CC: 0.41 CT: 0.55 TT: 0.04	11	0.28	Keep
rs41665257	9: 4,581,947	33,954	33,076	CC: 0 CT: 3 TT: 32	CC: 0.0 CT: 0.09 TT: 0.91	5	0.13	Keep
rs41594141	9: 73,368,179	27,358	49,534	AA: 14 AG: 22 GG: 2	AA: 0.37 AG: 0.58 GG: 0.05	2	0.05	Keep
rs109484182	9: 93,183,596	39,525	30,067	AA: 9 AG: 18 GG: 8	AA: 0.26 AG: 0.51 GG: 0.23	5	0.13	Keep
rs110216702	9: 100,163,987	25,797	140,058	AA: 27 AG: 12 GG: 0	AA: 0.69 AG: 0.31 GG: 0.00	1	0.03	Keep
rs41903759	10: 68,693,555	25,657	28,870	GG: 0 GT: 1 TT: 37	GG: 0.0 GT: 0.03 TT: 0.97	2	0.05	Remove
rs109629398	10: 74,011,076	31,092	21,743	AA: 39 AG: 1 GG: 0	AA: 0.98 AG: 0.02 GG: 0.0	0	0.00	Remove
rs41610542	11: 7,625,345	39,598	46,097	AA: 28 AG: 10 GG: 1	AA: 0.72 AG: 0.26 GG: 0.02	1	0.03	Keep
rs110866336	12: 77,746,796	51,412	69,844	AA: 28 AG: 11 GG: 0	AA: 0.72 AG: 0.28 GG: 0.0	1	0.03	Keep
rs109788678	13: 11,784,421	51,340	47,059	CC: 22 CT: 11 TT: 5	CC: 0.58 CT: 0.29 TT: 0.13	2	0.05	Keep
rs41565622	13: 24,688,809	43,157	43,307	AA: 9 AG: 25 GG: 5	AA: 0.23 AG: 0.64 GG: 0.13	1	0.03	Keep
rs110804434	13: 78,901,415	43,289	26,890	AA: 0 AG: 7 GG: 3	AA: 0.0 AG: 0.70 GG: 0.30	30	0.75	Keep
rs110199901	14: 2,524,432	32,876	29,093	AA: 10 AG: 17 GG: 12	AA: 0.25 AG: 0.44 GG: 0.31	1	0.03	Keep

Table A- 1: Continued

SNP (rs)	Locus (Chr: bp)	Dist. to 3'SNP (bp)	Dist. To 5' SNP (bp)	# Genotypes	Genotypes Freqs.	# Missing	Missing Freq.	Status
rs55617160	14: 4,468,478	30,211	25,426	AA: 3 AG: 18 GG: 14	AA: 0.09 AG: 0.51 GG: 0.40	5	0.13	Keep
rs109427689	15: 25,622,211	86,644	47,186	AA: 4 AC: 25 CC: 4	AA: 0.11 AC: 0.78 CC: 0.11	7	0.18	Keep
rs41576666	16: 9,941,492	57,705	39,509	CC: 19 CT: 12 TT: 8	CC: 0.49 CT: 0.31 TT: 0.20	1	0.03	Keep
rs41660125	16: 10,389,306	30,896	36,406	AA: 5 AG: 20 GG: 11	AA: 0.14 AG: 0.56 GG: 0.30	4	0.10	Keep
rs41596554	16: 10,740,538	37,074	30,430	GG: 7 GT: 18 TT: 7	GG: 0.22 GT: 0.56 TT: 0.22	8	0.20	Keep
rs41803475	16: 34,994,367	30,219	29,812	CC: 39 CT: 1 TT: 0	CC: 0.97 CT: 0.03 TT: 0.0	0	0.00	Remove
rs29019936	17: 562,366	86,303	85,971	CC: 2 CT: 26 TT: 8	CC: 0.06 CT: 0.72 TT: 0.22	4	0.10	Keep
rs41831243	17: 5,221,649	30,026	25,007	CC: 0 CT: 3 TT: 32	CC: 0.0 CT: 0.09 TT: 0.91	5	0.13	Keep
rs41570578	17: 64,142,204	92,914	22,597	GG: 20 GT: 10 TT: 3	GG: 0.60 GT: 0.31 TT: 0.09	7	0.18	Keep
rs29015249	18: 24,862,274	79,879	35,464	CC: 1 CT: 20 TT: 17	CC: 0.03 CT: 0.52 TT: 0.45	2	0.05	Keep
rs109290019	18: 53,260,898	36,260	21,033	AA: 14 AG: 18 GG: 6	AA: 0.37 AG: 0.47 GG: 0.16	2	0.05	Keep
rs109070561	18: 53,281,931	21,033	26,408	AA: 6 AG: 20 GG: 11	AA: 0.16 AG: 0.54 GG: 0.30	3	0.08	Keep
rs109243453	20: 58,928,800	63,912	22,186	AA: 6 AG: 16 GG: 17	AA: 0.15 AG: 0.41 GG: 0.44	1	0.03	Keep
rs41951389	20: 59,627,852	28,119	36,539	CC: 6 CT: 19 TT: 14	CC: 0.15 CT: 0.49 TT: 0.36	1	0.03	Keep
rs109961240	20: 63,407,185	27,332	27,944	AA: 3 AG: 14 GG: 22	AA: 0.08 AG: 0.36 GG: 0.56	1	0.03	Keep
rs111021268	21: 23,530,389	44,067	81,952	CC: 38 CT: 2 TT: 0	CC: 0.95 CT: 0.05 TT: 0.0	0	0.00	Keep
rs42651398	21: 39,716,280	56,906	42,242	CC: 0 CT: 27 TT: 1	CC: 0.0 CT: 0.96 TT: 0.04	12	0.30	Keep
rs109417302	21: 58,323,035	23,590	86,031	GG: 19 GT: 1 TT: 0	GG: 0.95 GT: 0.05 TT: 0.0	20	0.50	Remove
rs110942101	22: 7,385,040	71,997	35,676	GG: 0 GT: 3 TT: 37	GG: 0.0 GT: 0.08 TT: 0.92	0	0.00	Keep
rs109747337	22: 58,790,629	42,027	25,750	AA: 37 AG: 3 GG: 0	AA: 0.93 AG: 0.07 GG: 0.0	0	0.00	Keep

Table A- 1: Continued

SNP (rs)	Locus (Chr: bp)	Dist. to 3'SNP (bp)	Dist. To 5' SNP (bp)	# Genotypes	Genotypes Freqs.	# Missing	Missing Freq.	Status
rs109473352	22: 59,325,992	48,161	33,959	CC: 0 CT: 1 TT: 39	CC: 0.0 CT: 0.03 TT: 0.97	0	0.00	Remove
rs110693161	24: 48,872,488	87,497	27,307	AA: 2 AG: 7 GG: 30	AA: 0.05 AG: 0.18 GG: 0.77	1	0.03	Keep
rs110747207	26: 50,933,887	56,392	19,618	AA: 0 AG: 8 GG: 31	AA: 0.0 AG: 0.21 GG: 0.79	1	0.03	Keep
rs110599808	28: 247,173	37,280	61,632	AA: 4 AG: 23 GG: 13	AA: 0.10 AG: 0.58 GG: 0.32	0	0.00	Keep
rs41612729	28: 5,052,476	46,198	70,546	AA: 12 AG: 23 GG: 4	AA: 0.31 AG: 0.59 GG: 0.10	1	0.03	Keep
rs110196308	28: 7,899,978	21,253	40,207	AA: 0 AG: 1 GG: 39	AA: 0.0 AG: 0.03 GG: 0.97	0	0.00	Remove
rs41653441	28: 28,149,081	50,031	36,613	CC: 8 CT: 18 TT: 7	CC: 0.24 CT: 0.55 TT: 0.21	7	0.18	Keep

Vita

Lydia Jean Siebert was born on July 13, 1988, in Baltimore, Maryland. She graduated class salutatorian from Harford Technical High School in Bel Air, Maryland, in 2006. After high school she attended the University of Tennessee, Knoxville, where she majored in microbiology and earned her Bachelor of Science degree, *cum laude* in Biology in December of 2010. She pursued her graduate education under the guidance of Dr. Gina Pighetti in Animal Science at the University of Tennessee, Knoxville. Her Master of Science degree entitled “Bovine mammary interleukin-8 receptor expression and genetic association with *Streptococcus uberis* based mastitis” was completed in December 2013. She continued to focus on mastitis for her Doctorate degree and completed her dissertation titled “Identifying associations with unique mastitis phenotypes in response to intramammary *Streptococcus uberis* challenge” in May of 2017.